Credit and Capital Markets, 49. Jahrgang, Heft 2, Seiten 193–220
Abhandlungen

A Markov Switching Approach to Herding

Martin T. Bohl, Arne C. Klein and Pierre L. Siklos*

Abstract

Existing models of herding suffer from the drawback that conventional measures assume it is constant over time and independent of the state of the economy. This paper proposes a Markov switching herding model which supports the view that herding is not constant. By means of time-varying transition probabilities, the model is able to link changes in herding behavior with proxies for sentiment, the VIX, and the term structure. For the US stock market our estimates reveals that during periods of high volatility, investors disproportionately rely on fundamentals rather than on market consensus. Existing theory supports such a conclusion. Some policy implications are also drawn.

Ein Markov-Switching Modell zur Analyse von Herdenverhalten

Zusammenfassung

Bestehende Modelle für Herdenverhalten haben den Nachteil, dass diese die Messvariablen über die Zeit als konstant und als konjunkturunabhängig betrachten. Die vorliegende Arbeit schlägt ein Markov Switching Modell vor, das das Herdenverhalten zeitvariierend modelliert. Mithilfe von zeitvariierenden Übergangswahrscheinlichkeiten können Veränderungen im Herdenverhalten mit *Approximationsgrößen* für die Marktstimmung, dem VIX und der Zinsstruktur

* Corresponding author: Martin T. Bohl, Department of Economics, Westfälische Wilhelms-University Münster, Am Stadtgraben 9, 48143 Münster, Germany, phone: +49 251 83 25005, fax: +49 251 83 22846, e-mail: martin.bohl@wiwi.uni-muenster.de.

Arne C. Klein, Department of Economics, Westfälische Wilhelms-University Münster, Am Stadtgraben 9, 48143 Münster, Germany, e-mail: arne.klein@wiwi.uni-muenster.de.

Pierre L. Siklos, Lazaridis School of Business & Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON, N2L 3C5, Canada, e-mail: psiklos@wlu.ca.

An earlier version of the paper was presented at the Northern Finance Association Annual Conference. We are grateful for financial support from a CIGI Collaborative research grant. We thank an anonymous referee for helpful comments on an earlier draft of this paper.

verknüpft werden. Für den US-amerikanischen Aktienmarkt zeigen unsere Schätzungen, dass Investoren in Zeiten hoher Volatilität unverhältnismäßig stark auf die Fundamentaldaten vertrauen und nicht auf den Marktkonsens. Diese Schlussfolgerung wird durch bereits bestehende Studien unterstützt. Politische Implikationen werden ebenfalls aufgezeigt.

Keywords: Herding Behavior, Markov Switching, US Stock Market

JEL Classification: G1

I. Introduction

Herding behavior has been the subject of considerable interest over the years. Nevertheless, there is no consensus on the correct definition to describe this kind of phenomenon. The present paper adopts an empirical definition. Hence, we observe differential amounts of dispersion in stock price movements depending on whether investors imitate each other when making investment decisions. As *Graham* (1999) observes, however, herding can be defined in several ways depending on its proximate source. For example, information cascades, changes in reputation, or when a first mover is followed by others all depict behavior consistent with herding behavior. All of these forms push investors to make similar choices for different reasons.

The extent to which investors discriminate between stocks is ordinarily believed to be reflected in how returns deviate from overall market performance. If investors follow the market, then dispersion in returns should disappear entirely. It is widely observed, however, that following the market may be conditional on, for example, whether the overall market is rising or falling. More importantly, one would expect market sentiment and volatility, as well as macroeconomic or financial conditions to have a significant influence on the extent to which investors follow the market.

Existing empirical studies that investigate herding behavior suffer from a number of drawbacks. Possibly the most important of these includes the failure to consider that herding can change over time as market conditions change. With a few exceptions only, for example, *Gebka* and *Wohar* (2013), *Stavroyiannis* and *Babalos* (2013), *Stavroyiannis*, *Babalos* and *Zarangas* (2013),[1] herding measures are generally treated as

[1] It should be noted, that other studies extend the notion of herding to cross-sectional characteristics of market participants to analyze herding behavior over time (*Bernhardt*, *Capello* and *Kutsoati* 2006, *Tang* 2013, *Naujoks* et al. 2009).

constant through time. While this is certainly a possibility this view is economically unreasonable especially since market sentiment varies such as when business cycle conditions change as between recessionary and expansionary periods.

The paper also notes that a unit root in a time series representing dispersion can be rejected in favor of the alternative of stationarity when we drop the assumption that the sample in question consists of a single regime. This has implications for the behavior of stock market returns over time and suggests that herding is time-varying. Indeed, we conclude that the data are reasonably well described when we assume that two regimes exist, each with stationary returns. Similar to the behavior of the VIX, one regime displays high market volatility while in the other regime volatility is relatively lower. The aftermath of the dot com bubble represents a highly volatile regime, as does the period of the 2007–2010 financial crisis. Much of the remainder of the sample considered is broadly characterized as being in the relatively low volatility state.

Existing models of herding behavior cannot be estimated from daily or weekly data, or are incapable of accommodating factors that determine investors' propensity to display herd-like behavior. Hence, it is not surprising that the extant literature treats herding as being constant over time.

Our study makes a start towards overcoming these deficiencies by specifying a Markov switching model based on the approach proposed by *Chang*, *Cheng* and *Khorana* (2000). We apply this model to daily data from the US stock market. Time-varying transition probabilities as derived in *Diebold*, *Lee* and *Weinbach* (1994) enable us to take account of additional economic and financial variables that drive changes in herding behavior over time. In particular, proxies for market sentiment, such as implied volatility and trading volume, the VIX, as well as term structure variables, which the literature considers to be closely linked to macroeconomic fundamentals, prove to be useful in this respect. Finally, because of some well-known properties of returns, non-normal distributions and generalized autoregressive conditional heteroskedasticity (GARCH) effects are also accounted for.

The remainder of the paper is organized as follows. The next section provides a brief review of the herding literature. Classical approaches to measure herd formation are described in the third section. Then, Markov switching models of herd behavior are introduced. The data are described in section IV. Section V. summarizes our empirical findings. The two-regime augmented Dickey-Fuller (ADF) test, useful in motivating the need to treat herding as being time-varying, is also briefly sketched here. The paper ends with a brief conclusion and suggestions for future research.

II. Survey of the Literature

The literature defines herding as a situation where investors imitate each other's buy and sell decisions, even though this kind of trading strategy might be at odds with their own information and beliefs. *Graham* (1999) points out that herding may also be initiated though information cascades as when investors act on information even when this contradicts their own beliefs. Alternatively, when an investor's reputation grows this too may produce herd like behavior. Notice that, in one case, investors are led to make the same choices based on the observable performance of others while, in the other case, herding is prompted by imprecise information. Other forms of herding also exist as when the decision by an investor to move first leads others to follow. Nevertheless, all forms of herding have empirical implications that can be seen by how far stock returns deviate from overall market returns. This paper interprets herding behavior in this sense. *Graham* also notes that the various definitions of herding are not mutually exclusive.

Still other definitions of herding have been proposed. For example, spurious herding refers to a "clustering" of investment decisions due to similar underlying information sets. Herding behavior can be either rational or irrational (*Devenow* and *Welch* 1996; *Bikhchandani* and *Sharma* 2001). Pure irrational herd behavior is closely related to the theory of noise trading (*De Long* et al. 1990 and 1991; *Jeanne* and *Rose* 2002), which assumes that a group of investors act irrationally or base investment decisions on some exogenous concerns over liquidity combined with some limits to arbitrage.

