Aggregate US Merger levels: An explained markov switching analysis

September 2008
Do not quote without permission

Abstract

This paper adds to the current literature on the wave nature of the Mergers and Acquisitions time series in the United States. We aim to bridge the link between the macroeconomic theory of “merger waves” and the econometric literature on modeling levels of merger activity. From economic theory we compile a list of macroeconomic variables widely believed to influence levels of merger activity. Building on previous econometric work we model the time series as a first order autoregressive two regime Markov switching process which incorporates a set of macroeconomic variables. Incorporating these variables causes a time variance in the regime switching probabilities, governed by the variables’ behaviour. We find support for a Markov Switching (MS) model with regime dependent means and variances. Our model supports the idea that improving economic conditions facilitate merger waves. We also find evidence that liquidity plays a role in merger activity. Finally, we find that equity price volatility and valuation discrepancies lead to persistence in merger waves.
1. Introduction

Throughout the last century the time series of M&A activity has been marked by several large bursts of activity interrupting otherwise steady levels. These have become known as merger waves, and are now a widely researched and documented phenomenon in financial markets. Five such waves are commonly accepted to have occurred over the last century within the US market. Despite the large literature on M&A behavior, there exists as yet no universally accepted theory of the driving forces behind these waves.

The literature to date can be loosely classified into two categories: measurement and explanations. One set examines the nature of the M&A time series from an econometric or statistical perspective while the other set develops theories to explain why these observed large surges in activity may occur. This second body of literature highlights a large number of macro-level economic variables, the behaviour of which may facilitate surges in M&A activity. From this literature we have compiled a large set of macro economic variables believed influence aggregate M&A behavior.

Our aims in this paper are to further the first branch of econometric literature by examining the application of Markov regime switching models to M&A activity while simultaneously bridging the link between these two fields of research regarding merger activity. We build on Town’s (1992) two regime switching model by incorporating a vector of conditioning variables which can directly influence the probability of a regime switch.

The outline of this paper is as follows; section 2 reviews the literature on merger waves, section 3 discusses our chosen set of conditioning variables, section 4 discusses
regime switching models, section 5 gives a description of our data sets, section 6 further outlines the methodology, section 7 gives out estimation results and some discussion, section 8 shows results from forecasts (to be added) and section 9 concludes.

2. M&A Literature to Date

Merger waves have been extensively researched yet there are still no definite conclusions as to why these bursts of activity occur when they do. The literature identifies five major waves in the United States since 1895 and four major merger waves in the UK since the 1920’s, though there is still some dispute over the exact timing of some of these waves commencement and conclusion. Typically these waves have been judged solely on an ad-hoc and usually post-facto basis; there is thus an inherent considerable systematic error in the accuracy of identifying and dating these waves.

The existing literature on merger waves can be broadly separated into two categories. The first category consists of work that attempts to find a link between acquisition activity and certain economic or behavioural variables at the macroeconomic, industrial and firm level while the second category consists of work that explores the nature of the merger activity time series itself.

Within the first category there are several existing theories on what drives surges in the M&A market. One of the earliest is the idea that merger waves occur in response to certain types of economic disturbances (Gort, 1969). Changes in external conditions result in differences of opinion concerning the value of firms and this, in turn, leads to increased levels of M&A activity. This shock hypothesis has gained renewed acceptance in recent
years (Mitchell and Mulherin 1996, Harford 2005) with evidence that shocks cluster with merger waves in an industry. The belief is that the collective reaction of firms to these shocks, both inside and outside the industry, is to reallocate assets through M&A. The activity clusters in time as managers compete for assets. There is also evidence that capital liquidity plays an important role in this process (Harford 1999).

Nelson (1959) first examined merger activity at the turn of the 20th century and found positive correlations between industrial activity, stock prices and levels of M&A activity. In the following years more papers started to appear linking macro-economic variables such as GNP, stock prices and the business cycles (see as examples Steiner 1975, Globe and White 1988, Beckatti, 1986) to M&A activity. This has lead to the widely supported belief that improving economic conditions facilitate merger waves (Melicher, Ledolter and D’Antonio 1983, Gort 1969). The observed positive correlation between stock valuations and merger activity bred the ‘misvaluation hypotheses’ whereby, in bull markets, companies with overvalued stock will use it to acquire the real assets of target companies (Shleifer and Vishny 2003). Several other papers using firm level data have suggested behavioural hypotheses such as “empire building” managers, the fear of competition and herding behaviour (Gorton, Kaul and Rosen 2005, Harford 2003, Toxvaerd, 2003).

