EUROPEAN CENTRAL BANK

EUROSYSTEM

MACROECONOMIC VULNERABILITY AND DISAGREEMENT IN EXPECTATIONS

by Cristian Badarinza
and Marco Buchmann

EUROPEAN CENTRAL BANK
EUROSYSTEM

MACROECONOMIC VULNERABILITY AND DISAGREEMENT IN EXPECTATIONS

by Cristian Badarinza[1]
and Marco Buchmann[2]

In 2011 all ECB publications feature a motif taken from the €100 banknote.

1   Goethe University Frankfurt, Department of Money and Macroeconomics, House of Finance,Grueneburgplatz 1,D-60323 Frankfurt am Main, Germany; e-mail: badarinza@wiwi.uni-frankfurt.de
2   European Central Bank, Kaiserstrasse 29, 60311 Frankfurt am Main, Germany; e-mail: marco.buchmann@ecb.europa.eu

CONTENTS

Abstract

In this paper we discuss the role of the cross-sectional heterogeneity of beliefs in the context of understanding and assessing macroeconomic vulnerability. Emphasis lies on the potential of changing levels of disagreement in expectations to influence the propensity of the economy to switch between different regimes, a hypothesis that finds robust empirical support from a regime-switching model with endogenous transition probabilities for output growth and realized stock market volatility in the US.

Keywords: Heterogeneous beliefs, business cycles, regime-switching, forecasting, endogenous transition probabilities

JEL classification: C53, D8, E32

Non-technical summary

The purpose of this paper is to examine whether changing levels of aggregate disagreement have the potential to render the economy more or less vulnerable to exogenous shocks.

With *agreement*, disagreement respectively, we refer to a measure of aggregate dispersion in agents' beliefs about the expected course of the real economy in the near future. The underlying expectation may for example be about changes in employment conditions or economy-wide output.

Our hypothesis involves the assumption that changing levels of agreement do not, as such, *cause* the economy to switch between high-, medium-, or low-growth regimes but instead have the potential to impact the *propensity of the economy to switch* regimes. We argue that higher levels of agreement increase the propensity to switch from high-growth to medium- or low-growth regimes. For an actual downturn to occur, we still have to await the arrival of exogenous shocks.

We propose a simple and transparent measure of disagreement that we derive from survey data (Michigan survey) on expected real economic activity. Troughs in disagreement, i.e. pronounced high levels of *agreement*, tend to appear prior to all recession periods dated since 1960 in the US, including the most recent world-wide recession of 2007-2009.

Based on a three-regime endogenous switching model for GDP growth in the US, we provide robust evidence in favor of our hypothesis: The estimated transition probability of moving from normal growth to recession periods is about three times as large when agreement is high (15% vs. 5% with high agreement, disagreement respectively). Higher levels of disagreement also appear to let a sustainable growth regime last longer, with the estimated probability of continuance in medium, positive growth equaling 57% and 92% when there is high agreement, disagreement respectively.

We also argue that the disagreement metric can be a meaningful indicator for the purpose of measuring and assessing *financial stability*. As of yet, focus has been on interest rates, interest rate spreads, various measures of volatility, and so forth, which

may well be able to signal *stress* in financial markets - stress having arisen, however, means that the market was vulnerable already before the arrival of shocks, and should therefore not be seen as measures of financial market fragility, vulnerability respectively. Disagreement measures are different in nature since they tend to signal *tranquility* prior to a potential downturn.

In addition to the in-sample estimation and inference of regimes, we provide an out-of-sample forecast evaluation of the various models. For GDP growth, both the regime-switching mechanism as well as the further incorporation of disagreement to condition the transition probabilities result in significant gains with regard to point and directional forecast accuracy up to a four-quarter horizon. Compared to a benchmark model forecast, GDP forecast errors can be reduced by up to 18% over the eight-year test period (2002Q1-2010Q3), and around 24% in the period covering the recession and the remainder of the test sample (2007Q4-2010Q3). While, on the other hand, allowing for regime-switching in the models for realized volatility allows to explain a significantly larger portion of its variation in-sample, the gain in precision out-of-sample remains rather insignificant.

Along with the empirical analysis to substantiate that disagreement levels are a useful vulnerability measure, the aim of the paper is to embed the underlying hypothesis in the theoretical literature on heterogeneity and leverage, the social value of public information, the conceptual relation between disagreement and uncertainty, as well as in the fields of forecasting in general, and early-warning systems in particular. A detailed literature survey from the various fields serves to establish a conceptual framework for us to think broadly about the role of aggregate belief heterogeneity.

1 Introduction

The dramatic unfolding of events in the autumn of 2008 revealed deeply rooted sources of systemic vulnerability within a financial system which had actually been praised for years to be itself an essential determinant of risk containment and stability. As it became apparent that the main source of the credit and liquidity crunches which brought the economy to the slippery slope lay in the balance sheet interlinkages which defined the system and not only in the risk management practices of individual agents, the need for a systemic approach to supervision and regulation quickly came to dominate the public discourse. Policy makers and regulatory authorities in the sequel strive to put in place a system of supervisory bodies with the task of recognizing and containing systemic risks. In this paper we attempt to contribute to this ongoing effort towards defining the limits of a conceptual framework within which one can think about system-wide vulnerability. Building on the theoretical advances which we review in Section 2, our goal is to highlight the fact that cross-sectional survey measures of belief heterogeneity could prove useful in terms of assessing the resilience of the economy as a whole.[1]

We aim to advance and substantiate empirically the hypothesis that levels of aggregate disagreement have the potential to impact the economy's vulnerability to exogenous shocks. Our hypothesis thus involves the assumption that changing levels of agreement do not, as such, *cause* the economy to switch between high-, medium-, or low-growth regimes but instead influence the *propensity of the economy to switch* between different regimes. The results suggests that higher levels of agreement about the future course of the economy make the transition from welfare-superior regimes to welfare-inferior ones more likely.

Appropriate measures of disagreement could thus also be meaningful indicators for the purpose of assessing financial stability. As of yet, focus has been on interest rates, interest rate spreads, various measures of volatility, and so forth, which may well be able to signal stress in financial markets - stress having arisen, however, means that the market was vulnerable already prior to the arrival of a shock. Disagreement measures are different in nature since they tend to signal tranquility prior

---

[1]The topic is also subject of interest in ECB (2011).

to a potential downturn, while this calmness can be reflective of high vulnerability.

The data and econometric methodology are described in Section 3. We propose a simple and transparent measure of disagreement that can be derived from survey data (the Michigan Survey of Consumers) on expected real economic activity. For quantifying disagreement we employ a measure of the cross-sectional dispersion in agents' beliefs about the expected course of the real economy (with regard to employment or business conditions) in the near future.

We estimate a three-regime endogenous switching model in Section 4, for two cases: US GDP growth and realized volatility in US stock markets. The transition probabilities that are associated with switching between the three regimes are allowed to be a function of economy-wide disagreement. Based on the estimated models, we then simulate the US economy under the counterfactual assumptions of constant high and constant low disagreement in the long-run, so as to shed further light on differences in distributional characteristics of GDP growth and realized stock market volatility conditional on the prevailing levels of belief heterogeneity. Moreover, we provide an evaluation of simulated out-of-sample predictions in order to assess how both the regime-switching as well as the disagreement measure to condition the transition process help improve forecast accuracy.

2   Literature review

2.1   Heterogeneity, leverage and volatility

Economic agents, though acting entirely rational in a stochastic environment, may still be able to process information only at a limited rate, therefore differing in their probability distributions. Indeed, especially since detailed household survey data has become available for most developed countries, the connection between the heterogeneity in information sets, expectations, financial sophistication, literacy or analytical ability and the market-relevant behavior concerning financial planning, consumption or labor allocation is increasingly under scrutiny. Our primary concern in this section is to present a series of ideas and arguments pertaining to these effects.

Early in the development of modern finance, it has been noted that the heterogeneity in beliefs across market participants can have important effects on prices because, for example in the presence of short-sales constraints, the price of the asset reflects the valuation of the most optimistic buyers and thus disagreement is associated with higher prices, speculation and increased volatility (Miller (1977) and Harrison and Kreps (1978)). Since then, a series of papers have analyzed the implications of disagreement on financial variables in more detail (see e.g. Scheinkman and Xiong (2003)). Nimark (2010) is part of the recent efforts to complement empirical evidence. He finds substantial evidence for heterogeneous beliefs and speculative behavior across the term structure of US bond yields and shows that even if the estimated amount of heterogeneity is small, speculative trade driven by the individually perceived mispricing of bonds can be quantitatively important. Yu (2011) arrives at similar conclusions by analyzing individual-stock analyst forecast dispersion.

The role of heterogeneous beliefs has also been advanced in terms of explaining the evolutions which led to the recent financial and economic crisis. Sims (2009) noted first that differences of opinion about the future course of inflation generates over-investment in the real asset, thus implying that monetary policy could contribute to the appearance of bubble-like phenomena by acting or communicating in such a way as to cause divergences in beliefs. The same idea is applied to housing market dynamics by Tomura (2009), who shows that boom-bust cycles as the one the US has experienced during the 2000-2010 decade in housing markets, can occur endogenously if credit-constrained borrowers become overly optimistic and this expectation is neither shared by lenders, nor realized ex post. Favara and Song (2009) note that the nature of the housing stock and the possibility of renting amounts to a de facto short-sale constraint. So, considering this non-linearity in the demand for housing, the equilibrium house prices then reflect only the expectations of optimistic agents and are consistently biased upwards.

The analysis of Tian and Yan (2009) concerns the financial roots of the crisis and the role of heterogeneous beliefs in this context: "[T]he market for forward-type securities is rather resilient in the sense that, for a reasonable range of parameters, there is almost always a seller of the security, especially if beliefs are diverse. The market for option-type contracts, however, appears to be more vulnerable in that the

supply of the security may dry up quickly when the underlying asset value experiences a drastic shock (p. 2)". Compared to the previously mentioned setups in which information flows were seen to be constant, Tian and Yan (2009) are among the first to ask about the effects of time-variation in expectations. First of all, in their framework it still holds that disagreement generates trade, but an additional second element enters in that a subsequent alignment of information sets, coupled with the presence of negative shocks may cause the market to freeze, break down and then resume as the fundamental conditions slowly improve. As quoted above, this effect is shown to be of acute relevance in particular for tranched securities like the ones that were intensively traded before the crisis.

The Geanakoplos (2009) view on the crisis as an end to a leverage cycle is also connected along similar lines with the presence of disagreement. The crucial element of his analysis is the fact that optimists bid the prices up in good times and bear all losses in the event of a crash, with leveraged buyers forced to sell in order to meet margin requirements, which depresses asset prices and tightens collateral requirements even further. In addition to this basic mechanism through which disagreement affects the economy, Geanakoplos (2009) opens the door for two more ideas which we will be concerned with in the next section. First, he notes that: "The point of the leverage cycle is that excess leverage followed by excessive deleveraging will cause a crash even before there has been a crash in the fundamentals, and even if there is no subsequent crash in the fundamentals (p. 26)." This connects with our focus on the notion of *vulnerability* in this paper, as opposed to systemic risk or financial stress. What matters to us are these expectational *ex ante* effects, namely the inherent vulnerability of the system with respect to the updating of information sets, the revision of beliefs and the consequent market-relevant actions. Second, he writes: "The idea is that at the beginning, everyone thinks the chances of ultimate failure require too many things to go wrong to be of any substantial probability. There is little uncertainty, and therefore little room for disagreement. Once enough things go wrong to raise the spectre of real trouble, the uncertainty goes way up in everyone's mind, and so does the possibility of disagreement (p. 7)." As we will show in the context of our empirical exercise, this way of thinking about news and coordinated actions is well matched by observed data.

2.2 Social value of public information

Since the seminal contribution of Hirshleifer (1971), the literature on the social value of public information has grown considerably. In recent years, in particular the welfare-reducing properties of public information have attracted research interest (see e.g. Morris and Shin (2002) and Angeletos and Pavan (2007)). While not focusing directly on disagreement, these models complement the discussion from the previous subsection with a decisive insight: If agents overweigh the public signal in their expectation formation process, this may lower the incentive to exert effort in collecting costly private information, thereby increasing the chance of individual investment mistakes and reducing social welfare. The collateral effect of this herding on public information is to decrease the cross-sectional dispersion of beliefs, which does not per se have an impact upon equilibrium dynamics, but is an indication of sub-optimal behavior.

In this regard, of particular importance is the contribution by Veldkamp (2006), who builds on the benchmark model of Grossman and Stiglitz (1980) and shows that information markets are prone to frenzies and herds as a direct consequence of the non-rival nature of information. Her main insight is that an increase in demand for a particular piece of information in a competitive market does not cause a rise in the market price like in the case of limited-supply assets, but instead the information gets traded at even lower prices, the more the desirability of this information gets acknowledged by the information providers. Agents thus end up buying the same information that others are buying - just like in the class of models mentioned above, except that in this case the public or private nature of any particular bit of information is determined endogenously. Again, disagreement decreases as more information of a particular sort becomes available, thereby indicating (but not implying, since the causality is not bi-directional) a possible herding tendency due to the abundance of this particular low-cost informational content and resulting in gradual booms and sudden crashes.

Gala and Volpin (2010) analyze the social value of information in a financial setup in which individual investment decisions impose a negative externality on others, due to the limited availability of capital and the need for refinancing. They highlight

the possible reduction in social welfare which is a result of the provision of public (low-cost) information. If debtors do not internalize the equilibrium effects that their choices have on others and competitive financiers are better off financing all projects, since they earn ex post rents in the event of capital shortages and costly refinancing, then a higher correlation in the ex ante perceived financing needs of agents may result in a reduction of overall resources available for refinancing: "Public information, while acting as *information equalizer* which reduces any information gaps among entrepreneurs, directs all entrepreneurs towards the same action and, thereby, may trigger systemic liquidity shortages (p. 24)."