In contrast, information-based herding rests on the presumption that investors face uncertainty about the quality or precision of the information they are able to access. Although information cascades attempt to address this kind of behaviour (*Bikhchandani*, *Hirshleifer* and *Welch* 1992; *Welch* 1992; *Banerjee* 1992) *Avery* and *Zemsky* (1998) are the first to propose a model that is applicable to the case of financial markets. However, even for an investor who has access to superior private information, it might be rational to ignore this information and to rely on herding, for example, as in the case of portfolio managers facing incentives to stick with a benchmark (*Scharfstein* and *Stein* 1990; *Froot*, *Scharfstein* and *Stein* 1992; *Graham* 1999).

The empirical literature on herding to which this study contributes can be sub-divided into two branches. The first deals with herding among institutional investors such as fund managers. Research of this kind relies on data about their trading behaviour. Work on this topic is mainly based upon the measure proposed by *Lakonishok*, *Schleifer* and *Vishny* (1992), who compare the actual share of managers' buy and sell decisions against the expected values under the assumption of independent trading.

A second strand of research deals with herding towards the market. This is explained by investors who base their investment decisions entirely on market consensus, thereby ignoring their own beliefs about the risk-return profile of particular stocks. *Christie* and *Huang* (1995) are the first to address this issue empirically. They test the conjecture that such a trading pattern is more likely to arise during times of market stress as evidenced by unusually high volatility. However, their evidence for the US market does not corroborate a significant clustering of returns during strong market movements.

Unlike *Christie* and *Huang* (1995), the approach proposed by *Chang*, *Cheng* and *Khorana* (2000) does not neglect investors' behavior during periods of low or average volatility. Their test specification aims to compare the actual dispersion of single stock returns around the market with the value implied by rational asset pricing. In particular, they exploit the fact that those pricing models imply a linear relationship between the absolute value of the market return and its dispersion. Their findings support an increased tendency to herd in emerging markets but reveal only little evidence for such a behavior in developed countries.

Tan et al. (2008) investigate herding in Chinese A and B stocks using the approach of *Chang*, *Cheng* and *Khorana* (2000).[2] They find evidence for herding in both the A stocks available for domestic investors and in the B shares that are dominated by foreign investors. Analyzing the Polish stock market, *Bohl*, *Gebka* and *Goodfellow* (2009) highlight differences in trading patterns between individual and institutional investors. While the former display herding behavior, particularly during market downturns, the latter are unlikely to be driven by herd behavior. An application to the ETF market can be found in *Gleason*, *Mathur* and *Peter-*

[2] A shares trade in the Chinese currency, the renminbi, while B shares are denominated in a foreign currency (e.g., US or Hong Kong dollars).

son (2004). They estimate the models of *Christie* and *Huang* (1995) and of *Chang*, *Cheng* and *Khorana* (2000) from New York Stock Exchange intraday data and find strong evidence for adverse herding in this market. Adverse herding refers to a situation where, unlike the case of herding, investors disproportionately discriminate strongly between individual stocks.

Unlike mutual funds, ETFs trade on exchanges. Nevertheless, *Clifford* et al. (2014) conclude that ETF investors seek returns much in the same way that mutual fund investors do. Since volatility regimes play an important role in the empirical analysis to follow it is worth noting that ETFs contribute to raising the price volatility of the underlying stocks but this need not be explained by the process of price discovery (e.g., see *Ben-David* et al. 2014). An alternative explanation is that ETFs moves markets and that greater volatility is associated with higher returns (*Caginalp*, *De Santis* and *Sayrak* 2014). Of course, ETFs have grown substantially in recent years, particularly in the US and they may impact the co-movement of stocks (*Da* and *Shive* 2013).[3]

The papers cited above deal with herd behaviour within a given market and ignore potential international linkages in account. *Chiang* and *Zheng* (2010), however, investigate the impact of the US market on herding formation in several stock markets around the world. They provide favourable evidence that both the volatility as well the cross-sectional dispersion of single stock returns in the US influence herding activities in the rest of the world. In contrast, *Tan* et al. (2008) are unable to find interactions between the herding behaviour in the Chinese stock markets in Shanghai and Shenzhen.

Hwang and *Salmon* (2004) are the first to derive a measure of herding that allows for time variation in herding behavior. Their approach is based on the assumption of time-varying monthly betas. Results for the US and South Korea show a tendency of herding to mitigate, or even become adverse, in the run-up to and during periods of turmoil as in, for example, the Asian and Russian financial crises as well as the tech bubble of the early 2000s in the US. In order to establish a theoretical rationale for these facts, *Hwang* and *Salmon* (2009) put forward a testable model that incorporates the effect of investor sentiment. In this frame-

[3] See *Wurgler* (2011) and *Anton* and *Polk* (2014) for surveys of the impact of institutional investors on expected stock returns and the implications for the correlation of returns.

work, herding occurs in a situation when investors broadly agree about the future direction of the market, whereas adverse herding is likely to arise when there is a high probability of divergences of opinion among market participants. These interpretations of herding behavior also play a role in the empirical analysis below.

III. Constant and Time-Varying Herding Models

Research on herding rests on the seminal work of *Christie* and *Huang* (1995). Their approach considers the dispersion of single stock returns around the market. They propose the following measure:

$$S_t = \frac{1}{N(t)} \sum_{i=1}^{N(t)} \left| r_{i,t} - r_{m,t} \right|, \tag{1}$$

where $N(t)$ is the number of stocks available at time t with T observations in the sample, $r_{i,t}$ stands for the return of stock i and $r_{m,t}$ for the market return, respectively.[4] The market, in turn, is defined as a value-weighted average of single stock returns. Equation (1) is designed to measure the average absolute deviation of single stock returns from the market return and, thus, provides insights into the extent to which market participants discriminate between individual stocks. If all investors act alike and follow the market, S_t must be equal to zero.

Christie and *Huang* (1995) then regress S_t upon a constant and two dummy variables that control for both extreme positive and negative returns, measured by certain outer quantiles of the return distribution. Although very clear-cut, this approach obviously depends heavily on the definition of the thresholds for extreme returns. In addition, differing investor behavior during times of low and average volatility is completely neglected.

The extension put forward by *Chang*, *Cheng* and *Khorana* (2000) aims to overcome these drawbacks. They highlight the notion that, under the assumption of rational asset pricing, i.e., CAPM-type pricing, equation (1) is linear and strictly monotonically increasing in the expected value of the absolute market return, $E(|r_{m,t}|)$. By contrast, herding behavior is

[4] Actually, *Christie* and *Huang* (1995) use (1) only as a robustness check and base their main inference upon the cross-sectional standard deviation. The advantage of the absolute deviation (1) over the standard deviation is that the former is less sensitive to outliers.

better captured by a function that is either non-linear or reaches a maximum at a certain threshold value of $E(|r_{m,t}|)$, declining thereafter. The following regression is designed to capture these effects:

$$S_t = \gamma + \delta \, | r_{m,t} | + \zeta r_{m,t}^2 + \varepsilon_t, \tag{2}$$

where the realized market return is used to proxy their expected value. Rational asset pricing implies a significantly positive δ and a ζ equal to 0. By contrast, a value of ζ that significantly differs from 0 indicates a violation of the linearity implied by rational asset pricing. Using daily returns, this implies that $Var(r_{m,t}) = E(r_{m,t}^2) - E(r_{m,t})^2 \approx E(r_{m,t}^2)$ holds, so that $r_{m,t}^2$ can be regarded as the market return variance. If, during periods of high volatility investors herd towards the market, this implies that the dispersion of returns around the market S_t becomes disproportionally low compared to the rational pricing model. This should show up as a negative coefficient for ζ.