The second category of research we explore concerns the nature of the time series itself. Nelson (1959) noted that M&A activity in the US was characterised by “large bursts of activity separated by lengthy intervals of low activity”. The view that mergers come in waves has been the subject of considerable interest since Nelson's work first appeared, yet views on how to best model this process are varied. Shughart and Tollison
(1984) were unable to reject the hypothesis that merger activity is the result of a random walk or AR(1) process. Globe and White (1993) tested for merger waves in US history by fitting sine curves to the data. They found that the parameters characterising the curves were statistically significant and supported their hypothesis that mergers do have a wave like pattern. Barkoulas, Baum and Chakraborty (2001) proposed an autoregressive fractionally integrated moving average process. They concluded that M&A activity is strongly auto correlated and they attribute the observed lack of periodicity to the presence of long-memory dynamics.

Town (1992) successfully applied a two regime Markov switching model to M&A activity in both UK and US time series data. In this framework a merger wave is defined as a large discrete increase in the mean of the series. In his paper he compared a linear ARIMA processes and an AR(2) process incorporating Markov regime switching. From statistical tests he concluded that the regime switching model was superior. His work identified the existence of two distinct regimes in M&A activity with the mean and variance of the model taking significantly different values in each regime. This specific Markov switching model applied has gained support in three more recent works. Linn and Zhu (1997) reject the hypothesis of a random walk in favour of a model almost identical to Town’s. Using this model Resende (1999) finds evidence of merger waves in sectoral data in the UK. Gartner and Halbheer (2007) successfully apply a slightly modified version of the model used by Town and Linn and Zhu to updated and more consistent data in the UK and US. All four papers reject the hypothesis of a random walk in favor of a two regime Markov switching model.
The papers mentioned above have all focused on investigating the nature of the time series data with little regard to the factors determining the likelihood of a regime switch. This meant that the probability matrix governing the regime switch has hitherto been time independent. Recently, the more flexible regime switching model with time dependent transition probabilities (TVTP) has been introduced. The methodological details of this process are outlined by Diebold, Lee, and Weinbach (1994). This TVTP model has been applied to different economic phenomena such as exchange rates (Engel and Hakkio, 1996), stock returns (Schaller, Huntley & van Norden, 1997) and the business cycle (Filardo, 1994). Building on the previous support of regime switching in modelling the business cycle, Filardo (1994), introduced a model with time varying transition probabilities and found conclusive evidence that the CLI (Composite Index of Eleven Leading Indicators) performed well as an explanatory variable in predicting business cycles. In his 1999 paper Resende notes the limitations of having time independent regime switching probabilities and suggests extending the model to the time varying case may be more apt. His 2005 paper makes an attempt at this by incorporating variables referring to real output growth, real growth in money supply and real stock market returns into the probability functions.

3. Conditioning Variables

After reviewing the literature on the links between economic conditions and the level of M&A activity, we identify below a list of macroeconomic variables which we believe from this research to be leading indicators in our regime switching model. These variables fall into four major theoretical categories.
1. Economic Conditions: A good economic climate is broadly believed to aid the market for corporate control. More specifically the literature points to several proxies for economic conditions believed to be linked to M&A levels. Early work in this field speculated positive links between rising stock prices and industrial production (Weston 1953, Gort 1969). In the early 1980’s Melicher, Ledolter, & D'Antonio, (1983) introduced their Economic Prosperity theory which states that activity leading to mergers will increase with expectations of economic growth and capital market conditions that are conductive to financing mergers. Their paper finds levels of M&A to be positively correlated with stock price movements and negatively correlated with the level of business failures. There is a long standing low positive correlation between movements in GNP and the level of M&A activity (Steiner 1975, Chung and Weston 1982, Beckatti 1985). The reason for this is believed to be rooted in the ties between GNP, the business cycle and the overall size of the economy, the argument for the latter bring that the larger the economy the greater the number of potential mergers. We consider the lagged level of M&A since the merger time series is believed to be strongly auto correlated (Barkoulas, Baum and Chakraborty, 2001). We include corporate profit levels since it is believed that cash-richness is a strong predictor of acquisitions (Harford, 1999). Finally we include levels of unemployment (Bjorklund, Gray and Boyle, 2005) in our analysis to give the most complete description of the economic climate regarding the market for corporate control.
1. **Valuation Discrepancy**: As noted above, securities prices have a history of being linked to merger levels. However, several competing theories exist regarding the effects of changes in securities prices on the M&A market. Weston (1953) suggests that gains from high securities prices promote merger activity while Gort (1969) suggests changes in securities prices are a form of economic shock to the market and these shocks in turn create valuation discrepancies which change the profitability of mergers. This theory suggests that either an acquirer will bid on a target which it views as undervalued (Beckatti, 1986) or else an overvalued acquirer will bid on a target using its overvalued shares as payment (Shleifer and Vishny 2003, Rhodes-Kropf and Viswanathan 2004). Within this paper we use three separate measures to attempt to capture this behavior; securities prices (already included in the list of potential explanatory variables for their use as a proxy of economic conditions), the corresponding price-earnings ratios and price volatility. The price volatility is used as an average measure of the dispersion and uncertainty in stock prices. It serves as a measure of the magnitude of possible errors in managers valuations of their own firms stock or the stock of a target firm.