We conclude this subsection by referring to an additional strand of the literature that we deem relevant for interpreting our empirical findings: The connection between diversity, innovation and technological progress (Akcigit and Kerr (2010) and Acemoglu (2011)). A well established fact in the economics of innovation is the heterogeneity of R&D choices in terms of creating new products and opening new markets on one side (exploration) as well as improving and extending current product lines or market structures (exploitation) on the other, with the exploratory activities of entrants and small firms generating spillover effects and thereby stimulating economic growth more than proportionally (Akcigit and Kerr (2010)). Acemoglu (2011) shows that the diversity of research patterns is one of the factors contributing to more activities of this exploratory type. He also shows that there are natural market mechanisms (entry costs, profit maximization, short-term orientation of management) leading to too little research diversity which is detrimental to social welfare. However, he also identifies a potential remedy to the problem by suggesting that "a counteracting force against the potential lack of diversity in research may be the diversity of researchers: because of different competences, beliefs or preferences, researchers may choose to direct their research towards areas that are under-explored by others and this may partially redress the inefficiently low level of diversity of research in the market economy (p. 1)." In other words, the market may often work such as to reward too much conformity and too little diversity, with all or the majority of R&D activities devoted to developing the same research agenda - while higher diversity of researchers, for example with respect to their interests and competences, can induce a more diverse research portfolio and thus foster economic growth. We

use these ideas somewhat detached from their initial breeding ground in order to understand the influence of heterogeneous beliefs on the probability to switch between different macroeconomic or financial regimes, while the point remains however the same: learning from the success of others (herding or informational herding) can make economic activities (or investments or insurance schemes) more concentrated and thus more vulnerable, while heterogeneity can spur innovation, exploration and growth.

2.3   Disagreement and uncertainty

It has become standard in economic theory to consider as a base unit of analysis the dynamic decision of agents acting in a stochastic environment. The perceived uncertainty associated with future realizations of the stochastic processes is then the decisive element in terms of understanding the determinants of market prices and the policy trade-offs. Ex post, one can then always measure the realized volatility of a random variable by using in-sample statistical analysis. However, the assumption that ex ante uncertainty can be proxied by ex post volatility is quite heroic. In response to this problem, a wide series of measures of ex ante uncertainty have been proposed, by referring to either market data such as futures or spreads or survey data. For our purposes, it is particularly relevant that a close connection between belief heterogeneity and perceived uncertainty has been documented.

Zarnowitz and Lambros (1987) found evidence for this positive relationship by using the NBER-ASA survey; Bomberger (1996) confirmed the findings by using the Livingston survey - Rich and Butler (1998) and Bomberger (1999) are examples of the debate ensueing in the literature in this regard - with Giordani and Soderlind (2003) and Lahiri and Sheng (2010) recently refining our understanding of both the theoretical perspective and the empirical evidence. Pfajfar and Zakelj (2011) complement the discussion by providing experimental evidence.

Overall, most studies find that about 50% of the variation in uncertainty is explained by appropriate measures of disagreement; thus we conclude that disagreement is an imperfect yet useful proxy for ex ante perceived uncertainty. We interpret, at least up to a certain degree, periods of high disagreement to be associated with high

perceived uncertainty and periods of low disagreement to reflect a perception of the future as rather stable.

2.4   Forecasting and early warning systems

Forecasting business cycle dynamics from reduced form econometric models has a long tradition, a very wide field in the literature to which Stock and Watson (2003) is a useful entry point. A subset of the forecasting literature is the field about *early warning systems*, which are models that aim to provide estimates of the probability of a binary event, such as of a crisis in foreign exchange markets, the banking system, or at the micro level the bankruptcy of individual firms or financial institutions. Probit and logit, as well as discriminant analyses are the dominant econometric approaches in this field. Cox-proportional hazard models (see e.g. Cox (1972) and Whalem (1991)) have the advantage that besides providing a measure of the probability of a crisis event, they also allow an assessment as to its timing, while only weak assumptions have to be imposed on the distribution of the time to failure. Additional benchmark contributions to the early-warning literature are Flood and Marion (1999), Borio and Lowe (2002), Worrell (2004), Davis and Karim (2008), and more recently Almanides (2011).

The conclusion from surveying the literature on forecasting and early-warning systems is that measures of cross-sectional heterogeneity in beliefs are virtually absent from the analysis; when included however, they prove to have rather strong predictive capacity and valuable informational content.

Buraschi and Whelan (2010) show that Blue Chip survey measures of belief heterogeneity are positively correlated with bond term premia and are able to forecast bond returns, with the marginal forecasting contribution proving to be very robust after controlling for a range of risk and liquidity measures. Their conclusion is that disagreement contains important information about expected returns that is not captured in historical levels of macro or financial variables.

Legerstee and Franses (2010) also find the cross-sectional distribution of beliefs to contain useful information for forecasting, in particular when used in Markov-

switching models. They use data from the Survey of Professional Forecasters and they are also to the best of our knowledge the only ones having documented empirically that the degree of disagreement could signal upcoming structural or temporal changes in an economic process.

Finally, Buchmann (2010) uses data from the European Commission Business and Consumer survey and shows that disagreement can well compete with a wide set of other possible explanatory variables (macroeconomic and financial market variables) when it comes to forecasting real activity in the euro area. A supplementary analysis that aims to extract probabilities of the 2009 recession from a pre-crisis perspective proves that the very low level of disagreement just before the advent of the financial crisis (a level that was observed to that extent the last time in May 1993) leads the model to predict a strong downturn for the years ahead, at a time when conditions in financial markets were still relatively calm. The evaluation of simulated out-of-sample forecasts based on the models presented in the present paper will confirm that disagreement measures contain useful information for predicting real activity also for the US economy.

In the present paper, we propose to employ measures of belief heterogeneity as additional tools in the broad analytical framework used to assess financial and economic vulnerability. The underlying rationale for why higher levels of agreement, i.e. less heterogeneous beliefs, shall elevate the vulnerability of the economy as a whole is that economic agents tend to effect insurance against possible future shocks in a more homogeneous fashion the more they agree about the future course of the economy. As an example, say the vast majority of the population was considering a severe downturn in the near future rather unlikely (maybe because we were experiencing a prolonged period of robust positive growth) - precautionary savings to ensure against a sudden fall in real income, e.g. as a result of loss of employment, would in the sequel be more homogeneously chosen, compared to times when such events are perceived more likely to occur. Should an exogenous shock then materialize, a relatively higher fraction of agents will experience a relatively more pronounced loss in real income. Aggregate activity would possibly start to deteriorate quite sharply, more so compared to when there had been more disagreement, and hence more diverse insurance, before the arrival of the shock. An exogenous shock might e.g. come as

a sudden, higher than expected rise in interest rates so that the re-financing of debt becomes more costly, which some agents may not be able to afford any longer so that they might have to declare bankruptcy.

On the other hand, as we have mentioned in the previous subsection, disagreement measures contain information which seem to go beyond those contained in macroeconomic aggregates. Thus, our point is not only that cross-sectional belief heterogeneity can determine cross-sectional variation in insurance mechanisms, but that the structure of beliefs per se has the potential to impact macroeconomic outcomes, through generating e.g. correlated belief revisions and liquidity shortages.

3   Data and econometric methodology

3.1   Data

As a measure of real activity we employ annual growth in quarterly real GDP in the US covering the period from 1978Q1 - 2010Q3 (131 observations). The second main ingredient is a measure of dispersion in expected changes in real activity which we derive from selected questions contained in the Michigan Surveys of Consumer Attitudes (SCA).[2]

The benchmark model estimates for output growth to be presented in Section 4 are based on Question 17 from the *business condition* section of the survey which is phrased as follows:

How about people out of work during the coming 12 months – do you think that there will be more unemployment than now, about the same or less?[3]

---

[2]The SCA is conducted by the Survey Research Center at the University of Michigan. The survey is monthly, including at least 500 interviews, and contains about 50 questions addressing in particular the consumers' own financial situation as well as their views on the outlook for the economy as a whole in both the near future and the long term. We convert the monthly survey data to quarterly, which is the frequency with which we set up and estimate all models. Other sources for expectations, such as the Survey of Professional Forecasters (SPF) are not considered in this paper. The justification for using consumers' instead of professional forecasters' expectations as a reference for the analysis is that we consider the consumer expectations more representative of the population as a whole, and therefore more relevant for determining macroeconomic outcomes.

[3]The SCA also contains a question related to expectations at the 5-year horizon. However, we

Answers from among which the respondent can choose include 1) more unemployment, 2) about the same, 3) less unemployment, or 4) do not know. From micro-level data we first compute the number of respondents in each category, excluding however the fourth category.[4]

Since the answer to the question is categorical in nature, we do not compute a plain standard deviation for obtaining a measure of disagreement since we would have to assume that the distances between answer categories be quantifiable and equal. We refrain from assuming that answer categories are equally far from one another and employ an ordinal measure of dispersion:[5]

$$d_t = \sum_{i=1}^{2} F_t^i \left(1 - F_t^i \right) \tag{1}$$

where the F_t^i are the cumulative relative frequencies for the i-th category at time t. Note that i runs only up to 2 because the cumulative frequency including the third category equals 1 and does therefore not contain additional information about the distribution of the response shares. We refer the reader to Lacy (2006) for further details about this ordinal dispersion measure.

A time series plot of GDP growth along with the disagreement measure and NBER dated recessions can be found in Figure 1. All five NBER recession periods in the sample have been preceded by a continual rise in agreement. The disagreement index falls to a local trough in 1979Q3 ($d = 0.2$), with the peak of the business cycle following in 1980Q1. The expansion period between the trough in 1980Q3 and 1981Q4 is associated with the index falling from 0.90 down to 0.54. All subsequent expansion periods between 1982Q4 - 1990Q3, 1991Q1 - 2001Q1, and until the outbreak of the most recent recession period in 2007Q4 are accompanied by pronounced upward trending levels of agreement. During the intermediate contraction periods,

---

consider the shorter (here 1-year horizon) the more relevant planning horizon for households and therefore a more meaningful measure for the analysis that we aim to conduct.

[4]Answer shares from the remaining three categories are re-scaled so that they sum to 1 in every period.

[5]We use the same measure in Badarinza and Buchmann (2009), where the ordinal dispersion is also compared to alternative standard deviation type measures that are based on *quantified* survey expectations.

there is a clear tendency for more disagreement to build up.

To substantiate the hypothesis that varying levels of disagreement can as well be reflective of vulnerability of financial markets, we will estimate a second regime-switching model, with the target variable being a measure of *realized volatility* in US stock markets. The measure is derived from daily S&P500 index data covering the period from 03/01/1978 - 31/03/2010 (8122 observations), from which realized volatilities are computed using a 1-quarter window of compounded returns. The resulting quarterly sample covers the same period as the GDP model (131 observations).[6] Figure 1 (lower panel) shows the realized volatility measure along with the measure of disagreement in expectations which is, in this case, derived from an alternative question (Q13) from the SCA:

> *How about a year from now, do you expect that in the country as a whole, business conditions will be better, or worse than they are at present, or just about the same?*

As a benchmark, the business condition related question is in our view better suited for the analysis in the context of financial market developments.

NBER recession periods tend to be associated with rising volatility in stock markets. An exception is the fourth quarter of 1987, when volatility appeared to jump significantly without having been accompanied by a downturn in real activity.[7]

An alternative measure that has been used to measure disagreement in expectations is the Herfindahl index, see e.g. Li and Li (2010). The negative Herfindahl indices based on the two survey questions Q13 and Q17 are correlated with the Lacy-type disagreement measures by 99.2% and 93.9%, respectively. Model results presented in Section 4 remain robust when replacing the Lacy-based by the Herfindahl-based disagreement measures.

---

[6]The rationale for estimating the realized volatility model with a quarterly frequency is that we consider it the business cycle frequency. It has also been chosen to have a model structure that is comparable to the models for GDP growth presented in the paper.

[7]On 19 October 1987, stock markets around the world crashed, starting in Hong Kong and then spreading west-wards through European markets to the US. The S&P 500 dropped by 20.4%; other indices, alike, fell by amounts exceeding historic records. Soon after the crash, markets started to rally again, with gains that had not previously been observed within a single day. The fact that the collapse of markets in the US was not to trigger a recession has been attributed, by and large, to the FED's policy.

3.2 Regime-switching models with endogenous transition probabilities

For the purpose of assessing the extent to which the level of agreement has an impact on the propensity of the economy to switch regimes, we employ a regime-switching (RS) model approach, that is, we allow the model dynamics that govern the evolution of y_t to depend on a discrete-valued, latent regime process s_t.

$$y_t = \beta_{0,s_t} + \sum_{p=1}^{P} \beta_{p,s_t} y_{t-p} + \epsilon_t \tag{2}$$

where β_{0,s_t} is a constant, β_{p,s_t} are up to P autoregressive coefficients, and $\epsilon_t \sim N\left(0, \sigma_{s_t}^2\right)$. The regime indicator s_t can assume integer values between $1, ..., R$ where R is the number of regimes that the intercept, the autoregressive dynamics and the residual variances are allowed to switch between.

For the latent s_t we assume the following extended Markov-process

$$P\left(s_{t+1} = k | s_t = j, Z_t\right) = p_{jk,t} \tag{3}$$

The transition probability $p_{jk,t}$ is time-varying as a result of its endogenous dependence on a vector-valued $\mathbf{Z_t}$, which shall contain our measure of disagreement, an intercept, and more conditioning variables if so desired. We follow Diebold *et al.* (1994) and employ a logit model scheme for parameterizing the transition process, while however generalizing it to the R-regime case, that is,

$$p_{jk,t}\left(\mathbf{Z}_t, \gamma\right) = \frac{\exp\left(\mathbf{Z}_t \gamma_{jk}\right)}{\sum_{k=1}^{R} \exp\left(\mathbf{Z}_t \gamma_{jk}\right)} \tag{4}$$

where γ_{jk} is a vector of parameters per each cell (j, k) in the transition matrix, capturing the dependence of $p_{jk,t}$ on $\mathbf{Z_t}$.