In all of the models outlined so far, herding behaviour is assumed to be constant over time. However, as noted above, the extant literature relates herding to investors' sentiment (*Shiller*, *Fisher* and *Friedman* 1984; *Lee*, *Shleifer* and *Thaler* 1991; *Devenow* and *Welch* 1996; *Hwang* and *Salmon* 2009) which, by definition, is time-varying. Therefore, the assumption that herding is constant does not seem reasonable. Furthermore, it is conceivable that, due to the crisis-laden environment prevailing during the last decade, including the tech bubble, 9/11 and the most recent global financial crisis of 2008–2009, the dispersion measure, S_t, is unlikely to be stationary, as required, in the single regime setting consistent with constant hedging. Instead, herding might be better characterized by a two-state model allowing for different dynamics in tranquil and volatile periods.[5]

To account for time-varying effects, *Hwang* and *Salmon* (2004) propose the following state space model, which, while similar in spirit, does not directly make use of the dispersion measure (1). First of all, the model assumes that market betas are changing over time. Inference about herding can then be obtained from the cross-sectional standard deviation of the betas. For instance, a situation where the betas of all stocks in the

[5] In principle, herding can vary over several states. However, since the paper aims to demonstrate that herding is time-varying and the assumption of two states (e.g., recessions versus expansions) is commonly employed in the finance literature (*Naes*, *Skjeltorp* and *Oodergaard* 2011), we restrict our investigation to this case. See, however, below.

market are approaching the value one implies that this cross-sectional standard deviation gets close to zero. In contrast, when all investors disproportionately strongly differentiate between stocks, such that the betas more strongly diverge from one than is implied by the CAPM equilibrium condition, this is referred to as adverse herding and would lead to a higher standard deviation.

For these reasons, *Hwang* and *Salmon* (2004) estimate standard OLS betas on a monthly basis in a first step. In a second step, the cross-sectional standard deviation of these betas is calculated for all periods. The deviation is then modelled within a state space framework where the changes in the dispersion of the betas are governed by a latent herding variable. Assuming an AR(1) process describes its movements, the latter can be extracted by using the Kalman filter.

Although the above approach produces a continuously evolving herding variable, it suffers from several disadvantages. First, the model cannot be estimated from daily or weekly data but relies on monthly beta estimates. Monthly betas, however, are strongly driven by "noise," for example, during periods of substantial financial turmoil, such as in the case of the recent financial crisis. Reducing noise requires expanding the estimation period for the market betas, which, in turn, reduces the number of observations for the state space model. Furthermore, if herding dynamics actually take place in the very short run, say on a daily or weekly basis, the model cannot capture the sought after phenomenon. Second, the model is unable to link changes in herding to investor sentiment or macroeconomic fundamentals. Third, assuming a zero mean for the latent herding variable, the model, by definition, implies swings between herding and so-called adverse herding. Thus, this measure is unable to describe a market where investors are switching between herding, no herding or adverse herding forms of behavior. In contrast, a Markov switching adaptation of equation (2) aims to remedy these problems.

1. Markov Switching Herding Measures

Our principal aim is to model time-varying herd behavior, rely on daily data and, additionally, to allow variations in herding to be driven by exogenous variables that capture changing market sentiment and macroeconomic and financial market fundamentals. A straightforward way of introducing time-varying behavior is to assume that it is subject to regime switches. As argued above, we assume there are two regimes. One associated with high volatility, the other is characterized by low volatility. Equation (2) is modified to allow for switching between two regimes, $j \in \{1, 2\}$:

$$S_t = \gamma_j + \delta_j \mid r_{m,t} \mid + \zeta_j r_{m,t}^2 + \varepsilon_{j,t}, \tag{3}$$

where $\varepsilon_{j,t} \sim N(0, \sigma^2)$ and the other variables were previously defined. It is well known that financial time series often display leptokurtosis. Therefore, the model given in equation (3) is re-estimated allowing one or even both regimes to be governed by a fat-tailed distribution. To this end, the distribution is relied on as well as on the generalized error distribution (GED).[6] We assume the latent state variable to be driven by a first-order Markov process, with transition probabilities, $p_{ij,t} = \Pr(S_t = j \mid S_{t-1} = i)$, $i, j \in \{1, 2\}$, which can either be constant or time-varying. For the sake of inferring the regime the process is in at time t, based on all information available up to the end of the sample period, , smoothed probabilities $p_{i,t|T} = \Pr(S_t = j \mid \Gamma_T)$ were calculated as given in Kim (1994).

As stated previously, time-varying transition probabilities can provide insights into the factors driving changes in herding behavior over time. This means making $p_{11,t}$ and $p_{22,t}$ dependent on a set of exogenous variables X_{t-1} including a constant.[7] Variables suitable in explaining the switches in investors' herding behavior include investor sentiment and macroeconomic conditions relying on data available at the daily frequency.[8] Implied volatility, here measured using the Chicago Board Options Exchange Market Volatility Index (VIX), is used. Motivated by the branch of literature on sentiment (*Jones* 2002; *Baker* and *Stein* 2004; *Baker* and *Wurgler* 2006), the share turnover relative to market capitalization is also used.

---

[6] The GED may provide further insights into the distributional properties of the dispersion of single stock returns since, unlike the t distribution, it also allows for thinner tails than in the case of the normal distribution.

[7] These variables are lagged because the transition probabilities governing switches from $t-1$ to t must be determined at time $t-1$.

[8] Other candidate variables are, of course, also possible. For example, based on *Campbell*, *Hilscher* and *Szilagyi* (2008), variables representing firm performance may also be considered. Nevertheless, their work also highlights the important connection between the behavior of the VIX and the financial distress they seek to empirically measure. Since it is preferable, under the circumstances, to estimate our Markov Switching model with daily data some of alternative determinants are not available at the daily frequency.

Proxies for macroeconomic conditions can be derived from term structure data (*Estrella* and *Hardouvelis* 1991; *Estrella* and *Mishkin* 1997; *Estrella* and *Mishkin* 1998). *Litterman* and *Scheinkman* (1988) and *Knez*, *Litterman* and *Scheinkman* (1994) show that the variation in money as well as capital markets can be very well described by models that contain from one to four common factors. Based on zero bond returns, principal components analysis is used to extract common factors. Only those principal components with eigenvalues greater than one are included in X_{t-1}. This ensures that each factor has more explanatory power than any return series. If the coefficients are assembled in a vector θ_j, the transition probability associated with state j, $p_{jj,t}$ can be modelled as:

(4)
$$p_{jj,t} = \frac{e^{X'_{t-1}\theta_j}}{1 + e^{X'_{t-1}\theta_j}}.$$

Turning to the estimation procedures, the models that assume a normal distribution can be estimated using the expectation maximization (EM) algorithm (*Dempster*, *Laird* and *Rubin* 1977). A closed form solution for all parameters was put forward by Hamilton (1990), while the solutions for θ_j, the parameters for (4), are derived in *Diebold*, *Lee* and *Weinbach* (1994). The specifications using t and GED-distributed errors are also estimated using the EM algorithm. Unlike the case of the normal distribution, no analytic solutions for the regression parameters are available. Nevertheless, since the conditions for the closed-form solution for the transition probabilities, $p_{ij,t} = \Pr(S_t = j \mid S_{t-1} = i)$, given in *Hamilton* (1990) still hold, these can be calculated as a by-product of the smoothed probabilities, $p_{j,t|T}$. Thus, obtaining estimates for the remaining regression and distributional parameters requires a whole numeric optimization in each iteration of the EM algorithm relying on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.[9]

In the case of the t distribution, equation (3) is first estimated by assuming $E_{j,t} \sim t(0,\sigma_j^2)$, with υ_j regime-dependent degrees of freedom parameter governing the kurtosis. In principle, this parameter can take on any value in the region $]2,\infty[$. Nevertheless, it is well known that the t distribution is empirically indistinguishable from the normal one for degrees of freedom greater than 30 (*Hansen* 1994; *Jondeau* and *Rockinger*

[9] *Watanabe* and *Yamaguchi* (2004), and *Azzalani* and *Capitanio* (2014) are two sources that deal with the statistical properties of the various assumed error distributions considered here.