3. **Financing Mergers**: Since a significant number of M&A deals are financed either partially or fully by debt\(^1\), it is reasonable to infer that changes in interest rates directly affect the financing and hence the numbers of mergers (Bjorklund, Gray and Boyle 2005, Melicher, Ledolter, & D'Antonio 1983, Globe and White 1988). Furthermore Harford (1999) find evidence that a capital liquidity component must be included when modeling mergers to reflect the ease of financing. The literature predicts an inverse relationship between interest rates and the level of merger activity. Here we proxy interest

---

\(^1\) Estimates of the levels of debt financing in our dataset to be added.
rates two ways; via the spread between corporate bond rate and the Treasury bond rate and also via the level of corporate bank loan interest rates. This reflects the cost of money. The interest rate spread is also a well acknowledged leading indicator of economic conditions, bankruptcy and the business cycle (Friedman and Kuttner 1992, Kwark 2002). Explanations for this leading behavior are based on investors' perceptions of the future economy (Friedman and Kuttner, 1994). If investors expect future economic conditions to be favorable then the bankruptcy risk on risky investments should accordingly, this then means that investors require a relatively small risk premium over the risk-free interest rate, which then reduces the interest rate spread. We also include the M2 measure of Money Supply in the US as a proxy for liquidity to reflect the ease of financing.

4. Sentiment: Finally our model incorporates several variables to give an idea of market sentiment. To date literature focusing on the macroeconomic levels of mergers fails to include variables of this nature. However we feel that the inclusion of these variables will add to the explanatory power of our model by reflecting some of the motivations for merging at the firm level. We include investor and consumer sentiment to reflect the effect that shareholders can have on decisions to merger. Purchasing managers’ sentiment reflects the production, employment and sales levels of firms and hence can be used to proxy the growth potential of firms since growth and synergy are cited in the firm level literature as one of the main motivations to merge (Mukherjee, Kiymaz, and Baker 2004). The University Of Michigan Index Of Consumer Sentiment we use has also been established as a leading predictor of GDP growth and in turn the business cycle in the US.
The relationship between consumer sentiment and stock prices also shows positive correlation, with stock prices as leading indicators for consumer sentiment (Jansen and Nahuis 2003, Otoo, 1999).

4. Regime Switching Models

Modeling the dynamics of any long term time series of macroeconomic quantities faces the challenge that series quite often undergo “structural breaks” marked by distinct shifts in behavior. These breaks can occur in response to episodes such as wars, depressions, hyperinflations or other financial shocks. Changes in behavior may be temporary and/or recurrent. These changes indicate that constant parameter models may not be optimal in modeling long-term time series.

Recently regime switching models have become a popular framework for capturing the non-linear behavior seen in these time series. These models are based on the idea that the parameters of the time series, such as the mean and variance, assume different values within different time periods or “regimes”. The time series switches between these different regimes in accordance with a probability law. First introduced by Hamilton (1989) to explain business cycles, regime switching models have since been applied to a great number of phenomena including interest rates (Gray, 1996), exchange rates (Engle, 1994), inflation (Simon, 1996), the volatility of equity returns (Dueker, 1997) and more recently merger and acquisitions activity (Town, 1992).