For the purpose of estimation of the RS model, we employ the Expectation-Maximization (EM) algorithm, as outlined in Diebold *et al.* (1994), yet developing it

for the generic R-regime case and embedding a direct numerical optimization for the sake of solving the first-order conditions to derive estimates for γ.[8] The EM algorithm involves two iterative procedures to obtain *filtered* and one-step-ahead filtered probabilities (via a forward iteration), as well as a series of *smoothed* probabilities (via a backward iteration) that let us infer with what probability the economy prevails in each of the R regimes throughout the sample period.[9]

4   Empirical findings

Model estimates that we present in the following subsections will involve the $\mathbf{Z_t}$ comprising either a raw measure of disagreement, as defined in equation (1), or a *filtered* version thereof. The filtered disagreement measure is a residual from a regression of the raw disagreement series d_t on a comprehensive set of variables capturing macro and financial conditions in the US economy, including e.g. GDP growth, private investment growth, a personal savings rate, a measure of capacity utilization, and others. See Table 1 for the results from two auxiliary regressions, respectively for disagreement based on Question 13 and 17 from the Michigan Survey.

The rationale for employing a residual from such filter equations as an alternative disagreement metric is that it shall be free of information that other indicators convey about the state of the economy. The residual disagreement is by construction orthogonal to all variables included in the filter equation. Note that the two filter equations comprise also a level expectation variable (and its square), constructed from the respective survey question, which is computed as the share to the positive answer minus the share to the negative answer, excluding the share to the 'remain equal' category.

Figure 2 shows the raw and residual disagreement series. Residual disagreement appears still to have the cyclical profile that we can observe for raw disagreement;

---

[8]Diebold *et al.* (1994) provide closed-form solutions for the case when $R = 2$ and $\mathbf{Z}$ contains no more than three variables by employing a first-order Taylor series expansion around the parameters that govern the transition probabilities.

[9]For details about the EM algorithm in general see in particular Dempster *et al.* (1977). For its use in the time-series model context, see e.g. Hamilton (1994) for the exogenous and Diebold *et al.* (1994) for the case of endogenous transition processes.

with more agreement gradually building up throughout expansion periods, and quick adjustment to higher disagreement levels in the course of a recession.

4.1   Real activity

Tables 2 and 3 present the model estimates for US GDP growth based on the sample from 1978Q1-2010Q3. Figure 3 shows the corresponding smoothed regime probability estimates. While allowing the intercept and residual variance to switch across three regimes, the autoregressive dynamics are yet constrained to be equal across regimes, so as to mimic the original Hamilton (1989) model scheme.[10] Instead of referring to estimates for γ, we report the γ-implied transition probabilities at selected moments of the disagreement distribution.

For the model involving the raw disagreement measure (Table 2, Panel A), the estimated conditional means per regime equal [4.6,2.5,-0.6]%, thus we refer to them as *strong expansion*, *medium growth*, and *recession*. For the model based on filtered disagreement, the conditional mean estimates equal [4.8, 2.7,-0.8]%.

Based on the model involving the raw disagreement measure, the endogenous transition matrix estimates (Table 2, Panel B) reveal the following patterns: The probability of moving from medium growth to a recession equals an estimated 84% at times of high agreement, as opposed to virtually 0% at times of high disagreement; The probability of remaining in a medium-growth regime equals an estimated 16% and 45% at times of maximum agreement and disagreement, respectively. Continuance within a strong expansion regime is more likely at times of maximum agreement (91% when full agreement vs. 65% when full disagreement). Estimated error bounds (Figures 6 and 7) suggest that the implied probabilities respectively at minimum and maximum disagreement are statistically different at conventional levels of significance.

---

[10]To facilitate the estimation, in particular to allow for the application of the EM algorithm, we follow a two-step approach to estimation. Step 1: Estimate the 3-regime switching model (with endogenous transition process via **Z**) while yet excluding autoregressive lags. Step 2: Take the residuals from Step 1 and estimate an autoregressive model (of optimal order chosen according to the Bayesian information criterion) without regime-switching. Further note that the autoregressive coefficient estimates are not reported in the tables; only the mean estimates are reported.

When instead employing the filtered disagreement measure (Table 3, Panel B), the Z-dependent transition matrix estimates suggest that switching from medium-growth to recession at times of high agreement has a 15% probability, while at times of full disagreement that probability would fall to 5%. Compared to the estimates based on raw disagreement, the variation in that probability induced by disagreement has therefore become tighter (and somewhat less significant - see also Figures 8 and 9), while however still suggesting a threefold rise in that probability when migrating from high disagreement towards high agreement. More meaningful appear now the estimates for switching from strong expansion right to a recession period; when based on raw disagreement, we obtain an estimated probability of 13% for that scenario under high disagreement, while the model estimates based on filtered disagreement suggest a quasi-zero probability all along the disagreement distribution.

These differences in estimates for the transition probabilities are of course conform with the estimated smooth regime probabilities (since the former are a function of the latter), as depicted in the Figure 3. The model involving raw disagreement suggests a rather direct switch from strong expansion to recession at some points over the sample period, e.g. in 1981Q3-Q4. The model based on filtered disagreement considers the period from 1979-1981Q3 rather a medium growth period that is then followed by a recession state in 1981Q4 with 61%, and in 1982Q1 with 99.9% probability. Otherwise, results for staying in medium growth regimes conditional on high agreement or disagreement prevailing appear robust when employing the filtered disagreement measure: at times of high agreement, we remain in medium growth with 57% probability, opposed to 92% at times of high disagreement.

For the model including raw disagreement, the ergodic probabilities, that is, the long-run means of the growth-regimes that are now depending on the prevailing level of agreement equal, under the assumption of high disagreement in the long-run, to [54,33,13]%, and under the assumption of full agreement in the long-run to [0,2,98]% (Table 2, Panel B). These estimates would suggest that in 100 years, we would be experiencing 54, 33, and 13 years respectively in strong positive, medium positive growth, and recession regimes if there was constant high disagreement. If instead there was steady agreement in the long-run, we would spend two years cumulative in either medium or strong expansion periods and about 98 years in recession. Er-

godic probabilities implied by the median level of disagreement equal [42,46,12]%, i.e. imply a quite balanced distribution for strong and medium-growth episodes, and 12 out of 100 years to be spent in recession.

Ergodic probabilities should certainly be interpreted with caution, since the assumption of from now on ever-lasting constant full agreement or disagreement is obviously verily hypothetical and would bring about structural changes that may imply an evolution of growth regimes that differ from such estimates. They do, however, lead to the same conclusions that we draw from the estimated endogenous transition matrix and are meant to clarify the role of dispersion in beliefs for qualifying transition probabilities: disagreement appears to render the economy less vulnerable, recessions less likely, and growth more sustainable in the long-run.

In the top panel of Table 8 we report in-sample measures of fit for the GDP growth variable, extended by further benchmark models where we restrict the regimes back to one. We see that allowing for regime-specific dynamics significantly improves the fit of the model, with the marginal contribution from identifying the three regimes separately amounting to 15pp of additional variation in GDP growth that the model can explain. Allowing the transition process to depend endogenously on the disagreement level leads to a further improvement in the model fit, with the log-likelihood ratio to the AR benchmark being in all cases statistically significant at least at the 1% confidence level.

4.2   Volatility in stock markets

The RS model estimates and resulting endogenous transition probabilities are reported in Table 4 and plotted along with estimated error bounds in Figures 8 and 9. Unlike for GDP growth, now not only the intercept and residual variances but all autoregressive coefficients are allowed to switch, i.e. the entirety of the model's parameter space can be distinct across regimes.[11]

---

[11]Two autoregressive lags have been judged to be sufficient to capture persistence within the regimes. With more than two lags, our experience was that for the high volatility regime that is identified for about six quarters in the sample, the models would produce too high, that is, almost 'perfect', in-sample fit, thus rendering residuals in that regime virtually zero and in the sequel causing the likelihood to diverge to infinity.

The conditional means per regime which are obtained as a weighted average of observed realized volatility, where weights are the smooth probabilities per regime, equal [36.8,17.7,12.0]pp (annualized) for the model involving the raw disagreement measure, thus we refer to the states as *high*, *medium*, and *low volatility* regimes. There is a clear tendency for the autoregressive dynamics to become more persistent, that is, more lags become significant when migrating from higher towards lower volatility regimes. For the model including filtered disagreement, the conditional means per regime equal [38.3,17.8,11.4] percentage points; which is close to the estimates based on raw disagreement.

As regards the endogenous transition probabilities, the following patterns can be observed for the raw disagreement model: The probability of moving from medium to high volatility regimes equals 14% and 3% at times of maximum agreement, disagreement respectively. The probability of switching from low directly to high volatility regimes equals 68% and virtually 0% at times of full agreement, disagreement respectively. The probability of staying in the medium volatility regime equals 58% and 84% respectively at times of high agreement and disagreement. After all, also prevalence within the low volatility state is more sustainable at times of full disagreement (97% compared to 13% at times of full agreement). All nine probabilities in the transition matrix (Figure 8), appear to depend on agreement levels to a statistically significant extent. When using the filtered disagreement measure instead (see Table 5, Panel B and Figure 9), six of nine probabilities in the transition matrix appear to depend significantly on disagreement levels.

The estimated disagreement-dependent ergodic probabilities, based on raw disagreement (Table 4, Panel B), once more confirm the findings: The thought experiment of living through 100 years would imply that we spend 30, 32, and 38 years respectively in high, medium, and low volatility regimes when there was constant high agreement, as opposed to 1, 18, and 81 years when there was high disagreement in the long run. These estimates compare to [32,37,32]% and [1,13,86]% respectively under constant high agreement and disagreement, for the model including the filtered disagreement measure. They are broadly in line with the estimates based on raw disagreement.

The hypothesis that disagreement levels are reflective of the degree of vulnera-

bility of financial markets therefore receives empirical support, too. In particular, switches straight from a low to a high volatility regime, without intermediate visit of a medium volatility regime, appear to be more likely at times of high agreement, with high disagreement levels being able to let that probability approach zero.[12]

In view of the finding that one of the three regimes is explosive, and to further assess the robustness of the findings presented so far, we also estimate a version of the model where realized volatility enters in first differences. Results are presented in Tables 6 and 7 and Figures 10 and 11. The model dynamics are here allowed to switch between only two regimes, the reason being that a third regime could not be separately identified.[13] Moreover, one autoregressive lag has now been sufficient to capture persistence, which is actually significant in only one of the regimes. We refrain from discussing the results for that additional model in detail, and merely conclude that the role of disagreement in qualifying the transition probabilities remains robust.

As we report in the middle panel of Table 8, allowing for regime-specific dynamics leads to a strong improvement in the in-sample fit of the model: The portion of variation in volatility explained by the model increases from 27% in the case of a simple autoregressive process to 81% when considering the three distinct endogenous regimes. This fact is mirrored by the log-likelihood ratio relative to the AR benchmark, which suggests a statistically significant improvement at least at a 1% confidence level. The bottom panel of Table 8 shows the in-sample measures for the models comprising the first differences of volatility. The results suggest that the RS model for volatility in differences (opposed to the model in levels) results in a relatively smaller gain with regard to in-sample fit, with the additional portion of variation explained being merely 7pp.

---

[12] As a robustness check, we have as well changed the underlying questions and used Q17 for the volatility and Q13 for the GDP growth model. Results presented in Section 4 remain robust. Moreover, we have as well considered disagreement measures based on other forward-looking questions from the survey: Q6 about the expected financial situation and Q9 about real family income. Based on Q9, results for either model context remain robust, both qualitatively/quantitatively. The analysis based on Q6 disagreement (for both models) is the only context in which transition probability dependencies were found to not be significantly different when conditioning on high disagreement, high agreement respectively.

[13] Again, imposing three regimes would here result in too high in-sample fit for one of three regimes which eventually causes the likelihood to diverge to infinity.

4.3 Stochastic simulations

We now employ one model estimate per context (the GDP growth and realized volatility in US stock markets), both of them based on the respective filtered disagreement measures, and simulate artificial data for a series of 10,000 periods under the counterfactual assumption of steady high versus steady low disagreement. The resulting simulated paths as well as respective kernel density distribution charts are presented in Figures 12 and 13. Selected moments of the distributions are summarized in Table 9.

For the GDP growth model, we have chosen the 1% and 99% quantile of the filtered disagreement distribution under which the economy is simulated. The long run medians under high agreement and high disagreement equal 0.8% and 2.7% GDP growth year-on-year. The dispersion around the mean is about 2.5 times as large under full agreement, i.e. the amplitude of the business cycle is on average considerably higher under high agreement. The estimated interquartile ranges, 4.9pp and 1.6pp under high agreement and disagreement, respectively, confirm this finding.

For the volatility model, the 25% and 75% quantiles of the filtered disagreement distribution have been taken to derive the implied transition matrix, under which the volatility in US stock markets has then been simulated. The medians under high agreement and disagreement suggest that high agreement induces higher volatility on average in the long run (14.6 vs 12.0 annualized pp volatility). The distribution under agreement is more skewed to the right, i.e. towards higher volatility (skewness equals 11.8 and 5.9 under high agreement and disagreement respectively). The estimated kurtosis confirms that the simulated distribution under agreement has much longer tails, that is, a much larger portion of the variance is driven by infrequent extreme deviations from the center of the volatility distribution.

4.4 Out-of-sample forecasting

In addition to examining the models' performance in capturing the variables' dynamics in-sample, we assess the models' performance in forecasting out-of-sample. The respective first of the three pairs of models for GDP growth, realized volatility and

volatility in first differences, i.e. the ones involving the raw disagreement measures, are used to produce recursive out-of-sample forecasts over the period from 2002Q1-2010Q3. An expanding window scheme is chosen, that is, we let the sample start in 1978Q1 and only the forecast origin to move recursively, with the models being re-estimated at each point to then produce a set of one up to four quarter ahead forecasts.[14]

Five model schemes are considered for each variable:

1. AR model containing the target variable only

2. Bivariate VAR model containing the target variable and the disagreement measure

3. RS model for the target variable only

4. RS model for the target variable, with the transition probabilities being a function of disagreement, conditional on a predicted out-of-sample disagreement path (via an auxiliary AR model), which is used to derive the transition matrix (recursively over the horizon) and thereby the out-of-sample regime probabilities

5. RS model for the target variable, with the transition probabilities being a function of disagreement and disagreement held fix along the out-of-sample horizon to derive the transition matrix and thereby the out-of-sample regime probabilities.