2003). Thus, if for state j, the estimate for υ_j takes on a value above 30, it again fits the model, this time with one regime being governed by a t and the other one by a normal distribution. When applying the GED distribution to the errors, $\varepsilon_{j,t} \sim GED(0,\sigma_j^2)$, a one-step estimation procedure can be followed since this distribution reduces to the normal for a tail thickness parameter, κ_j, equal to 1.

To account for autocorrelation, we make use of the covariance matrix proposed by *Newey* and *West* (1987) where a lag length equal to eight is set as suggested by the *Newey* and *West* (1994) criterion. Since the construction of this error matrix and the selection of the appropriate lag length rests on several assumptions that might be crucial for the results, a robustness check is conducted by performing the analysis based on different numbers of lags. Since the autocorrelations in S_t are in general found to be relatively large (*Chang*, *Cheng* and *Khorana* 2000), all models are re-estimated for 6, 10, 12 and 14 lags.

For some markets, studies report different herding dynamics during falling and rising markets (*Chang*, *Cheng* and *Khorana* 2000; *Bohl*, *Gebka* and *Goodfellow* 2009). In addition, evidence from fund managers trading reveals differences in their herding behaviour between buying and selling decisions (*Keim* and *Madhavan* 1995; *Grinblatt*, *Titman* and *Wermers* 1995). These phenomena are also accounted for by estimating an asymmetric version of the baseline model:

$$\begin{aligned} S_t = {} & \gamma_j + \gamma_j^{asy} I^{Rm,t<0} + \delta_j \mid r_{m,t} \mid + \delta_j^{asy} I^{Rm,t<0} \mid r_{m,t} \mid \\ & + \zeta_j r_{m,t}^2 + \zeta_j^{asy} I^{Rm,t<0} r_{m,t}^2 + \varepsilon_{j,t} \end{aligned} \tag{5}$$

where $I^{Rm,t<0}$ is a dummy variable that is equal to 1 if the market return is negative and equal to 0 otherwise and $E_{j,t} \sim N(0,\sigma^2)$. To keep the econometric model as simple as possible we make no attempt to model asymmetries in higher moments.

Finally, it is sensible for financial data which exhibit volatility clustering and skewness to control for ARCH effects as proposed by *Gray* (1996b). Therefore, equations (3) and (5) are estimated as Markov switching GARCH(1, 1) (MSGARCH(1, 1)) models. Again for sake of simplicity we rely on the simple, but well-established GARCH(1,1) specification. The first lag of S_t is included, to take into account autocorrelation since the *Newey* and *West* (1987) errors cannot be used for GARCH models. To model skewness, a skewed t distribution, is applied as proposed by *Fernandez* and *Steel* (1998). The density function is given as follows:

$$f_t(\varepsilon_{j,t}) = \frac{2\beta_j}{1+\beta_j^2}\left(t(0,\beta_j\varepsilon_{j,t},\upsilon_j)I^{\varepsilon_{j,t}<0} + t(0,\frac{\varepsilon_{j,t}}{\beta_j},\upsilon_j)(1-I^{\varepsilon_{j,t}<0})\right) \tag{6}$$

where $I^{\varepsilon_{j,t}<0}$ is an indicator function that is equal to 1 if $\varepsilon_{j,t}$ is negative and equal to 0 otherwise. $\beta_j > 1$ indicates a distribution that is skewed to the right while β_j is smaller than one in case of a left skewed density. For $\beta_j = 1$, (6) reduces to a standard t distribution. The MSGARCH(1, 1) model is estimated using numerical optimization according to the BFGS algorithm. As the forward-looking algorithm provided in *Gray* (1996a) is used to calculate smoothed probabilities, $p_{j,T} = \Pr(S_t \mid \Gamma_T)$, this approach can also be considered as a robustness check for *Kim's* (1994) smoother.[10]

IV. Data

The analysis covers the entire US stock market for the period 2001–2010. The 2001 to 2003 period captures the fallout from the dot com bubble while the 2007 to 2010 period marks the era of the so-called global financial crisis.[11] Figure 1 plots the estimates for S_t for the full sample while the vertical bars delineate the beginning and end of recessions as dated by the NBER (http://www.nber.org/cycles.html). Total returns were obtained for all listed stocks and a capitalization weighted market index from the Center for Research in Security Prices (CRSP) at the University of Chicago. To ensure that the results are not sensitive to the selection of the sample period, our models were also estimated for the periods 1999–2010 and 2003–2010. The second sample omits the period of the 2001 tech bubble period. The analysis is also carried out based on a sample that is free of cross-listings, listings in foreign currencies, stocks from minor exchanges, ETFs and preferred stocks as well as stocks that are not marked as major securities.

Additional data were obtained from Thomson Reuters Datastream. The principal components of the term structure are extracted using the Datastream Zero Curve with maturities of 0, 3, 6 and 9 months as well as

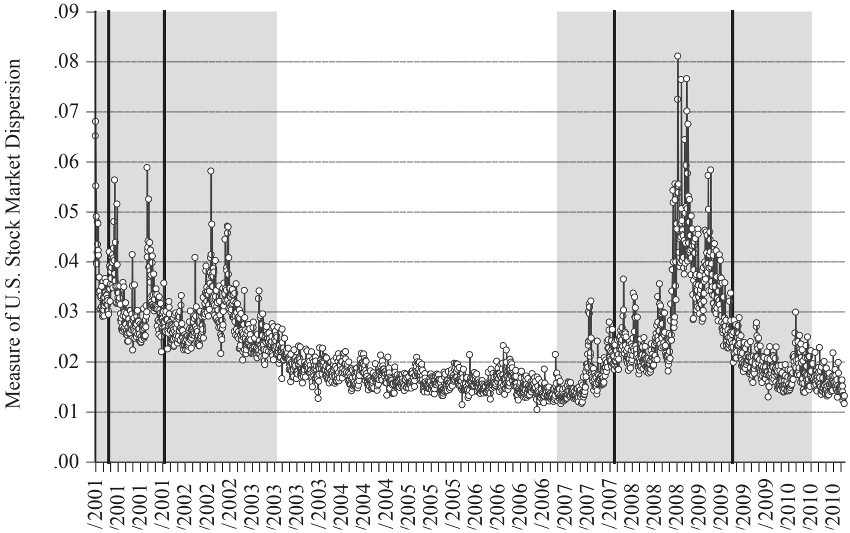
[10] To control for potential over-parameterization, a simple MSGARCH(1, 1) is also estimated with normally distributed errors and without lagged dependent variables.

[11] Dating these events is, of course, imprecise. For the global financial crisis we relied on the St. Louis Federal Reserve's chronology (https://www.stlouisfed.org/financial-crisis/full-timeline#2011) which begins February 27, 2007. The end is dated when President Obama signs Dodd-Frank financial reform legislation on 21 July 2010.

1–10, 12, 15, 20, 25 and 30 years. Aggregated trading volume and market capitalization for the US-Datastream Market are employed, and the VIX is also obtained from Thomson Reuters Datastream. The term structure data is frequently used as a macro-financial indicator or anticipated future economic conditions while there is growing evidence that the VIX is not only a reliable indicator of financial volatility but is also liked to overall economic performance via its relationship to monetary policy (e.g., *Bekaert* and *Horeova* 2014, *Bekaert*, *Horeova* and *Lo Luca* 2010). Meanwhile the term structure or term spreads has long been a staple of macro and financial models because of its predictive ability for future economic conditions, especially recessions (e.g., see *Wheelock* and *Wohar* 2009, and references therein).

V. Empirical Results

Under rational asset pricing (see equation 2) S_t should be stationary. Hence, the absence of stationarity is an indication that the process has changed over time. Clearly, as shown in Figure 1, the dispersion of stock prices experiences a slow moving downward trend roughly from the end



Note: Time series of S_t, see equation (1) in the text. The shaded areas represent the dot com and global financial crises periods, respectively. The vertical dashed lines mark the beginning and end of recession dates, as published by the NBER.