---

2 In this paper we refer several times to structural breaks. In this context we define structural breaks as changes in the data generating process of our dependant variable as observed by inferred regime switches. Results of more definitive tests for structural breaks in the time series to be added.
As discussed, this paper aims to build on Town’s (1992) work by extending his original two regime model to consider the case where the probability law governing the regime switch is itself a function of a set of macro economic variables. There are several reasons why we speculate that this kind of model is suitable for modeling aggregate M&A activity. Firstly, the success of previous studies strongly suggests that regime switches are characteristic of the M&A time series. Secondly, the extensive literature on the determinants of merger waves supports the idea that waves are driven by exogenous economic variables, shocks as it were. The TVTP model can merge these two areas of research to provide further and more definite insight into the causes of merger waves. Thirdly, the TVTP has increased flexibility over its fixed counterpart. As noted by Filardo (1994) and Gordon (1993) with the FTP model the expected durations of regimes are fixed, while the TVTP model allows for the duration of regimes to vary with time and changes in the conditioning variables.

**The TVTP Model**

Many economic time series exhibit structural breaks marked by abrupt changes in the behavior of the series. Quite often, certain variables can behave quite differently during events such as market bubbles, economic downturns and so on and regime switching models provide an intuitive way to capture this non linear behavior. Here we model the number of announced mergers per month in the United States, \( y_t \), as a two regime auto regressive process. In general the regime in operation at a point in time is an unobserved variable. The process which determines the probability of a regime switch is
incorporated into the stochastic structure of the model, hence it is possible to draw inference about which regime is in operation. Markov models assume that regime shifts evolve according to a Markov chain.

One of the most widely used Markov models is a variation of the autoregressive (AR) model of order k.

\begin{equation} y_t = \alpha + \sum_{j=1}^{k} \phi_j y_{t-j} + \varepsilon_t \tag{1} \end{equation}

\begin{equation} y_t = \alpha_{s_t} + \sum_{j=1}^{k} \phi_{j s_t} y_{t-j} + \varepsilon_t \tag{2} \end{equation}

(1) Shows the AR(k) models without regime switching and (2) shows the same model incorporating regime switching. The disturbance term is assumed to be normally distributed. \(^3\)

\[ \varepsilon_t \sim (N, \sigma^2) \]

The difference in (2) is that the parameters of the model are dependant on a state variable \( S_t \) which takes a value \( S_t = i, \ i=0,1, \ldots, N-1; \) and denotes which of the N regimes are in operation at time \( t \). The model is simply a constant parameter linear regression conditional on the regime in operation at the time. These Markov models assume that the

\[^3\] This assumption of a normal distribution may not be true. As seen in other financial time series from work on ARCH and GARCH models variance distributions are not always normal. Nonetheless, it is common, and computationally much simpler, to assume a normal distribution, which we do here.
state variable $S_t$ follows an N-state Markov chain. The evolution of the transition probabilities of switching from one regime to another is described in (3).

\[
P(S_t = i | S_{t-1} = j, S_{t-2} = k, \ldots) = P(S_t = i | S_{t-1} = j) = p_{ij}
\]

Supposing $y_t$ can be observed directly but the value of $S_t$ can only be inferred based on the behavior of $y_t$. For a two regime model ($S_t = 0, 1$) this inference takes the form of two probabilities which sum to unity.

\[
\xi_{jt} = Pr(s_t = j | \Omega_t; \theta)
\]

(4)

\[
\Omega_t = \{y_t, y_{t-1}, \ldots, y_1, y_0\}
\]

\[
\theta = (\sigma, \phi_1, \phi_2, \alpha_1, \alpha_2, p_{11}, p_{22})'
\]

Here $\Omega_t$ is the set of observations as of time t and $\theta$ is the vector of population parameters. The inference is performed iteratively for each $t = 1, 2, \ldots, T$ with step $t$ having equation (5) as input for each $i = 0, 1$. 

\[
\xi_{jt-1} = Pr(s_{t-1} = i | \Omega_{t-1}; \theta)
\]

(5)

To perform this estimation the densities under the two regimes are needed (6).

\[
\eta_{jt} = f(y_t | S_t = s_t, \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}c} exp \left[ -\frac{(y_t - \phi_j - \phi_{y-1})^2}{2\sigma^2} \right]
\]

(6)
From this the conditional density takes the form of (7) and the desired output is in (8).

\begin{equation}
 f(y_t|\Omega_{t-1}; \theta) = \sum_{i=0}^{1} \sum_{j=0}^{1} p_{ij} \xi_i, t-1 \eta_{jt}
\end{equation}

\begin{equation}
 \xi_{i,t} = \frac{\sum_{j=0}^{1} p_{ij} \xi_i, t-1 \eta_{jt}}{f(y_t|\Omega_{t-1}; \theta)}
\end{equation}