The lag lengths of the AR and the VAR models, including the auxiliary AR model for predicting disagreement in model Scheme 4, are allowed to vary over time. The Bayesian information criterion is used to choose the lag length between a minimum zero and maximum four over the test period.

Tables 10 and 11 show the forecast evaluation results for the three groups of models, respectively for the full test period (35 periods) and a reduced window for

---

[14]At the outset of the test period, the forecast origin is adjusted (the very first to 2001Q1) so as to obtain the same number of one- up to four-period ahead forecasts over the 35-quarter test sample.

the recession period, starting at the peak of the US business cycle in 2007Q4 and then covering the remaining quarters until 2010Q3 (12 periods).

Referring to the results for GDP growth in Table 10, the VAR model with disagreement reduces Root Mean Square Forecast Errors (RMSE) relative to those resulting from the AR benchmark by 10%-12% at short horizons of up to two quarters. For the three and four-quarter horizon, the RS model including disagreement to condition the transition probabilities performs best, with the RMSE reduced by up to 18%. The same ranking of best-performing models results for the crisis test sample (Table 11), for which the RMSE ratios suggest an even stronger role for disagreement with a statistically significant 24% improvement at the four-quarter horizon.

The proportions of correctly predicted directions of change in GDP growth confirm that incorporating disagreement in the model improves forecast performance markedly, with e.g. 93% of the directions of change over the three-quarter horizon in the recession part of the test sample foreseen correctly.

Overall, the forecast evaluation for GDP let us conclude that both the disagreement variable as well as the regime-switching specification help improve out-of-sample performance significantly. While at short horizons the regime-switching does not yet improve performance, where joint dynamics with disagreement are better modeled in a one-regime VAR, the regime-switching can clearly improve performance closer to the one-year horizon.

Evaluation results for volatility in levels reveal that the RS models' explosive third regime indeed appears to dominate over a significant portion of the test sample, thereby resulting in large errors of up to 628pp at the four-quarter horizon and eventually leading us to conclude that the model is of no avail for forecasting out-of-sample. It has however proven useful to consider the specification for volatility in first differences, for which the three RS models' regimes are both stable over the recursive test period. The VAR model's RMSE ratio indicates a 7% improvement at the one-quarter horizon, which is comparable however to the RS model without the additional disagreement variable as input. Otherwise, the five model schemes' performances appear rather equal, both over the full test period and the reduced recession sample, with all ratios surrounding one. Regarding direction of change

measures, too, all five model schemes' performances appear limited. At the shortest horizon, a somewhat gain can be reported for the VAR's performance, where the presence of disagreement in the model results in 83% of one-quarter changes predicted correctly over the recession period.

5 Conclusions

Both with respect to the evolution of the real economy and to realized volatility in financial markets, the hypothesis that the level of economy-wide disagreement has the ability to qualify the propensity to continue within or switch between different regimes has received clear empirical support. Rising levels of agreement make a medium-growth expansion less probable to prevail and a transition to lower-growth regimes more likely. Also continuance within medium-growth expansions has been found to be more sustainable with more disagreement. Likewise, concerning the volatility regimes in financial markets, the model results suggest that switches from low to high volatility regimes become considerably more likely at times of high aggregate agreement.

Along with in-sample criteria of model performance, the regime-switching models' out-of-sample ability has been assessed based on an eight-year test period. The results suggest that the regime-switching model for GDP growth, with disagreement used to condition the transition probabilities, improves the precision of GDP projections significantly at horizons up to four quarters, in particular so during the recession period of 2007-2009. Out-of-sample performance of the models for realized volatility, on the other hand, appears limited. Steps towards more appropriate endogenous modeling of disagreement (and therefore potentially towards providing more accurate predictions for disagreement) can certainly help to further the forecast performance in either setting. As concerns the models presented for realized volatility, despite their failure to provide more precise out-of-sample projections, we would argue that the models can still be of avail for the purpose of producing counterfactual scenarios (with the projections being conditional e.g. on different disagreement assumptions, or regime assumptions, or both).

We shall emphasize, once more, that the role of disagreement in our model framework is to qualify the *propensity* of the economy to switch regimes, i.e. it may impact the associated transition *probabilities*, while not having a direct channel through which it could drive economy-wide output, or volatility in financial markets. For the economy or some particular market to switch regimes, we still have to await the arrival of some exogenous shock, which then might cause a regime shift with a disagreement-level dependent transition probability.

The results have been found to be robust to the deployment of raw as well as filtered disagreement measures, with the filtering being accomplished by auxiliary models that relate disagreement to a comprehensive set of variables that capture macro and financial conditions in the US economy.

While we do propose in this paper a method for controlling for eventual feedback effects from aggregate macroeconomic conditions to the level of disagreement, there is still further research needed in order to properly understand these mechanisms. Along the lines discussed in the literature review of our paper, the theoretical macro model frameworks should be further developed to account for endogenous information flows, learning, signal extraction, and time-varying attention levels on the one hand, and a sufficient degree of agent heterogeneity on the other. It is only if the modeling and empirical identification techniques explicitly embed a transmission channel from the macroeconomy to the expectation formation process that one could also derive implementable normative implications for central bank communication, financial reporting and supervision.

References

ACEMOGLU, D. (2011). *Diversity and technological progress.* NBER Working Papers 16984, National Bureau of Economic Research.

AKCIGIT, U. and KERR, W. (2010). *Growth through heterogeneous innovations.* NBER Working Papers 16443, National Bureau of Economic Research, Inc.

ALMANIDES, P. (2011). Banking crises, early warning models, and efficiency. *Department of Economics, Rice University, Unpublished working paper.*

ANGELETOS, G.-M. and PAVAN, A. (2007). Efficient use of information and social value of information. *Econometrica*, **75** (4), 1103–1142.

BADARINZA, C. and BUCHMANN, M. (2009). Inflation perceptions and expectations in the euro area - the role of news. *ECB Working Paper No. 1088.*

BOMBERGER, W. (1996). Disagreement as a measure of uncertainty. *Journal of Money, Credit and Banking*, **28** (3), 381–92.

— (1999). Disagreement and uncertainty: A reply. *Journal of Money, Credit and Banking*, **31** (2), 273–76.

BORIO, C. and LOWE, P. (2002). Asset prices, financial and monetary stability: Exploring the nexus. *BIS Working Paper No. 114.*

BUCHMANN, M. (2010). Corporate bond spreads and real activity in the euro area - least angle regression forecasting and the probability of the recession. *ECB Working Paper No. 1286.*

BURASCHI, A. and WHELAN, P. (2010). *Macroeconomic uncertainty, difference in beliefs, and bond risk premia.* Working paper series, Imperial College London.

CLARK, T. and WEST, K. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, **138** (1), 291–311.

COX, D. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.

DAVIS, E. and KARIM, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, **4**, 89–120.

DEMPSTER, A., LAIRD, N. and RUBIN, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**, 1–38.

DIEBOLD, F. X., LEE, J.-H. and WEINBACH, G. (1994). *Regime switching with time-varying transition probabilities, Nonstationary Time Series Analysis and Cointegration. (Advanced Texts in Econometrics, C.W.J. Granger and G. Mizon, eds.).* Oxford University Press.

ECB (2011). Financial Stability Review, December 2011.

FAVARA, G. and SONG, Z. (2009). *House price dynamics with heterogeneous expectations.* Tech. rep., University of Lausanne.

FLOOD, R. and MARION, N. (1999). Perspectives on the recent currency crisis literature. *International Journal of Finance and Economics*, **4**, 1–26.

GALA, V. and VOLPIN, P. (2010). Social value of information in a levered economy. *AFA 2011 Denver Meetings Paper.*

GEANAKOPLOS, J. (2009). *The leverage cycle.* Cowles Foundation Discussion Papers 1715, Cowles Foundation for Research in Economics, Yale University.

GIORDANI, P. and SODERLIND, P. (2003). Inflation forecast uncertainty. *European Economic Review*, **47** (6), 1037–1059.

GROSSMAN, S. and STIGLITZ, J. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, **70** (3), 393–408.

HAMILTON, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.

— (1994). *Time series analysis.* Princeton University Press.

HARRISON, J. and KREPS, D. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics*, **92** (2), 323–36.

HIRSHLEIFER, J. (1971). The private and social value of information and the reward to inventive activity. *American Economic Review*, **61** (4), 561–74.

LACY, M. (2006). An explained variation measure for ordinal response models with comparisons to other ordinal R^2 measures. *Sociological Methods and Research*, **32**, 469–520.

LAHIRI, K. and SHENG, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, **25** (4), 514–538.

LEGERSTEE, R. and FRANSES, P. (2010). Does disagreement amongst forecasters have predictive value? *Tinbergen Institute Discussion Paper No. 088/4*.

LI, D. and LI, G. (2010). Belief dispersion, dispersion of belief changes, and trading volume: Evidence from surveys of consumers and professional forecasters. *Working paper available at SSRN: http://ssrn.com/abstract=1572439*.

MILLER, E. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance*, **32** (4), 1151–68.

MORRIS, S. and SHIN, H. (2002). Social value of public information. *American Economic Review*, **92** (5), 1521–1534.

NIMARK, K. (2010). *Speculative dynamics in the term structure of interest rates*. Economics Working Papers 1194, Department of Economics and Business, Universitat Pompeu Fabra.

PFAJFAR, D. and ZAKELJ, B. (2011). *Uncertainty and disagreement in forecasting inflation: Evidence from the laboratory*. Working paper series, Tilburg University.

RICH, R. and BUTLER, J. (1998). Disagreement as a measure of uncertainty: A comment on Bomberger. *Journal of Money, Credit and Banking*, **30** (3), 411–419.

SCHEINKMAN, J. and XIONG, W. (2003). *Heterogeneous beliefs, speculation and trading in financial markets.* Tech. rep., Paris-Princeton Lectures in Mathematical Finance, Springer.

SIMS, C. (2009). *Inflation expectations, uncertainty and monetary policy.* BIS Working Papers 275, Bank for International Settlements.

STOCK, J. and WATSON, M. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, **41**, 788–829.

TIAN, W. and YAN, H. (2009). *Heterogeneous beliefs, competition, and the frailty of financial innovation.* Working paper series, University of North Carolina at Charlotte.

TOMURA, H. (2009). *Heterogeneous beliefs and housing-market boom-bust cycles in a small open economy.* Working Papers 09-15, Bank of Canada.

VELDKAMP, L. (2006). Media frenzies in markets for financial information. *American Economic Review*, **96** (3), 577–601.

WHALEM, G. (1991). A proportional hazards model of bank failure: An examination of its usefulness as an early warning tool. *Economic Review*, **27**.

WORRELL, D. (2004). Quantitative assessment of the financial sector: An integrated approach. *IMF Working Paper No. 04/153.*

YU, J. (2011). Disagreement and return predictability of stock portfolios. *Journal of Financial Economics*, **99** (1), 162–183.

ZARNOWITZ, V. and LAMBROS, L. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, **95** (3), 591–621.

Appendix

Table 1: Filter equation estimates

DIURX	coef.		DIBUSC1	coef.	
const.	0.3178	(0.00)	const.	0.3385	(0.00)
LEXPURX	-0.0147	(0.00)	LEXPBUSC1	-0.0034	(0.14)
LEXPURX^2	-0.0064	(0.00)	LEXPBUSC1^2	-0.0011	(0.36)
GDP	-0.0100	(0.00)	GDP	-0.0148	(0.00)
GPDIY	0.0016	(0.00)	GPDIY	0.0019	(0.00)
PCECTPIY	0.0026	(0.01)	PCECTPIY	0.0032	(0.01)
PSAVERTY	-0.0017	(0.18)	PSAVERTY	0.0002	(0.91)
TCUY	-0.0008	(0.28)	TCUY	-0.0004	(0.68)
UNRATE	0.0046	(0.00)	UNRATE	0.0024	(0.14)
TOTALSLY	0.0001	(0.87)	TOTALSLY	-0.0003	(0.53)
YLDC	0.0062	(0.00)	YLDC	0.0071	(0.00)
M2SLY	0.0013	(0.10)	M2SLY	0.0018	(0.04)
R^2	0.69			0.69	
obs.	131			131	

Variable list	
DIURX	Disagreement based on Michigan Survey Question 17
DIBUSC1	Disagreement based on Michigan Survey Question 13
LEXPURX	Level expectations based on Michigan Survey Question 17
LEXPBUSC1	Level expectations based on Michigan Survey Question 13
GDP	Real GDP, YoY
GPDIY	Gross private domestic investment, YoY
PCECTPIY	Personal Consumption Expenditures: Chain-type Price Index, YoY
PSAVERTY	Personal Saving Rate, 4-quarter differences
TCUY	Capacity Utilization: Total Industry, YoY
UNRATE	Civilian Unemployment Rate
TOTALSLY	Total Consumer Credit Outstanding, YoY
YLDC	Yield curve, 10y - Fed funds rate
M2SLY	M2 Money Stock, YoY

Table 2: Regime-switching model estimates for GDP growth (including raw disagreement)

A: Estimated parameters

	$\beta_{0,s}$		σ_s^2
Expansion	4.585	(0.00)	1.364
Medium growth	2.523	(0.00)	0.379
Recession	-0.554	(0.09)	2.860

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	Expansion			Medium growth			Recession			Ergodic prob.		
	Exp	Med	Rec	Exp	Med	Rec	Exp	Med	Rec	Exp	Med	Rec
1%	.91	.09	.00	.00	.16	.84	.00	.02	.98	.00	.02	.98
25%	.87	.12	.00	.07	.79	.14	.00	.07	.93	.14	.27	.59
50%	.84	.15	.01	.15	.83	.02	.00	.13	.87	.42	.46	.12
75%	.80	.18	.03	.25	.75	.00	.00	.23	.77	.52	.42	.07
99%	.65	.21	.13	.55	.45	.00	.07	.50	.43	.54	.33	.13

Note: p-values are reported in parentheses.