Figure 1: Stock Market Dispersion in the U.S.

of the dot com period until approximately 2006, that is, shortly before the onset of the global financial crisis.

To investigate the stationarity properties of a time series, it is common practice to rely on a unit root test such as ADF tests. Typically, these tests ignore possible regime switching effects often present in financial time series. To take these effects into account, *Hall* and *Sola* (1994) and *Hall*, *Psadarakis* and *Sola* (1999) propose a Markov switching ADF test. Allowing for deterministic trending and a regime-depending variance, the test equation is given as follows:

$$\Delta S_t = \phi_j S_{t-1} + \sum_{d=0}^{D} \alpha_{d,j} t^d + \sum_{h=1}^{H} \rho_{h,j} \Delta S_{t-h} + \eta_{j,t}\,, \tag{7}$$

where $\eta_{j,t} \sim N(0, \pi^2)$. $j \in \{1,2\}$ again denotes the state the process is in at time t, while D = 0, 1, 2, 3 indicates the degree of the polynomial defining the deterministic trend. Obviously, if D = 0, (7) this reduces to a Markov switching ADF (MSADF) test with a constant. When D = 1, a regime-depending linear trend is added while, in case of D = 2, the trend can be changing and for D = 3, this trend may have a turning point. H indicates the number of lags included. Due to strong autocorrelations in S_t,[12] the maximal lag length is set at a relatively high value of 25 and then the number of lags is successively reduced until the coefficient of the last lag H is found to be statistically significant at the 10 percent level in at least one state.[13] The MSADF test is estimated using the EM algorithm.[14]

First, the stationarity properties of the time series of the cross-sectional absolute deviations given in (1) are considered. To this end, the ADF and the MSADF tests described above are applied to the series of dispersion, S_t. Dickey-Fuller test statistics are given in Table 1.

When the standard (single regime) ADF test is considered, a unit root can only be rejected by the version of the test that does not account for deterministic trending. By contrast, the two-state MSADF test rejects

[12] The first four autocorrelations are 0.97, 0.87, 0.86, and 0.84. The correlations remain high for various sub-samples (results not shown).

[13] As a robustness check, the procedure is also performed for a maximal lag length of 15.

[14] At the suggestion of a referee we also estimated a version of equation (7) assuming 3 regimes (results not shown) but could not obtain unique coefficients for a third regime.

Table 1

Dickey-Fuller Test Statistics

	Single State	2 States
$S_t = 1$		
$D = 0$	–2.990**	–10.145***
$D = 1$	–3.051	–15.116***
$D = 2$	–3.083	–18.047***
$D = 3$	–3.790	–17.751***
$S_t = 2$		
$D = 0$		–3.559***
$D = 1$		–3.851**
$D = 2$		–3.988**
$D = 3$		–4.330**

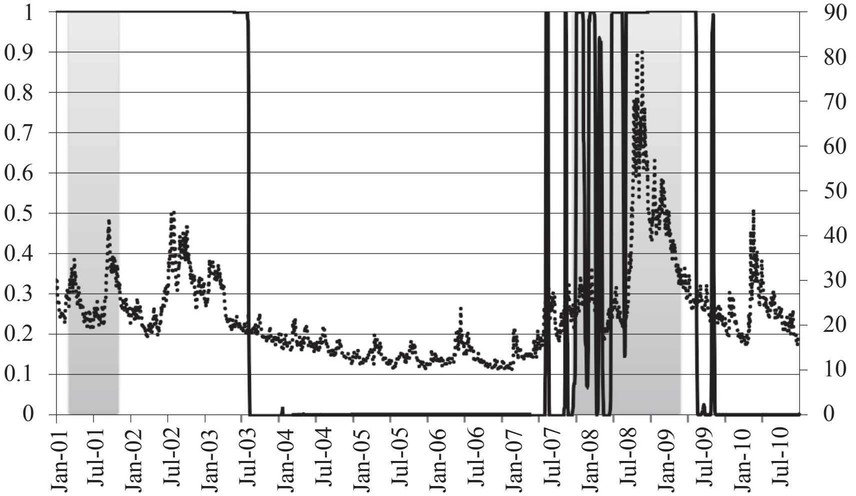
Notes: Single State refers to a standard ADF test where 2 States stands for the two regimes version of the test outlined in the section on Markov switching herding measures. The test statistic provided is the pseudo t-statistic. ***, **, and *, denote statistical significance at the one percent, five percent and 10 percent level, respectively. Significance is based on asymptotic critical values obtained by Monte Carlo simulation.

the null in both states and all specifications for the deterministic trend.[15] These findings corroborate the use of a time-varying herding measure since the assumption of constant herding is not only economically unreasonable but in the present context also ignores structural shifts in the time series dynamics of S_t such as when markets move from a low- to a high-volatility state.

Next, we turning to the baseline model, namely equation (3) with normal errors. The smoothed probabilities, $p_{i,t|T}$, which are plotted in Figure 2 (left-hand side scale), reveal two clearly distinct herding states. The smoothed probabilities are plotted against the VIX (right-hand side scale). It is immediately clear that the high volatility state typically coincides with deteriorating investors' sentiment, that is, a relatively high implied volatility.[16] The high-volatility regime can be related to periods

[15] A maximal number of lags, H, equal to 25 are used, but these results also hold for $H = 15$.

[16] Indeed, a plot of the VIX against S_t (not shown) suggests that the two are strongly positively correlated (simple correlation coefficient is 0.65 for the full sample).



Notes: VIX (dotted line) and smoothed probabilities calculated as proposed by *Kim* (1994). The smoothed probabilities are plotted on the left-hand side axis while the right-hand side displays the VIX. The shaded areas represent NBER recession dates.

Figure 2: Smoothed Probabilities and Implied Volatility

of large market movements and is characterized by a herding parameter, ˆ, that is significantly positive, indicating adverse herding. Unlike the case of herding, this indicates that investors differentiate more strongly between particular stocks than implied by rational asset pricing behavior. By contrast, the second regime seems to prevail during more tranquil times. Here, $\hat{\zeta}_2$, is found to be positive, but statistically insignificant, which is in line with CAPM-type models. Parameter estimates are reported in Table 2.

The high-volatility state initially prevails from the beginning of our sample in 2001, that is the bursting of the tech bubble, the time around 9/11 as well as the start of wars in Afghanistan and Iraq. Shortly thereafter, in mid-August 2003, a switch into the calmer regime takes place. This switch coincides with a strong US economy. The second regime then ends in mid-2007 when, around August 6, a switch into the high volatility state takes place, a couple of days before central banks around the world started to intervene in order to stabilize the money market at the onset of the financial crisis. Subsequently, the process switches several times between both regimes, consistent with uncertainty about the existence of a grave crisis prevailing among market participants during this period. Again, we also see this reflected in the behavior of the VIX. On

Table 2

Estimation Results for Four Herding Models

	OLS		Markov Norm		Markov t Norm		Markov GED	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
$S_t = 1$								
$\hat{\gamma}_1$	0.017***	(0.000)	0.025***	(0.000)	0.024***	(0.000)	0.023***	(0.000)
$\hat{\delta}_1$	0.549***	(0.043)	0.356***	(0.042)	0.375***	(0.015)	0.378***	(0.014)
$\hat{\zeta}_1$	0.622	(0.746)	1.676**	(0.735)	1.494***	(0.205)	1.510***	(0.197)
$\hat{\sigma}_1^2$	4.064 10^{-5}		3.536 10^{-5}		3.548 10^{-5}***	(0.000)	3.347 10^{-5}***	(0.000)
$\hat{\upsilon}_1/\kappa_1$					4.750***	(0.972)	1.615***	(0.107)
$\hat{p}_{11}$			0.991		0.991		0.991	
$S_t = 2$								
$\hat{\gamma}_2$			0.015***	(0.000)	0.015***	(0.000)	0.015***	(0.000)
δ_2			0.253***	(0.021)	0.253***	(0.012)	0.247***	(0.014)
ζ_2			0.873	(0.769)	0.987**	(0.440)	1.051**	(0.453)
$\hat{\sigma}_2^2$			4.042 10^{-6}		3.884 10^{-6}***	(0.038)	3.869 10^{-6}***	(0.000)
$\hat{\upsilon}_2/\kappa_2$							0.842***	(0.090)
$\hat{p}_{22}$			0.995		0.994		0.993	

Notes: Autocorrelation and heteroskedasticity consistent standard errors as proposed by Newey and West (1987) are provided in brackets. OLS refers to the basic ordinary least squares estimates while Markov norm, Markov t norm and Markov GED stand for the Markov switching models with normal, t-/normal and generalized error distribution, respectively. ***, **, and * denote statistical significance at the one percent, five percent and ten percent level, respectively. OLS estimates assume there is only one state.