For a specified value of the conditional log likelihood is (9) and an estimate of \( \theta \) is obtained by maximization of this numerically.

\begin{equation}
 \log f(y_1, y_2, ..., y_T|y_o; \theta) = \sum_{t=1}^{T} \log f(y_t|\Omega_{t-1}; \theta)
\end{equation}

A common restriction on these models is that transition probabilities evolve independently of the lagged values of the series itself. Regimes are said to evolve “exogenously” of the series. Hence these models are the best choice when the aim is to investigate regime shifts in the data without tying shifts to any observable variable. A full discussion on the estimation of the parameters of Markov switching models can be found in Hamilton (1994). Based on our data the model we choose to use is a two regime first order autoregressive AR(1) model with state dependant means and variances. (This is explained in more detail in section 6). We aim to capture two distinct regimes which we believe to be present; regime one, a high mean/variance state and regime zero, a low mean/variance state.
Our model builds on the simple two regime switching process by allowing the transition probabilities to be dependant on our set of economic variables $z_t$.

\begin{equation}
P(S_t = s_t | S_{t-1} = s_{t-1}, z_t) = \begin{bmatrix}
    q(z_t) & 1 - p(z_t) \\
    1 - q(z_t) & q(z_t)
\end{bmatrix}
\end{equation}

The parameters of the model and the transition probability parameters are jointly estimated via the conditional joint density distribution, $f$. The conditional density for a first order process, $f^*$, is

\begin{equation}
f^*(y_t | y_{t-1}, z_t)
\end{equation}

\begin{equation}
= \sum_{s_t=0}^{1} \cdots \sum_{s_{t-1}=0}^{1} f(y_t, S_t = s_t, S_{t-1} = s_{t-1} | y_{t-1}, z_t)
\end{equation}

\begin{equation}
= \sum_{s_t=0}^{1} \cdots \sum_{s_{t-1}=0}^{1} \hat{f}(y_t, S_t = s_t, S_{t-1} = s_{t-1} | y_{t-1})
\end{equation}

\begin{equation}
\times P(S_t = s_t | S_{t-1} = s_{t-1}, z_t)
\end{equation}

\begin{equation}
\times P(S_{t-1} = s_{t-1} | y_{t-1}, z_{t-1})
\end{equation}

And the log likelihood function is

\begin{equation}
L(\theta) = \sum_{t=1}^{T} ln[f^*(y_t | y_{t-1}, z_t; \theta)]
\end{equation}

Equations (11) and (12) show how the level of mergers and the conditioning variables affect the estimation of the model; both sources enter the model in two ways. The level of mergers directly influences the likelihood through the normal density and
indirectly through the information they provide about past states. The conditioning variables affect the transition probability matrix directly and the distribution of states indirectly through the lagged values.

The Markov switching model allows us some inference about the unobserved state of the model. In the TVTP case past and present information on the level of mergers as well as the information contained in the set of conditioning variables help identify which regime is dominant at any given time. To assess the effect of the conditioning variables on the inference about the state of the time series we must make the link between the transition probabilities $P(S_t = s_t | S_{t-1} = s_{t-1}, z_t)$ and the inferred probabilities $P(S_t = s_t | y_t, y_{t-1}, z_t)$. This is calculated by integrating out the effects of the past states in the joint density distribution,

$$P(S_t = s_t | y_t, y_{t-1}, z_t) =$$

$$= \sum_{s_{t-1}=0}^{1} \cdots \sum_{s_{t-1}=0}^{1} P(S_t = s_t, S_{t-1} = s_{t-1} | y_t, y_{t-1}, z_t)$$

$$= \sum_{s_{t-1}=0}^{1} \cdots \sum_{s_{t-1}=0}^{1}$$

$$\times \frac{f(y_t, S_t = s_t, S_{t-1} = s_{t-1} | y_{t-1}, z_t)}{f^*(y_t | y_{t-1}, z_t)}$$

The transition probabilities influence the density-distribution $f$ and thus directly affect the inferred probabilities as seen in equation (13). Clearly the FTP model is a simple case of this whereby the conditioning variables are non-informative about the evolution of the regimes.
The switching model we assume in (2) characterizes two distinct regimes. These should be characterized by marked differences greater than zero in the means \( \mu_0 \) and \( \mu_1 \) as well as the variances \( \sigma_0 \) and \( \sigma_1 \).