Table 3: Regime-switching model estimates for GDP growth (including filtered disagreement)

A: Estimated parameters

	$\beta_{0,s}$		σ_s^2
Expansion	4.787	(0.00)	1.430
Medium growth	2.689	(0.00)	0.656
Recession	-0.756	(0.03)	2.742

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	Expansion			Medium growth			Recession			Ergodic prob.		
	Exp	Med	Rec	Exp	Med	Rec	Exp	Med	Rec	Exp	Med	Rec
1%	.96	.04	.00	.28	.57	.15	.00	.00	.99	.19	.03	.78
25%	.88	.12	.00	.15	.74	.11	.00	.07	.93	.32	.26	.42
50%	.83	.17	.00	.11	.79	.10	.00	.15	.84	.30	.44	.27
75%	.73	.27	.00	.08	.84	.08	.00	.40	.60	.20	.67	.13
99%	.44	.55	.00	.03	.92	.05	.01	.90	.09	.06	.90	.05

Note: p-values are reported in parentheses.

Table 4: Regime-switching model estimates for realized volatility (including raw disagreement)

A: Estimated parameters

	$\beta_{0,s}$		$\beta_{1,s}$		$\beta_{2,s}$		σ_s^2
High	2.252	(0.93)	1.892	(0.04)	0.085	(0.95)	224.893
Medium	9.532	(0.00)	0.448	(0.00)	0.000	(0.99)	8.250
Low	5.800	(0.00)	0.206	(0.00)	0.205	(0.00)	1.760

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	High volatility			Medium volatility			Low volatility			Ergodic prob.		
	High	Med	Low	High	Med	Low	High	Med	Low	High	Med	Low
1%	.00	.22	.78	.14	.58	.28	.68	.19	.13	.30	.32	.38
25%	.02	.53	.45	.09	.67	.24	.04	.31	.66	.06	.51	.43
50%	.07	.69	.23	.07	.72	.21	.00	.19	.81	.04	.44	.52
75%	.21	.70	.09	.05	.77	.17	.00	.10	.90	.02	.34	.64
99%	.63	.37	.01	.03	.84	.13	.00	.03	.97	.01	.18	.81

Note: p-values are reported in parentheses.

Table 5: Regime-switching model estimates for realized volatility (including filtered disagreement)

A: Estimated parameters

	$\beta_{0,s}$		$\beta_{1,s}$		$\beta_{2,s}$		σ_s^2
High	1.641	(0.96)	1.890	(0.06)	0.185	(0.90)	238.024
Medium	9.808	(0.00)	0.441	(0.00)	0.004	(0.92)	9.048
Low	5.788	(0.00)	0.206	(0.00)	0.211	(0.00)	1.845

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	High volatility			Medium volatility			Low volatility			Ergodic prob.		
	High	Med	Low	High	Med	Low	High	Med	Low	High	Med	Low
1%	.33	.12	.55	.06	.68	.26	.60	.25	.15	.32	.37	.32
25%	.23	.38	.39	.06	.69	.25	.02	.34	.64	.05	.53	.42
50%	.17	.53	.30	.06	.69	.25	.00	.25	.74	.04	.47	.49
75%	.10	.72	.18	.06	.70	.24	.00	.15	.85	.02	.37	.61
99%	.01	.96	.03	.06	.71	.23	.00	.03	.97	.01	.13	.86

Note: p-values are reported in parentheses.

Table 6: Regime-switching model estimates for realized volatility in first differences (including raw disagreement)

A: Estimated parameters

	$\beta_{0,s}$		$\beta_{1,s}$		σ_s^2
High	3.120	(0.35)	-0.576	(0.01)	270.00
Low	-0.247	(0.34)	0.036	(0.43)	6.68

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	High volatility High	High volatility Low	Low volatility High	Low volatility Low	Ergodic prob. High	Ergodic prob. Low
1%	.32	.68	.37	.63	.35	.65
25%	.45	.55	.16	.84	.23	.77
50%	.54	.46	.09	.91	.16	.84
75%	.63	.37	.04	.96	.10	.90
99%	.76	.24	.01	.99	.05	.95

Note: p-values are reported in parentheses.

Table 7: Regime-switching model estimates for realized volatility in first differences (including filtered disagreement)

A: Estimated parameters

	$\beta_{0,s}$		$\beta_{1,s}$		σ_s^2
High	3.139	(0.33)	-0.570	(0.01)	264.00
Low	-0.265	(0.31)	0.035	(0.45)	6.66

B: Estimated endogenous transition probabilities and ergodic probabilities

Disagr. quantiles	High volatility High	High volatility Low	Low volatility High	Low volatility Low	Ergodic prob. High	Ergodic prob. Low
1%	.75	.25	.26	.74	.52	.48
25%	.63	.37	.14	.86	.28	.72
50%	.58	.42	.10	.90	.20	.80
75%	.50	.50	.07	.93	.12	.88
99%	.30	.70	.02	.98	.03	.97

Note: p-values are reported in parentheses.

Table 8: Summary in-sample model performances

GDP growth

Model	Log likelihood	R^2	p-value
AR(1)	-194.17	0.77	
VAR(1) with Zraw	-184.08	0.80	0.00
VAR(1) with Zfiltered	-193.19	0.77	0.16
RSAR (R=3)	-125.05	0.92	0.00
RSAR (R=3) with Zraw	-124.50	0.92	0.00
RSAR (R=3) with Zfiltered	-117.73	0.93	0.00

Realized volatility levels

Model	Log likelihood	R^2	p-value
AR(2)	-442.93	0.27	
VAR(2) with Zraw	-442.65	0.27	0.76
VAR(2) with Zfiltered	-440.50	0.29	0.09
RSAR (R=3)	-355.21	0.81	0.00
RSAR (R=3) with Zraw	-360.42	0.80	0.00
RSAR (R=3) with Zfiltered	-355.18	0.81	0.00

Realized volatility first differences

Model	Log likelihood	R^2	p-value
AR(1)	-454.20	0.12	
VAR(1) with Zraw	-452.70	0.14	0.08
VAR(1) with Zfiltered	-451.63	0.15	0.02
RSAR (R=2)	-447.04	0.21	0.16
RSAR (R=2) with Zraw	-448.08	0.20	0.73
RSAR (R=2) with Zfiltered	-446.86	0.21	0.55

Note: All models are estimated based on the 1978Q1-2010Q3 sample. For the VAR models, the reported R^2 and log likelihood values refer to the first equation only (i.e. to GDP or volatility levels/differences). The p-values refer to likelihood ratio test statistics that are based on the restricted AR model and respective unrestricted alternative model schemes, with the number of degrees of freedom for the underlying asymptotic chi-square distributions of the test statistics being set accordingly.

Table 9: Summary statistics for simulated distributions

	GDP growth		Realized volatility	
	Agreement	Disagreement	Agreement	Disagreement
Mean	1.06	2.66	17.35	13.94
Median	0.79	2.70	14.56	12.02
STD	3.08	1.24	15.82	7.37
Skewness	0.08	-0.22	11.83	5.90
Kurtosis	-0.73	0.88	233.30	60.59
Min	-9.19	-3.64	0.15	2.65
Max	9.83	8.95	537.38	141.27
75% Quantile	3.59	3.47	18.55	15.82
25% Quantile	-1.28	1.89	11.64	10.21
InterQ-Range	4.87	1.58	6.91	5.61

Note: The statistics are based on a sample of 10,000 periods simulated from the calibrated GDP growth and realized volatility models presented in Tables 3 and 5, respectively.

Table 10: Out-of-sample forecast evaluation, test sample 2002Q1-2010Q3

|	RMSE (in pp)				RMSE ratios				CW t-stats				Changes correctly pred.			
	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4
GDP growth																
AR	0.80	1.46	1.96	2.25									49%	63%	66%	66%
VAR	0.72	1.28	1.76	2.08	0.90	0.88	0.90	0.92	1.94	1.62	1.56	1.55	69%	66%	60%	60%
RSAR	1.07	1.74	2.12	2.32	1.34	1.19	1.08	1.03	-0.94	-0.86	-0.50	0.02	51%	51%	57%	49%
RSARZp	0.91	1.39	1.68	1.93	1.14	0.95	0.86	0.86	0.59	1.49	1.51	1.35	57%	57%	66%	60%
RSARZc	0.91	1.36	1.63	1.85	1.14	0.93	0.83	0.82	0.59	1.54	1.52	1.38	57%	60%	63%	57%
Realized volatility levels																
AR	9.78	11.24	11.89	12.11									60%	60%	54%	51%
VAR	9.88	11.83	12.42	12.34	1.01	1.05	1.04	1.02	-0.75	-1.07	-1.12	-0.78	57%	60%	54%	51%
RSAR	10.07	16.11	80.64	628.27	1.03	1.43	6.78	51.86	1.25	0.38	-0.05	-0.14	54%	57%	43%	49%
RSARZp	13.35	16.81	26.04	149.27	1.37	1.50	2.19	12.32	-0.09	0.65	-1.36	0.56	51%	54%	43%	46%
RSARZc	13.35	18.11	27.36	119.00	1.37	1.61	2.30	9.82	-0.09	0.50	-1.32	0.69	51%	54%	43%	46%
Realized volatility first differences																
AR	11.22	10.58	10.08	10.02									77%	69%	83%	69%
VAR	10.46	10.50	10.05	10.04	0.93	0.99	1.00	1.00	1.16	0.61	0.78	0.36	77%	71%	80%	69%
RSAR	10.59	10.71	10.11	10.25	0.94	1.01	1.00	1.02	1.41	0.33	1.19	-0.61	71%	77%	74%	69%
RSARZp	11.77	10.82	10.06	10.12	1.05	1.02	1.00	1.01	0.81	0.58	1.53	0.34	77%	74%	80%	69%
RSARZc	11.77	10.84	10.08	10.11	1.05	1.02	1.00	1.01	0.81	0.59	1.59	0.65	77%	74%	80%	69%

Note: Root mean square forecast errors (RMSE) ratios are relative to the AR benchmark. Values smaller one indicate that a model outperforms the benchmark by one minus the ratio. The Clark-West (CW) test statistics indicate whether a gain in performance is significant from a statistical viewpoint (see Clark and West (2007)). Critical values are 2.33 for the 1% level, 1.64 for the 5% level, and 1.28 for the 10% level of significance. In the model acronyms, Zp means predicted and Zc means constant disagreement. See Section 4.4 for details.

Table 11: Out-of-sample forecast evaluation, test sample 2007Q4-2010Q3

|	RMSE (in pp)				RMSE ratios				CW t-stats				Changes correctly pred.			
	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4	H=1	H=2	H=3	H=4
GDP growth																
AR	1.11	2.29	3.16	3.66									42%	50%	58%	42%
VAR	0.95	1.94	2.77	3.30	0.86	0.85	0.88	0.90	2.16	2.23	2.18	2.10	75%	67%	67%	42%
RSAR	1.63	2.74	3.39	3.69	1.47	1.20	1.07	1.01	-1.55	-1.39	-1.19	-0.21	33%	50%	67%	42%
RSARZp	1.28	2.06	2.57	2.96	1.15	0.90	0.81	0.81	0.13	1.82	1.97	1.73	83%	75%	92%	75%
RSARZc	1.28	2.00	2.44	2.79	1.15	0.87	0.77	0.76	0.13	2.01	2.16	1.84	83%	75%	92%	75%
Realized volatility levels																
AR	15.17	17.31	18.26	18.61									75%	58%	58%	58%
VAR	15.28	18.35	19.11	18.88	1.01	1.06	1.05	1.01	-0.48	-1.20	-1.13	-0.68	67%	58%	58%	58%
RSAR	15.60	21.09	30.78	52.14	1.03	1.22	1.69	2.80	0.57	-0.47	-0.84	-0.52	83%	50%	50%	42%
RSARZp	21.77	26.86	30.35	36.70	1.44	1.55	1.66	1.97	-0.64	-1.12	-1.17	-0.54	50%	42%	42%	42%
RSARZc	21.77	29.34	37.58	48.59	1.44	1.69	2.06	2.61	-0.64	-1.15	-1.17	-0.52	50%	42%	42%	42%
Realized volatility first differences																
AR	17.76	16.64	15.73	15.69									75%	67%	83%	75%
VAR	16.63	16.56	15.72	15.68	0.94	1.00	1.00	1.00	1.00	0.40	0.46	0.42	83%	67%	83%	75%
RSAR	16.63	16.90	15.75	15.98	0.94	1.02	1.00	1.02	1.41	0.19	0.85	-0.50	75%	83%	75%	67%
RSARZp	18.76	17.20	15.76	15.82	1.06	1.03	1.00	1.01	0.34	0.39	1.63	0.66	75%	75%	75%	67%
RSARZc	18.76	17.23	15.80	15.80	1.06	1.04	1.00	1.01	0.34	0.39	1.71	1.02	75%	75%	75%	67%

Note: Root mean square forecast errors (RMSE) ratios are relative to the AR benchmark. Values smaller than one indicate that a model outperforms the benchmark by one minus the ratio. The Clark-West (CW) test statistics indicate whether a gain in performance is significant from a statistical viewpoint (see Clark and West (2007)). Critical values are 2.33 for the 1% level, 1.64 for the 5% level, and 1.28 for the 10% level of significance. In the model acronyms, Zp means predicted and Zc means constant disagreement. See Section 4.4 for details.