September 2, 2008, a switch into the high-volatility state is indicated, a couple of days before Lehman Brothers released news about severe losses for the first time. In mid-August 2009, the process then moves back into the calmer regime and remains there until the end of the sample at the end of 2010.[17]

Applying the procedures described above to allow for fat-tailed distributions produces virtually unchanged inferences about regimes, the smoothed probabilities and very similar parameter estimates compared to the assumption of normality. The only substantial difference found is that, $\hat{\zeta}_j$, is significantly different from zero in both states. Hence, adverse herding is significantly stronger during volatile periods but remains significant in more tranquil market phases. It is conceivable that volatile periods create the conditions necessary for investors to place relatively greater weight on private information rather than follow the lead of other investors.

The estimates for the model allowing for both t and normally distributed regimes indicate that, during the high volatility state, the errors follow a t distribution while they are well characterized by a normal distribution during calm periods.[18] Taken together, these findings highlight that in the first state the dispersion of returns around the market and, hence, investor sentiment, are not only relatively volatile but also subject to large shocks.

The results for the asymmetric specification given in equation (5) suggest that herding in US stocks does not differ as much between market upturns and downswings since the t values for $\hat{\zeta}^{asy}$ are highly statistically insignificant. This is shown in Table 3. By contrast, the use of the MSGARCH model is corroborated by the data since strong volatility clustering and ARCH effects are found. This should not be surprising in view of the time series behavior of the dispersion variable (S_t). In addition, the values of $\hat{\upsilon}_j$ and $\hat{\beta}_j$ are found to be significantly different from

[17] Robustness checks using the sample periods 1999–2010 and 2003–2010 broadly confirm the findings. The 12-year period actually reveals that already in 1999 and 2000, the process is in the high volatility state. This supports the interpretation of the first regime as being closely related to periods of strong market movements rather than only strong downward movements or crises periods. Moreover, the results using different lag lengths in calculating the Newey and West (1987) covariance matrix broadly confirm the results.

[18] The use of the GED reveals that the calm period regime displays tails being even thinner than those of a normal distribution. Filtered and smoothed probabilities for the non-normal models are available upon request.

Table 3

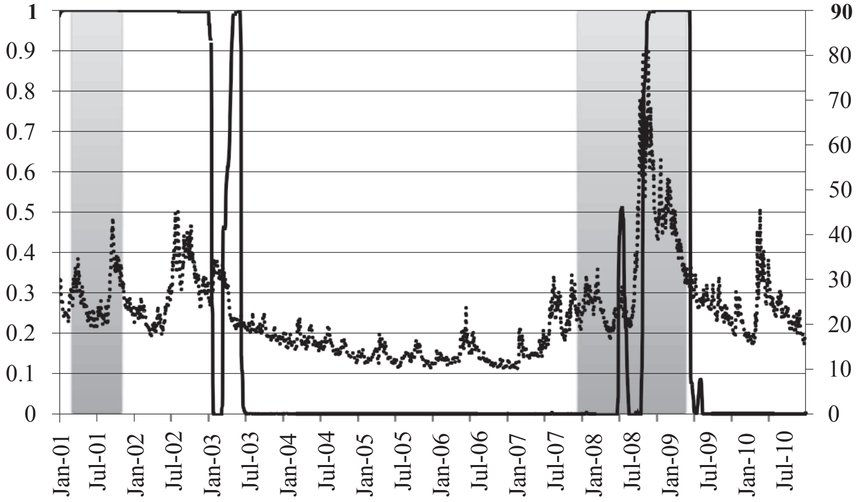
Estimation Results for the Asymmetric Herding Model

Markov norm asy		
Coeff.	Std. Error	
$S_t = 1$		
$\hat{\gamma}_1$	0.024***	(0.001)
$\hat{\gamma}_1{}^{asy}$	$8.643\ 10^{-4}$	($5.953\ 10^{-4}$)
$\hat{\delta}_1$	0.454***	(0.056)
$\hat{\delta}_1{}^{asy}$	–0.176***	(0.060)
$\hat{\zeta}_1$	1.248	(0.918)
$\hat{\zeta}_1{}^{asy}$	0.553	(0.798)
$\hat{\sigma}_1^2$	$3.411\ 10^{-5}$	
$\hat{p}_{11}$	0.991	
$S_t = 2$		
$\hat{\gamma}_2$	0.015***	($1.626\ 10^{-4}$)
$\hat{\gamma}_2{}^{asy}$	$3.523\ 10^{-4}$**	($1.750\ 10^{-4}$)
$\hat{\delta}_2$	0.275***	(0.023)
$\hat{\delta}_2{}^{asy}$	–0.057*	(0.034)
$\hat{\zeta}_2$	1.316*	(0.772)
$\hat{\zeta}_2{}^{asy}$	–0.308	(1.175)
$\hat{\sigma}_2^2$	$3.991\ 10^{-6}$	
$\hat{p}_{22}$	0.995	

Notes: Autocorrelation and heteroskedasticity consistent standard errors as proposed by *Newey* and *West* (1987) are provided in brackets. Markov norm asy refers to the Markov switching models with normally distributed errors given in equation (5). ***, **, and * denote statistical significance at the one percent, five percent and ten percent level, respectively.

the values implied by a normal distribution. The smoothed probabilities for the MSGARCH(1,1) specification with skewed t distribution, given in Figure 2, are even more clear-cut than those for the above models with constant variances. Again, we find adverse herding to be much stronger when the high volatility state prevails.[19]

[19] The smoothed probabilities for the MSGARCH (1, 1) with normal errors closely resemble those for the homoskedastic models and are available upon request, as are the parameter estimates for both MSGARCH models and the asymmetric specification. The results for the sample that is adjusted for minor stocks, ETFs, etc. are similar in spirit. The main difference is that the coefficients $\hat{\zeta}_j$ are



Notes: VIX (dotted line) and smoothed probabilities calculated as proposed by Gray (1996a). The smoothed probabilities are plotted on the left-hand side axis while the right-hand side displays the VIX. The shaded areas represent NBER recession dates.

Figure 3: Smoothed Probabilities for the Model MSGARCH(1, 1) Model with Skewed t Distribution and Implied Volatility

We now turn to a discussion of the approach relying on the time-varying transition probabilities,$p_{jj,t}$, now conditioned on additional determinants. These are conditioned which are made conditional on the VIX, the relative turnover and the principal components of the yield curve. The derivation of the principal components series was explained above. In what follows, only the first two components are of interest for the model since they are associated with eigenvalues greater than zero. Other interpretations, of course, are possible since they depend on the loadings of these components.[20] However, our selection is in line with the patterns known in the literature as shift and twist (*Litterman* and *Scheinkman* 1988; *Knez*, *Litterman* and *Scheinkman* 1994) components in the term structure. Indeed, the approach is also widely used in the recent literature that explores the impact of unconventional monetary policies (e.g.,

---

larger in absolute values. This suggests that adverse herding is stronger in the stocks of large and transparent firms.