Figure 1: Disagreement in expectations and macroeconomic aggregates
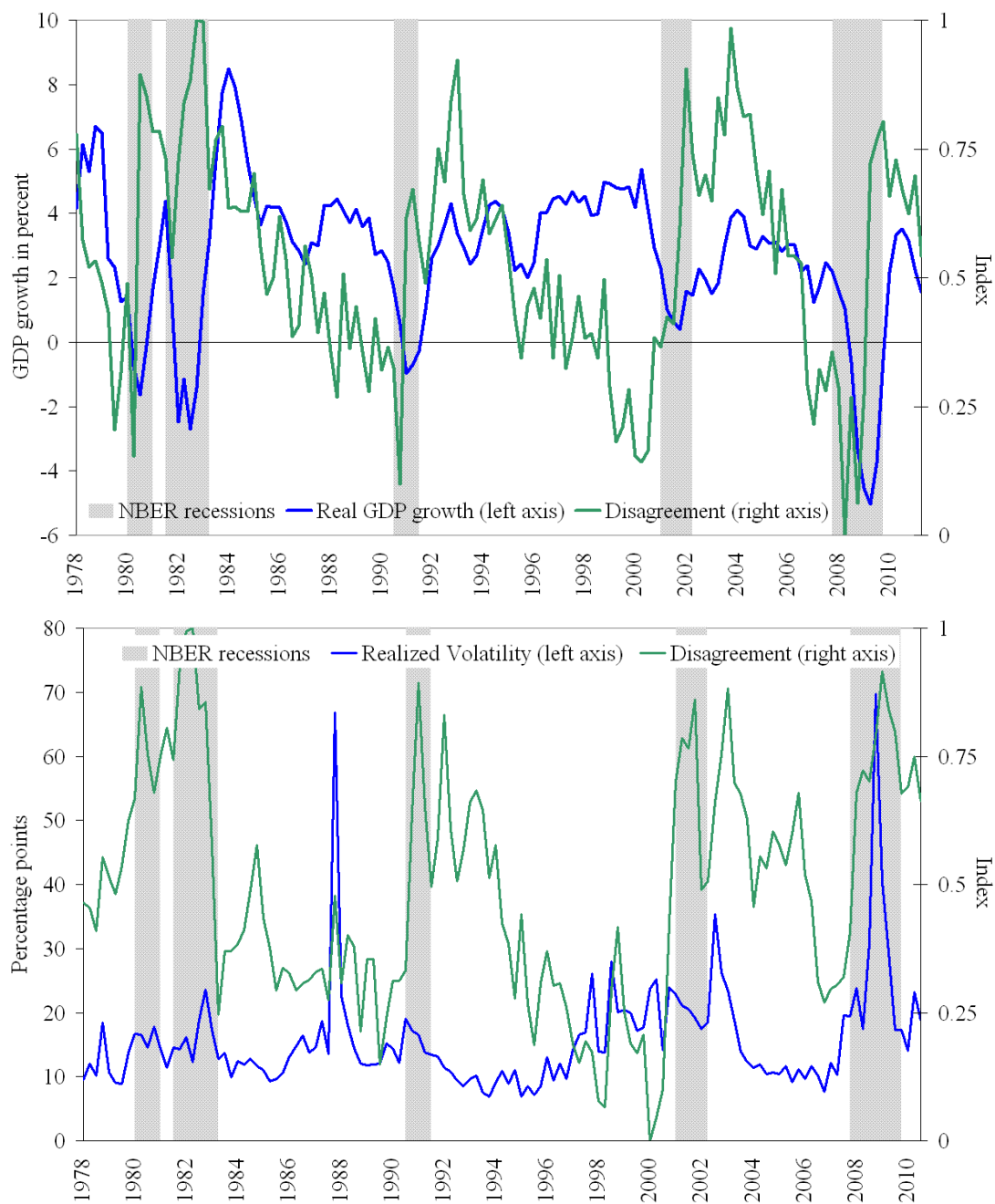
Note: The disagreement measure is raw (not filtered), and has been normalized so as to range between [0,1]. For details as to the construction of the disagreement measure see Section 3.1.

Figure 2: Raw vs. filtered disagreement measures based on Michigan Survey
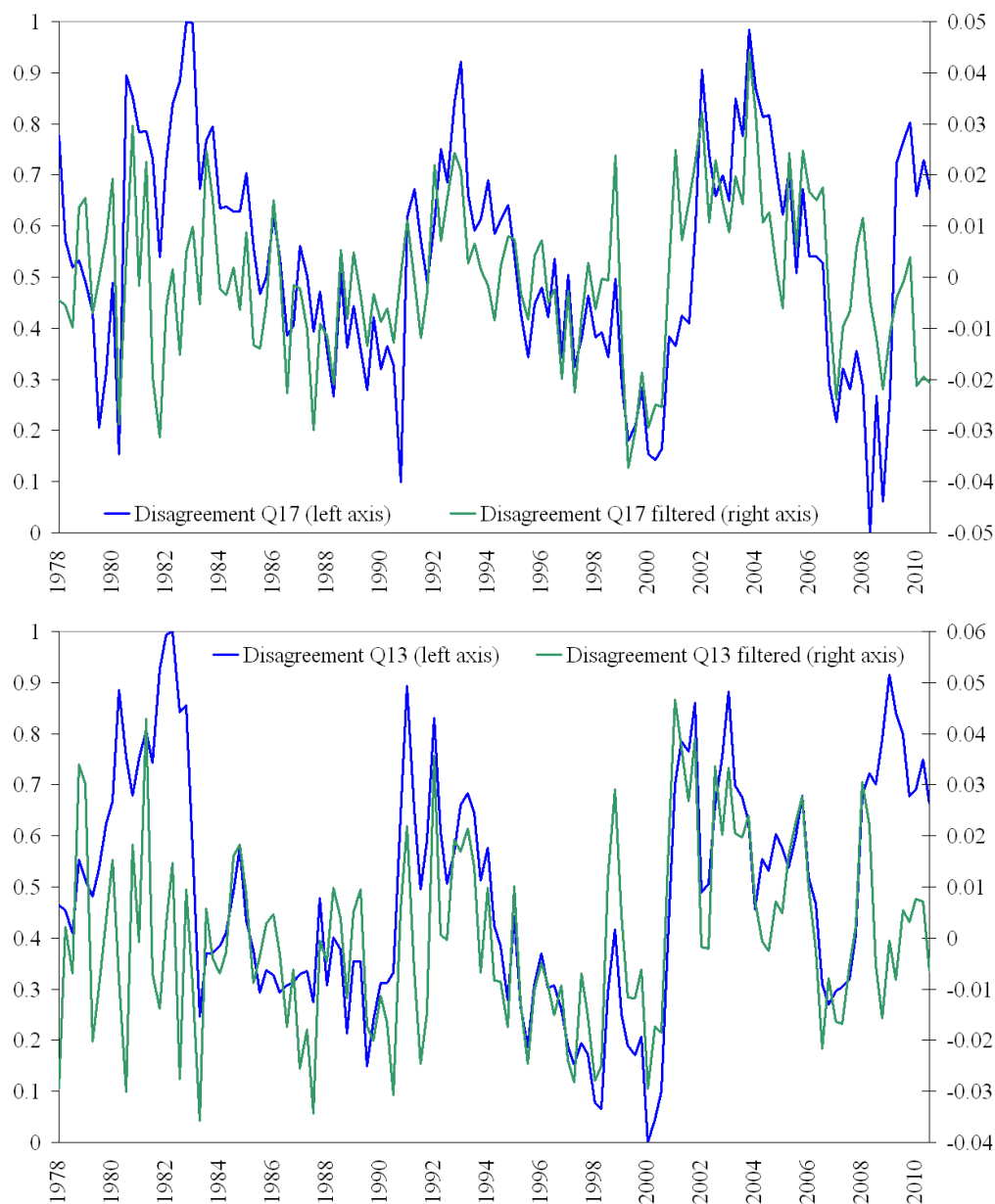
Note: The raw disagreement measure has been normalized so as to range between [0,1]. The filtered disagreement measure is the residual from an auxiliary model that relates raw disagreement to a set of variables capturing macro and financial conditions in the US. For further details see Section 4.

Figure 3: Estimated smooth regime probabilities - GDP growth model

Raw disagreement

Filtered disagreement
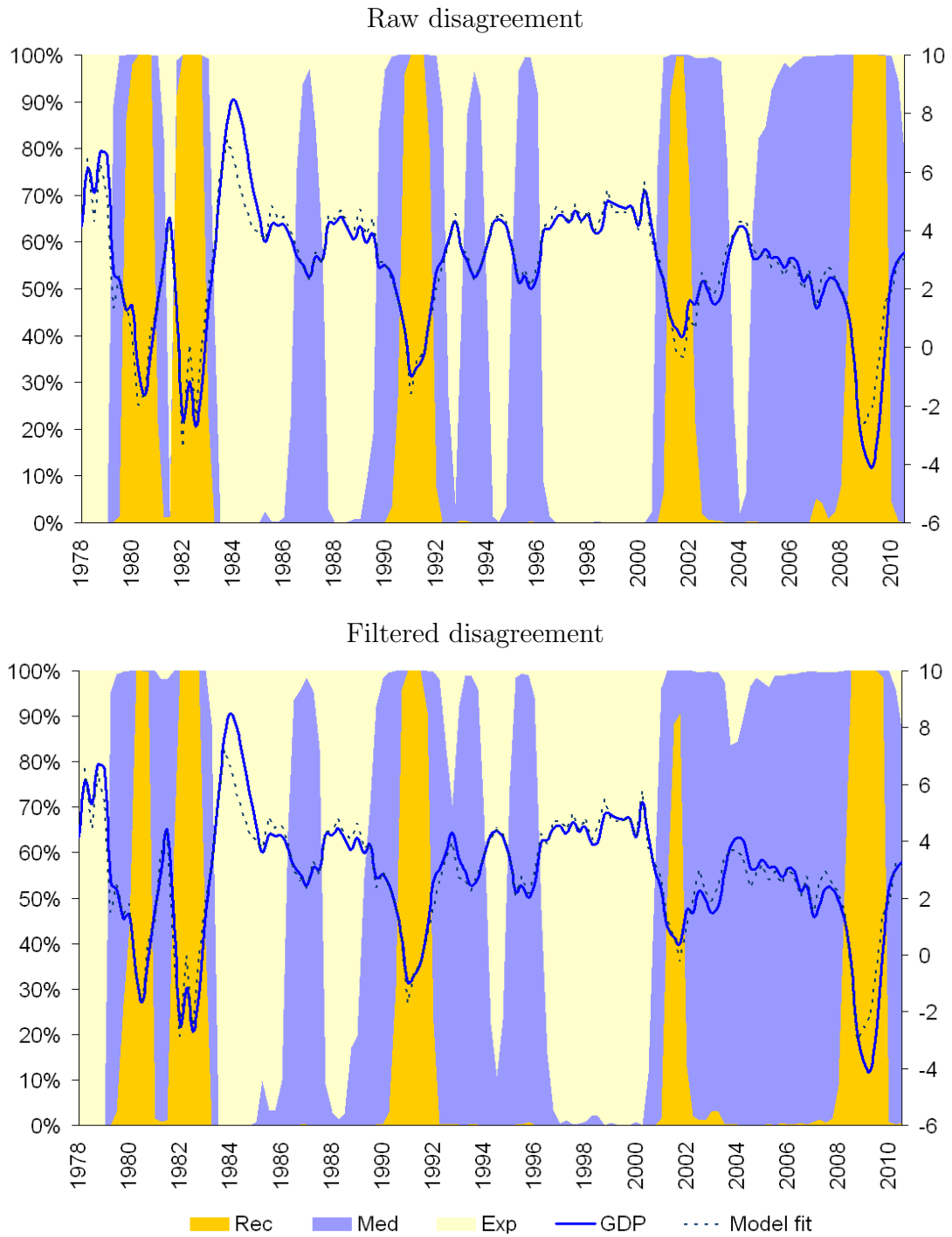
Figure 4: Estimated smooth regime probabilities - Realized volatility model

Raw disagreement

Raw disagreement

High   Med   Low   —— Realized volatility   ···· Model fit
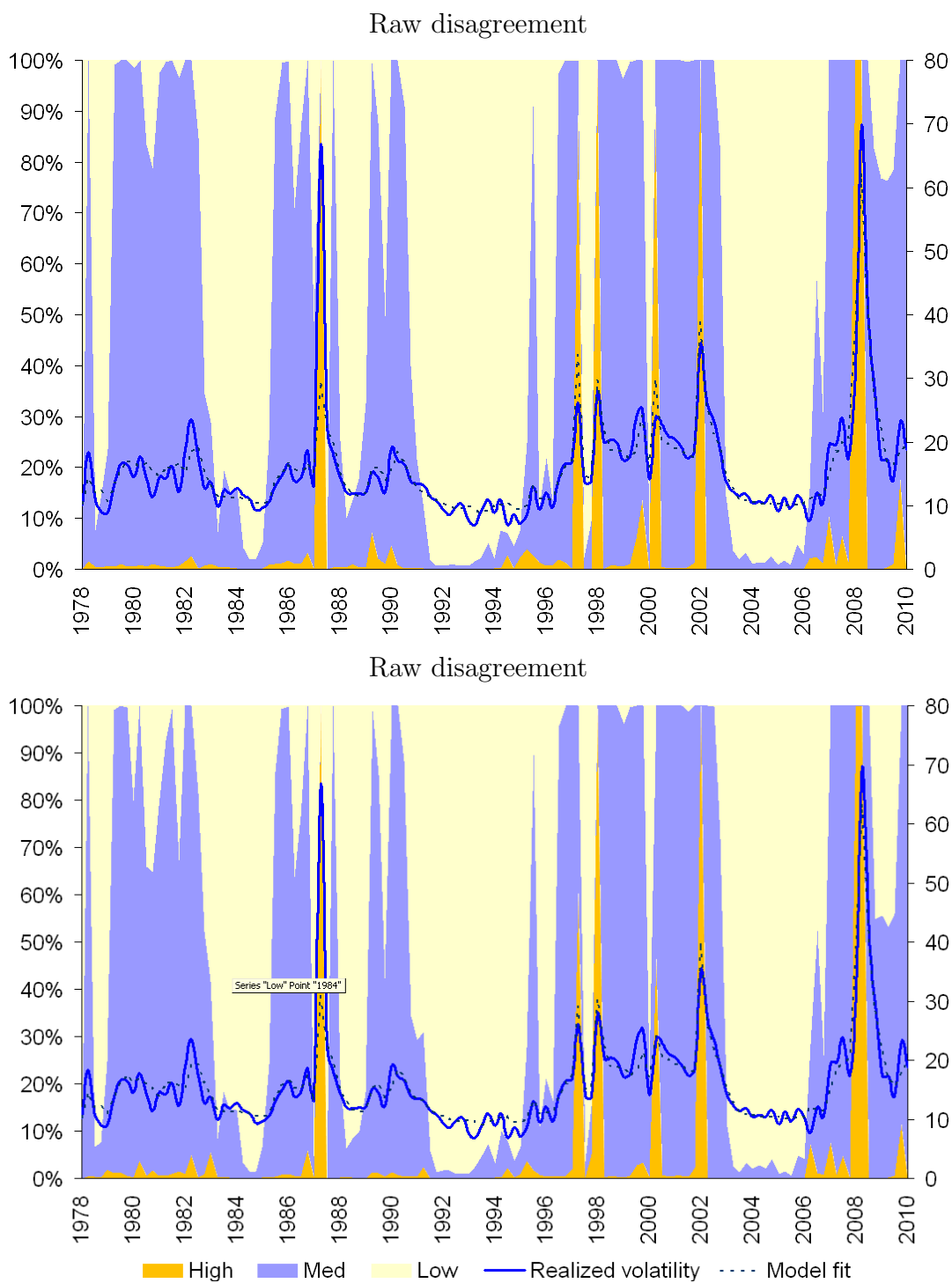
Figure 5: Estimated smooth regime probabilities - Model for realized volatility in first differences

Raw disagreement

Filtered disagreement
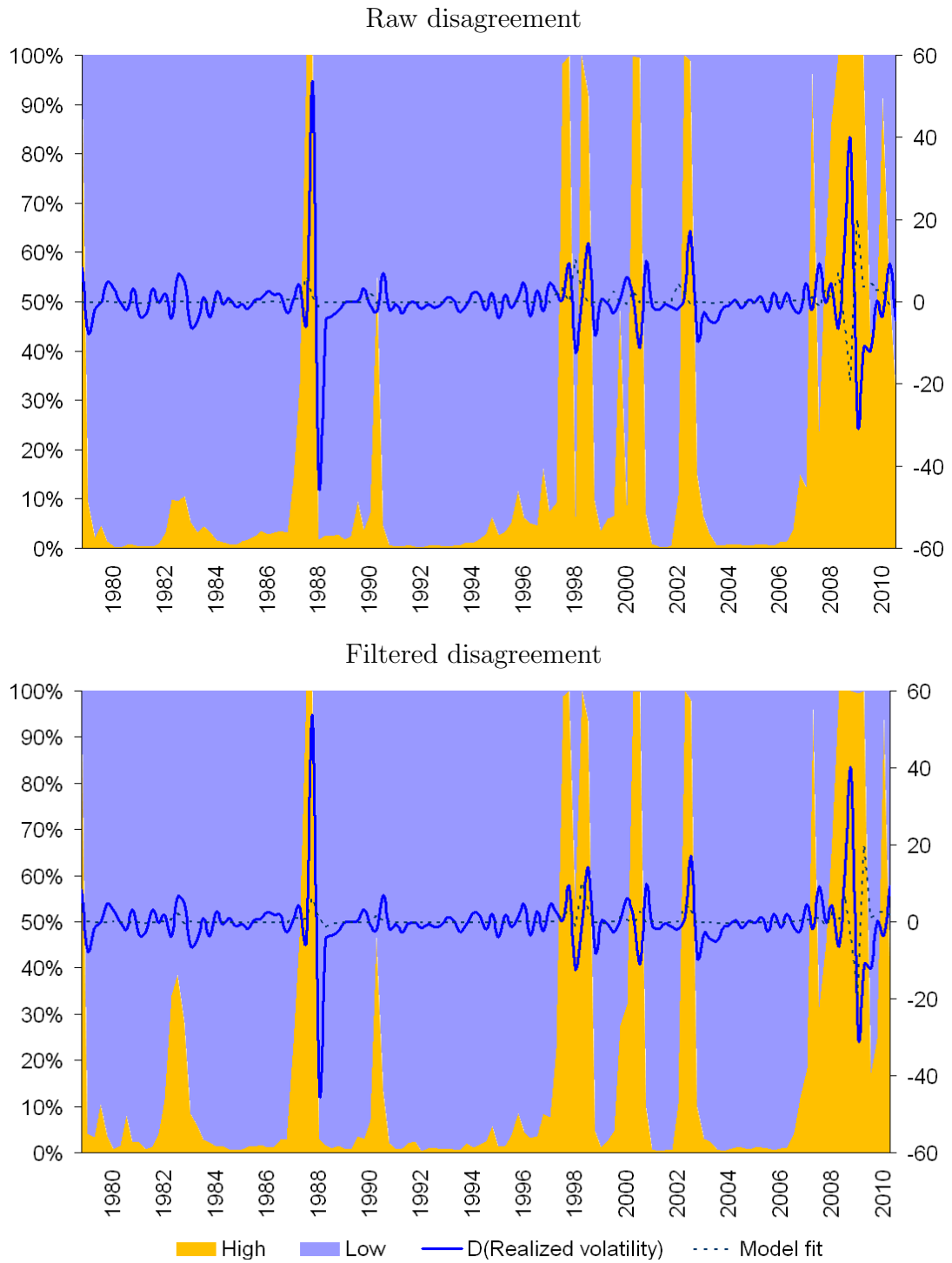
Figure 6: Endogenous transition probability estimates - GDP growth model
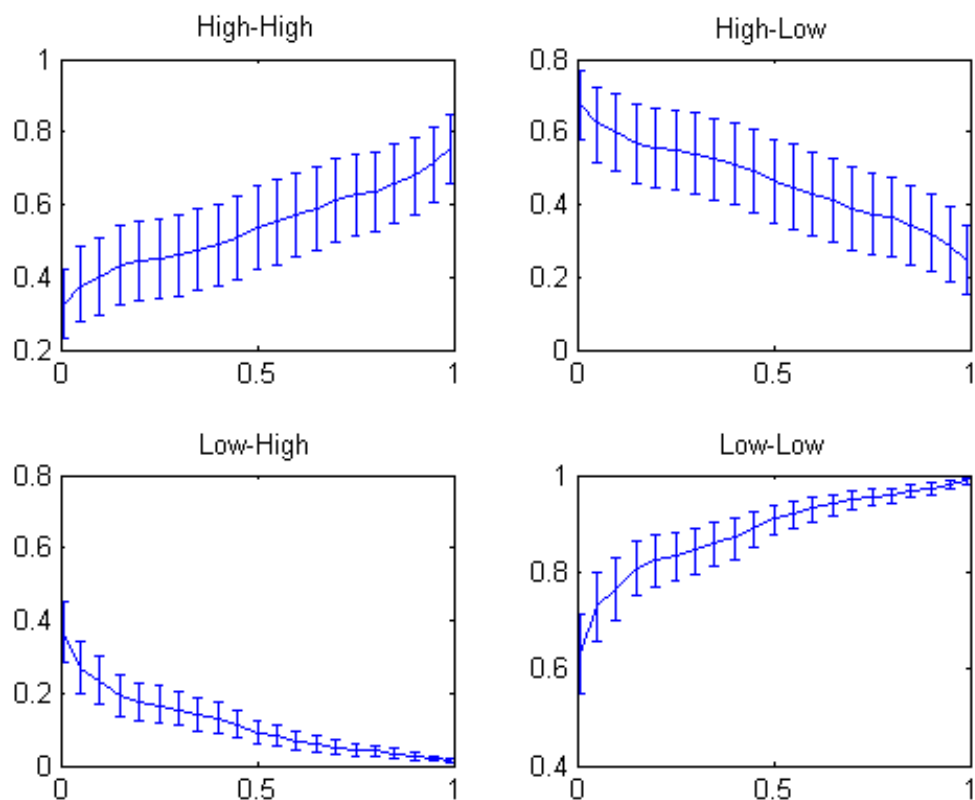(including raw disagreement)
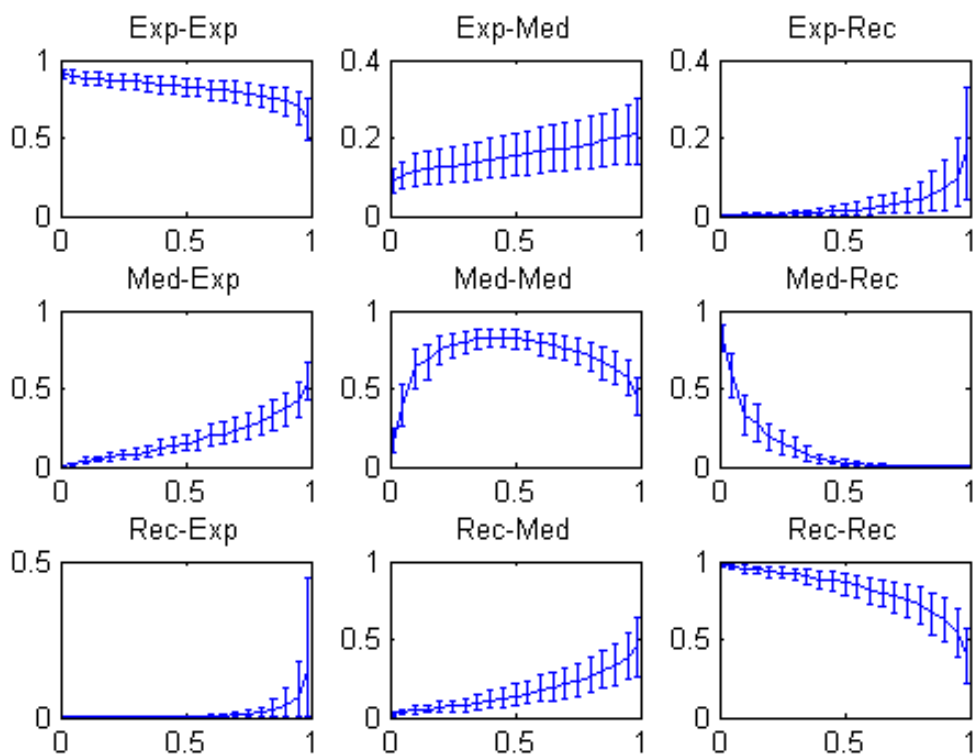
Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement). Vertical axes denote the transition probability estimates. Error bounds mark the 10% and 90% confidence levels, respectively.

Figure 7: Endogenous transition probability estimates - GDP growth model
(including filtered disagreement)
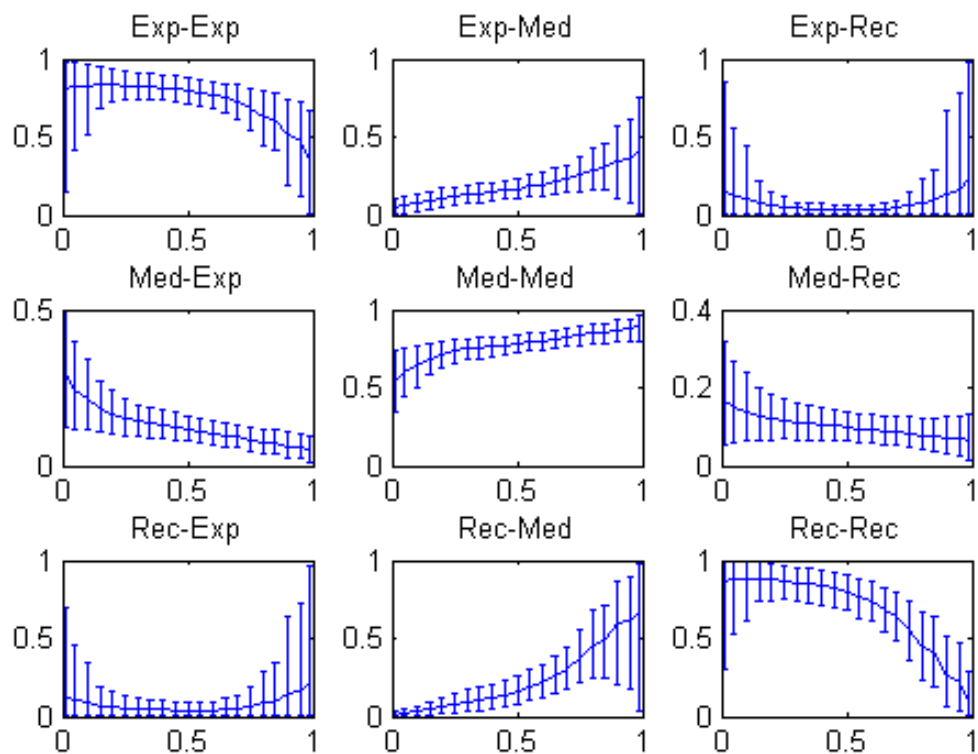
Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement). Vertical axes denote the transition probability estimates. Error bounds mark the 10% and 90% confidence levels, respectively.