[20] The components are linear combinations of the underlying zero bond rates and the respective parameters are referred to as loadings. There are always as many components as there are different zero rates.

Table 4

Parameters of the Time-varying Transition Probabilities

	$S_t = 1$	$S_t = 2$
Constant	–0.003	10.414
*PC*1	–1.651	–0.384
*PC*2	–2.603	0.361
VIX	0.225	–0.179
TURN	–124.680	–176.966

Notes: PC1 and PC2 refer to the first and second principle component of the yield curve while VIX and TURN stand for the Chicago Board Options Exchange Market Volatility Index and the relative volume measure defined by turnover in US$ divided by market capitalization times 1000.

Gürkaynak and *Wright* 2015). This means that the first component has very evenly distributed loadings and stands for a shift in the level of interest rates. By contrast, the second principal component is characterized by loadings, which monotonically decrease with a change in the slope of the term structure. In particular, a rise in short-term interest rates which is not accompanied by a proportional rise in long-term rates, or vice versa. So, starting from a normal yield curve, a rise in this factor is associated with a flattening of the term structure.

Using the principal components and the other two covariates in equation (4) and the logit specification for $p_{jj,t}$, the estimates for the parameters in (3) and the variances are found to be very close to those for the model with constant probabilities and a normal distribution. For this reason, Table 4 only reports $\hat{\theta}_j$, the parameters estimates for (4), the specification governing the transition probabilities.[21]

Those parameters estimates, which differ with respect to the signs between the states, are of particular interest since a change in a given exogenous variable always increases the probability for one regime while decreasing the one for the other regime, irrespective of the level of the covariates (see specification 4). Put differently, high values of such a variable could be linked to one state while below-average values would always be consistent with the other regime, independent of the behavior of other exogenous variables. While the parameters for the first principal

[21] It must be borne in mind that there are no standard errors for the parameters of the transition probabilities within the EM framework.

component representing the level as well as the turnover measure have a negative sign for both states, this is not the case for the second component and the VIX. The interpretation of the latter is straightforward, namely that state one is associated with a positive coefficient. Hence, an increase in the implied volatility makes it more likely to switch into or remain in regime one while the reverse holds for regime two.

More interestingly, the coefficient for the second principal component, which represents the slope of the term structure, is negative for the first state and positive for the second one. At first glance, this is counterintuitive since a flat (or even inverse) term structure is, in general, associated with a contracting economy. This should presage a switch into the high volatility. However, when the model is re-estimated employing two quarter lagged principal components, the sign turns negative for both states and, in particular, more negative for the second, the tranquil state.

VI. Conclusions

This paper proposes a time-varying model of herding behavior. We argue that extant empirical treatments of herding behavior, which often assume that herding behavior is constant, are inappropriate. Models that assume constant herding dynamics are economically implausible since the literature links herding to investor sentiment, among other economic and financial variables, that are by definition time varying. Unit root tests corroborate this view since one is only able to reject the unit root null unambiguously when the process is allowed to switch between two distinct regimes. Existing empirical models of herding behavior models of the time varying variety resort to data at monthly or even lower sampling frequencies. Hence, they cannot be used to generate evidence about investors' short-term behaviour.

The procedure proposed by *Christie* and *Huang* (1995) and *Chang*, *Cheng* and *Khorana* (2000) serves as our starting point to examine investors' herd behavior. However, their approach is modified by fitting a Markov switching model to allow for different dynamics between high and low volatility regimes. In addition, the time-varying transition probabilities proposed by *Diebold*, *Lee* and *Weinbach* (1994) are used to condition on economic and financial factors that may be used to explain changes in herd behavior over time driven by proxies for macroeconomic conditions and investor sentiment. Our paper is consistent with other approaches that take into account time herding behavior (*Gebka* and

Wohar 2013, *Stavroyiannis* and *Babalos* 2013, *Stavroyiannis*, *Babalos* and *Zarangas* 2013).

The Markov switching models are estimated for US stock market data, thereby controlling for non-normalities, autocorrelation and GARCH effects. The findings suggest that during times of high volatility in the market, investors discriminate more strongly between single stocks than during tranquil times and more strongly than implied by rational asset pricing models. In other words, there is considerable evidence of time-varying herding behavior and this can be associated with changes in economic conditions which themselves are related to other economic and financial variable such as the VIX or the behavior of the term structure of interest rates. As a general implication, our findings demonstrate the importance of using time varying parameter approaches when analyzing herding behavior of stock market participants. Furthermore, policy makers should be more aware that investor behavior does change when there is a significant shift in the current state of the economy. Therefore, regulatory responses should not assume that low volatility states are benign since herding is more likely to take place.

Needless to say, future research will have to consider a wider array of potential determinants of changes in herding behavior. For example, while the dot com and global financial crises both impact the likelihood of observing herding behavior it is quite likely that more refined explanations will require a more complex model. Moreover, institutional developments, such as the rise of ETFs and the growing importance of institutional investors over time are other factors that will require further study to assess their impact on herding behavior. From the methodological point of view one way to proceed is to allow for gradual regime switches depending on, for example, the stock market volatility rather than high-low volatility regime switches.

References

Avery, C./*Zemsky,* P. (1998): Multidimensional Uncertainty and Herd Behavior in Financial Markets, American Economic Review, Vol. 88(4), pp. 724–48.

Azzalani, A./*Capitanio*, A. (2014): The Skew Normal and Related Families (Cambridge: Cambridge University Press).

Baker, M./*Stein*, J. (2004): Market Liquidity as a Sentiment Indicator, Journal of Financial Markets, Vol. 7(3), pp. 271–99.

Baker, M./*Wurgler,* J. (2006): Investor Sentiment and the Cross-section of Stock Returns, The Journal of Finance, Vol. 61(4), pp. 1645–80.

Banerjee, A. (1992): A Simple Model of Herd Behavior, The Quarterly Journal of Economics, Vol. 107(3), pp. 797–817.

Bekaert, G./*Horeova*, M. (2014): The VIX, the Variance Premium, and Stock Market Volatility, Journal of Econometrics, Vol. 183(2), pp. 181–192.

Bekaert, G./*Horeova*.M./*Lo Luca*, M. (2010): Risk, Uncertainty, and Monetary Policy, paper presented at the 11th Jacques Polak Annual Research Conference, November 4–5, International Monetary Fund.

Ben-David, I./*Franzoni*. F. A./*Moussawi*, R. (2014): Do ETF's Increase Volatility?, NBER working paper 20071.

Bernhardt, D./*Campello*. M./*Kutsoati*, E. (2006): Who Herds?, Journal of Financial Economics, Vol. 80, pp. 657–675.

Bikhchandani, S./*Hirshleifer*. D./*Welch*, I. (1992): A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades, Journal of Political Economy, Vol. 100(51), pp. 992–1026.

Bikhchandani, S./*Sharma*, S. (2001): Herd Behavior in Financial Markets, IMF Staff Papers, Vol. 47(3), pp. 279–310.

Bohl, M. T./*Gebka*. B./*Goodfellow*, C. (2009): Together We Invest? Individual and Institutional Investors' Trading Behaviour in Poland, International Review of Financial Analysis, Vol. 18(4), pp. 212–21.

Cagilnap, G./*DeSantis*. M./*Syrak*, A. (2014): The Non-Linear Price Dynamics of U.S. Equity ETFs, working paper.

Campbell, J./*Hilscher*, J./*Szilagyi*, J. (2008): In Search of Distress Risk, Journal of Finance, Vol. 58(6), pp. 2899–2939.

Chang, E. C./*Cheng*, J. W./*Khorana*, A. (2000): An Examination of Herd Behavior in Equity Markets: An International Perspective, Journal of Banking & Finance, Vol. 24(10), pp. 1651–1679.

Chiang, T./*Zheng*, D. (2010): An Empirical Analysis of Herd Behavior in Global Stock Markets, Journal of Banking & Finance, Vol. 34(8), pp. 1911–1921.

Christie, W. G./*Huang*, R. D. (1995): Following the Pied Piper: Do Individual Returns Herd around the Market?, Financial Analysts Journal, Vol.51(4), pp. 31–37.

Clifford, C. P./*Fulkerson*, J. A./*Jordan*, B. D. (2014): What Drives ETF Flows?, Financial Review, Vol. 49(3), pp. 611–642.

De Long, B./*Shleifer*. A./*Summers*. L./*Waldmann*, R. (1990): Positive Feedback Investment Strategies and Destabilizing Rational Speculation, The Journal of Finance, Vol. 45(2), pp. 379–395.

– The Survival of Noise Traders in Financial Markets, The Journal of Business, Vol. 64(1), pp. 1–19.

Dempster, A. P./*Laird*, N. M./*Rubin*, D. B. (1977): Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society, Series B (Methodological), Vol. 39(1), pp. 1–38.

Devenow, A./*Welch*, I. (1996): Rational Herding in Financial Economics, European Economic Review, Vol. 40, pp. 603–615.

Diebold, F. X./*Lee*, J. H./*Weinbach*, G. C. (1994): Regime Switching with Time-varying Transition Probabilities, In: *Nonstationary Time Series Analysis and Cointegration*, edited by C. P. Hargreaves, Oxford: Oxford University Press, pp. 283–302.

Estrella, A./*Hardouvelis*, G. A. (1991): The Term Structure as a Predictor of Real Economic Activity, The Journal of Finance, Vol. 46(2), pp. 555–576.

Estrella, A./*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, Vol. 41(7), pp. 1375–1401.

– (1998): Predicting U.S. Recessions: Financial Variables as Leading Indicators, The Review of Economics and Statistics, Vol. 80(1), pp. 45–61.

Fernandez, C./*Steel*, M. F. J. (1998): On Bayesian Modeling of Fat Tails and Skewness, Journal of the American Statistical Association, Vol. 93(441), pp. 359–371.

Froot, K. A./*Scharfstein*, D. S./*Stein*, J. C. (1992): Herd on the Street: Informational Inefficiencies in a Market with Short-term Speculation, The Journal of Finance, Vol. 47(4), pp. 1461–1484.

Gebka, B./*Wohar*, M. E. (2013): International Herding: Does it Differ across Sectors?, Journal of International Markets, Institutions and Money, Vol. 23, pp. 55–84.

Gleason, K. C./*Mathur*, I./*Peterson*, M. A. (2004): Analysis of Intraday Herding Behavior among the Sector ETFs, Journal of Empirical Finance, Vol. 11 (5), pp. 681–694.

Graham, J. R. (1999): Herding among Investment Newsletters: Theory and Evidence, The Journal of Finance, Vol. 54(1): 237–268.

Gray, S. F. (1996a): An Analysis of Conditional Regime-switching Models, Working paper.

– (1996b): Modeling the Conditional Distribution of Interest Rates as a Regime-switching Process, Journal of Financial Economics, Vol. 42, pp. 27–62.

Grinblatt, M./*Titman*. S./*Wermers*, R. (1995): Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior, The American Economic Review, Vol. 85 (5), pp. 1088–1105.

Gürkaynak, R./*Wright*, J. (2015): Macroeconomics and the Term Structure, Journal of Economic Literature (forthcoming).

Hall, S. G./*Psadarakis*. Z./*Sola*, M. (1999): Detecting Periodically Collapsing Bubbles: A Markov-Switching Unit Root Test, Journal of Applied Econometrics, Vol. 14(2), pp. 143–154.

Hall, S. G./*Sola*, M. (1994): Rational Bubbles during Poland's Hyperinflation: Implications and Empirical Evidence, European Economic Review, Vol. 38(6), pp. 1257–1276.

Hamilton, J. D. (1990): Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics*, Vol. 45(1–2), pp. 39–70.

Hansen, B. E. (1994): Autoregressive Conditional Density Estimation, International Economic Review, Vol. 35 (3), pp. 705–730.

Hwang, S./*Salmon*, M. (2004): Market Stress and Herding, Journal of Empirical Finance, Vol. 11(4), pp. 585–616.

– (2009): Sentiment and Beta Herding, Working paper.

Jeanne, O./*Rose*, A. K. (2002): Noise Trading and Exchange Rate Regimes, Quarterly Journal of Economics, Vol. 117(2), pp. 537–569.

Jondeau, E./*Rockinger*, M. (2003): Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements, Journal of Economic Dynamics & Control, Vol. 27(10), pp. 1699–1737.

Jones, C. (2002): A Century of Stock Market Liquidity and Trading Costs, Working paper.

Keim, D. B./*Madhavan*, A. (1995): Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Traders, Journal of Financial Economics, Vol. 37 (3), pp. 371–398.

Kim, C. J. (1994): Dynamic Linear Models with Markov-Switching, Journal of Econometrics, Vol. 60(1–2), pp. 1–22.

Knez, P. J./*Litterman,* R./*Scheinkman*, J. (1994): Explorations Into Factors Explaining Money Market Returns, The Journal of Finance, Vol. 49(5), pp. 1861–1882.

Lakonishok, J./*Shleifer*, A./*Vishny*, R. W. (1992): The Impact of Institutional Trading on Stock Prices, Journal of Financial Economics, Vol. 32(1), pp. 23–43.

Lee, C. M./*Shleifer*, A./*Thaler*, R. H. (1991): Investor Sentiment and the Closed-end Fund Puzzle, The Journal of Finance, Vol. 46(1), pp. 75–109.

Litterman, R./*Scheinkman*, J. (1991): Common Factors Affecting Bond Returns, Journal of Fixed Income, Vol. 1, pp. 54–61.

Naes, R./*Skjeltorp*, J./*Oodergarrd*, B. A. (2011): Stock Market Liquidity and the Business Cycle, Journal of Finance, Vol. 66, pp. 139–176.

Naujoks, M./*Aretz*, K./*Kerl*, A./*Walter*, A. (2009): Do German Security Analysts Herd?, Financial Markets and Portfolio Management, Vol. 23, pp. 3–29.

Newey, W. K./*West*, K. D. (1987): A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, Vol. 55(3), pp. 703–708.

– (1994): Automatic Lag Selection in Covariance Matrix Estimation, The Review of Economic Studies, Vol. 61(4), pp. 631–653.

Scharfstein, D. S./*Stein*, J. C. (1990): Herd Behavior and Investment, The American Economic Review, Vol. 80(3), pp. 465–479.

Shiller, R. J./*Fisher*, S./*Friedman*, B. M. (1984): Stock Prices and Social Dynamics, Brookings Papers on Economic Activity, Vol. 2, pp. 457–510.

Stavroyiannis, S./*Babalos,* V./*Zarangas*, L. (2013): Time Varying Herding and Anti-Herding Behavior and DCC in European Market Indices, Working Paper.

Stavroyiannis, S./*Babalos*, V. (2013): On the Time Varying Nature of Herding Behavior: Evidence from Mayor European Indices, Working Paper.

Tan, L./*Chiang*, T./*Mason*, J. R./*Nelling*, E. (2008): Herding Behavior in Chinese Stock Markets: An Examination of A and B Shares, *Pacific-Basin Finance Journal, Vol.* 16(1–2), pp. 61–77.

Tang, X. (2013): Do Professional Forecasters Herd?, Working Paper.

Watanabe, M./*Yamaguchi*, K. (Ed.) (2004): The EM Algorithm and Related Statistical Models (New York: Marcel Dekker).

Welch, I. (1992): Sequential Sales, Learning, and Cascades, The Journal of Finance, Vol. 47(2), pp. 695–732.