Figure 8: Endogenous transition probability estimates - Realized volatility model (including raw disagreement)
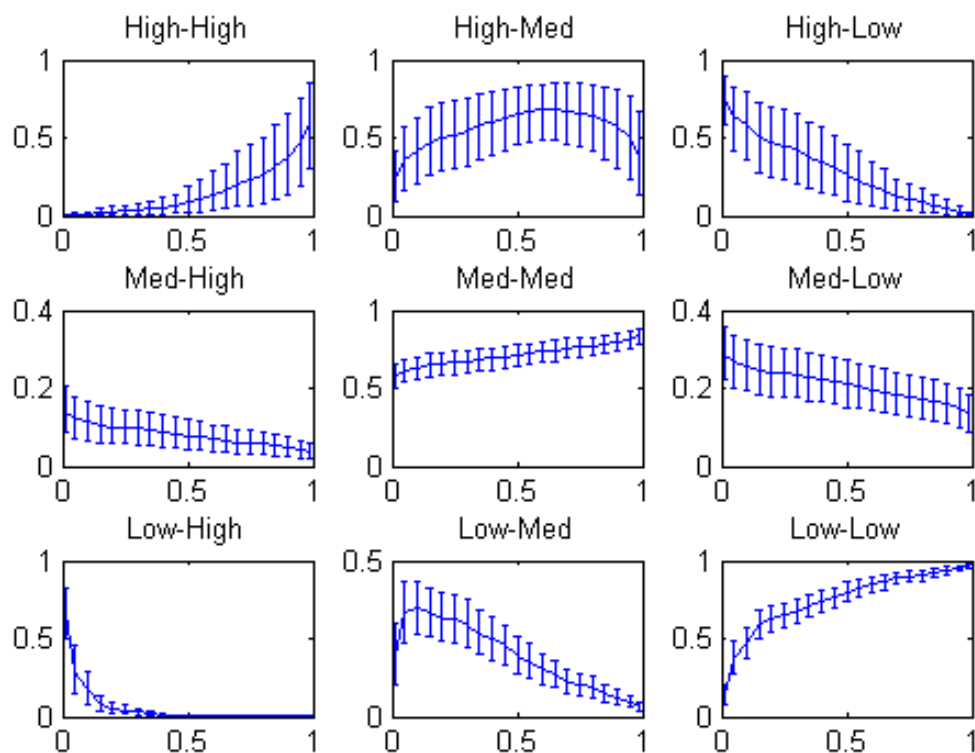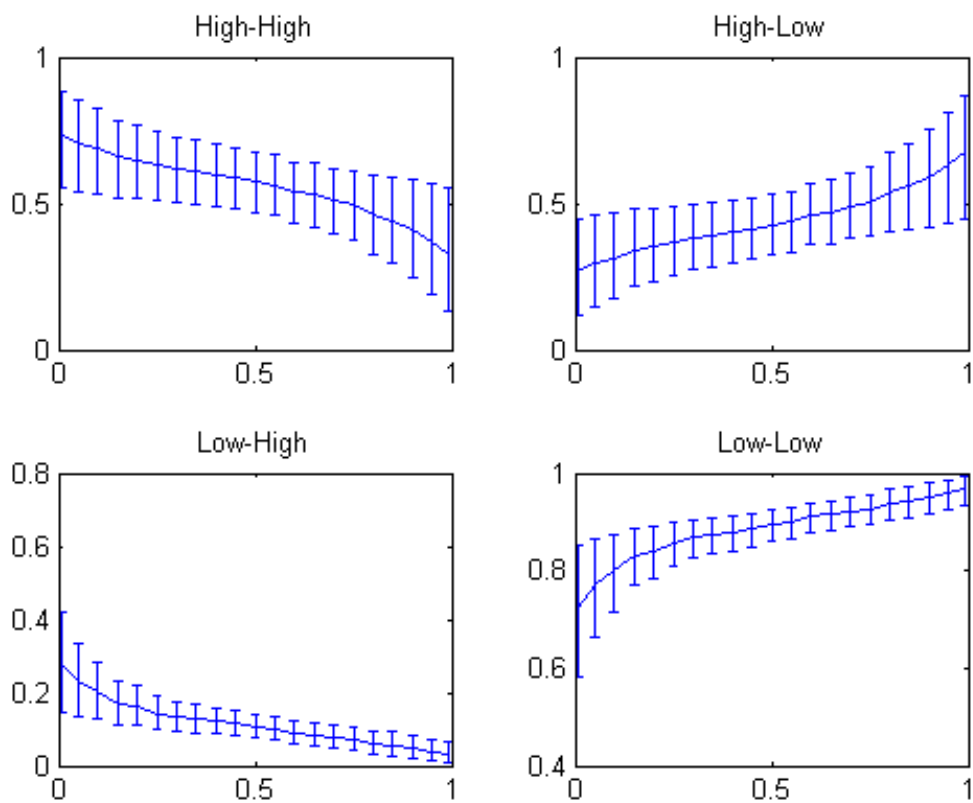

Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement). Vertical axes denote the transition probability estimates. Error bounds mark the 10% and 90% confidence levels, respectively.

Figure 9: Endogenous transition probability estimates - Realized volatility model
(including filtered disagreement)
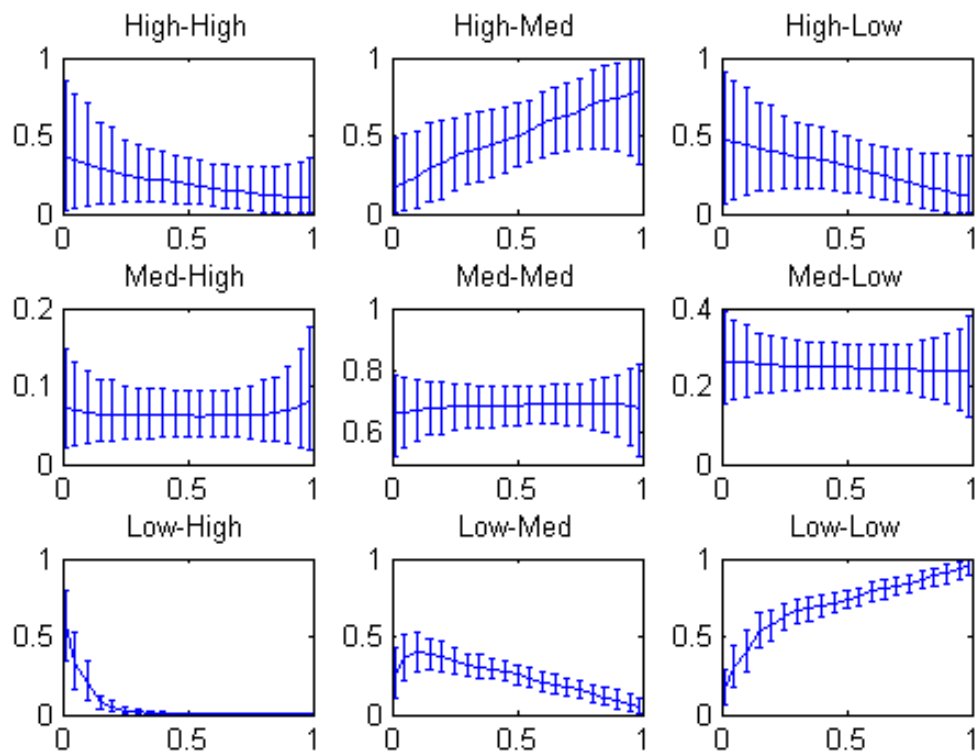
Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement). Vertical axes denote the transition probability estimates. Error bounds mark the 10% and 90% confidence levels, respectively.

Figure 10: Endogenous transition probability estimates - Model for realized
volatility in first differences (including raw disagreement)

Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement).
Vertical axes denote the transition probability estimates. Error bounds mark the 10%
and 90% confidence levels, respectively.

Figure 11: Endogenous transition probability estimates - Model for realized volatility in first differences (including filtered disagreement)


Note: Horizontal axes refer to the agreement level (0=agreement, 1=disagreement). Vertical axes denote the transition probability estimates. Error bounds mark the 10% and 90% confidence levels, respectively.

Figure 12: Path and distribution of simulated GDP growth model
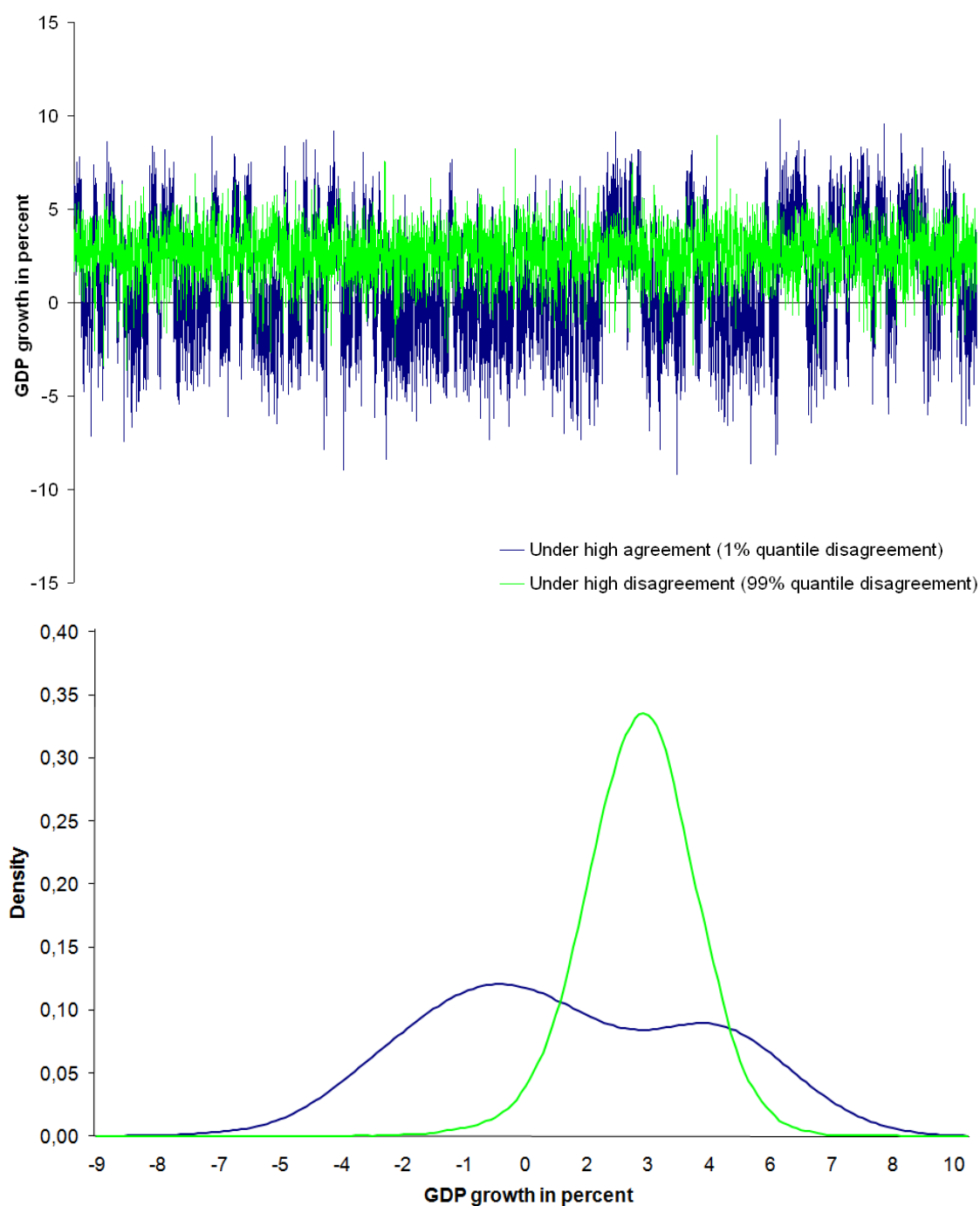
Note: The upper panel shows a sample of 10,000 periods simulated from the GDP growth regime-switching model presented in Table 3. An Epanechnikov kernel has been used to obtain a smooth estimate of the probability densities, which are presented in the lower panel.

Figure 13: Path and distribution of simulated realized volatility model
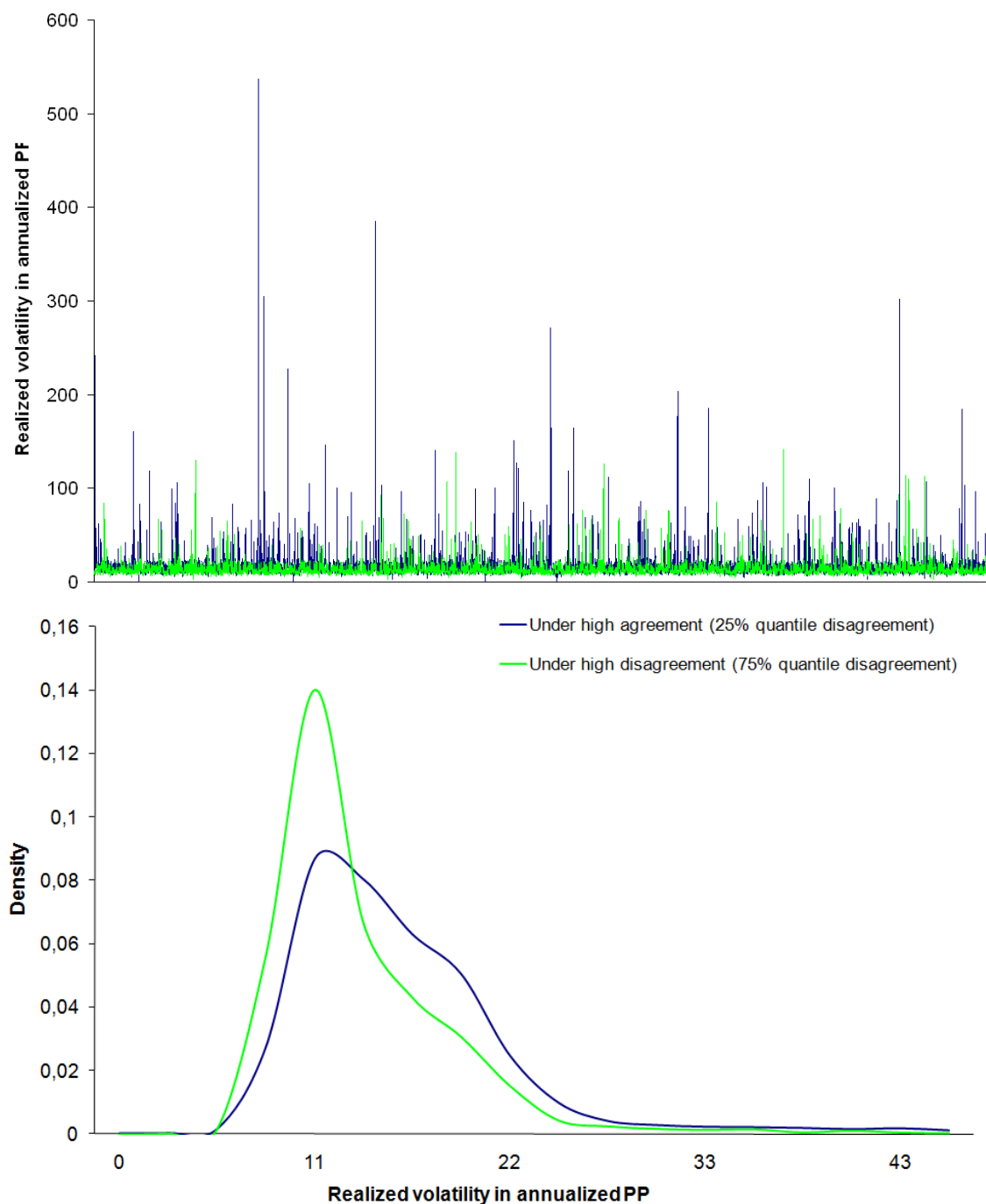
Note: The upper panel shows a sample of 10,000 periods simulated from the realized volatility regime-switching model presented in Table 5. An Epanechnikov kernel has been used to obtain a smooth estimate of the probability densities, which are presented in the lower panel.