To test for time variation, we must test jointly for the appropriateness of the functional form of the TVTP model as well as statistical significance of the coefficients on the conditioning variables. We can reject the null hypothesis of no time variation in the transition probabilities if \( \delta = 2(L(\theta)_{FTP} - L(\theta)_{TVTP}) \) exceeds \( \chi^2_{J_1+J_2, \alpha} \) where \( J_1 + J_2 \) is the number of restrictions. The functional form for the transition probabilities maps the conditioning variables into the open interval \((0,1)\). The parameterisation for the test is

\[
\begin{align*}
    p(z_t) &= \frac{\exp(\psi_{p0} + \sum_{j=1}^{J_1} \psi_{pj} z_{t-j})}{1 + \exp(\psi_{p0} + \sum_{j=1}^{J_1} \psi_{pj} z_{t-j})} \\
    q(z_t) &= \frac{\exp(\psi_{q0} + \sum_{j=1}^{J_1} \psi_{qj} z_{t-j})}{1 + \exp(\psi_{q0} + \sum_{j=1}^{J_1} \psi_{qj} z_{t-j})}
\end{align*}
\]

Here, \( p(z_t) \) is the probability of going from regime one to regime one in the next time step and \( q(z_t) \) is the probability of going from regime zero to regime zero. The special case of the FTP model corresponds to \( \varphi_{pj} = \varphi_{pq} = 0 \) for all \( j \neq 0 \). The effect of the conditioning variables is inferred from their effects on the movements of \( p(z_t) \) and \( q(z_t) \). For example, if \( p_t \) (the probability of going from regime 1 to regime 1) increases and \( q_t \) (the probability of going from regime 0 to regime 0) decreases as a result of \( z_t \) increasing, both the probability of going from 1 to 1 and 0 to 1 increase (i.e. 1- \( q_t \) increases), hence the probability of being in regime 1 and time \( t+1 \) increases overall.
The parameters of interest are estimated with maximum likelihood (ML) methods for mixtures of normals. All estimations were done using RATS 7.0 software and codes are available from the authors on request.

5. Data Description

Our M&A dataset for the United States consists of the number of announced M&A deals between public companies each month and spans from January 1985 to December 2007. We took all deals where the target company was a US domiciled entity. This resulted in 31,955 separate deals over the twenty-two year period and accounts for just over 47% of the world market during this time frame. This is shown in figure 1.

In total we have 16 conditioning variables listed in table 1. All data was obtained from Datastream except the proxy for investor sentiment which was taken from Yale Universities stock market confidence indices compiled under the direction of Dr. Robert Schiller (http://icf.som.yale.edu/Confidence.Index/).
**Table 1. List of 16 conditioning macro economic variables**

<table>
<thead>
<tr>
<th>Measuring</th>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securities Prices</td>
<td>PI</td>
<td>S&amp;P 500 composite price index</td>
</tr>
<tr>
<td>GNP</td>
<td>GNP</td>
<td>US Gross National Product</td>
</tr>
<tr>
<td>Industrial Productivity</td>
<td>IPROD</td>
<td>US Industrial Production Index</td>
</tr>
<tr>
<td>Unemployment</td>
<td>UNPLY</td>
<td>US Unemployment (initial claims)</td>
</tr>
<tr>
<td>Corporate Profits</td>
<td>CPR</td>
<td>US Corporate Profits</td>
</tr>
<tr>
<td>Bankruptcy</td>
<td>BKR</td>
<td>Bankruptcy Filings in the US</td>
</tr>
<tr>
<td>Economic Indicators</td>
<td>EI</td>
<td>The Conference Board Composite Index of Leading Indicators</td>
</tr>
<tr>
<td>Price Earnings Ratio</td>
<td>PE</td>
<td>S&amp;P 500 composite price earnings ratio</td>
</tr>
<tr>
<td>Price Volatility</td>
<td>PVOL</td>
<td>CBOE Volatility Index VXO</td>
</tr>
<tr>
<td>Liquidity</td>
<td>LIQ</td>
<td>Money Supply in the US (M2 measure)</td>
</tr>
<tr>
<td>Bank Loan Interest Rate</td>
<td>BLIR</td>
<td>Bank Prime Loan Rate (middle rate) in the US</td>
</tr>
<tr>
<td>Interest Rate Spread</td>
<td>SPIR</td>
<td>Spread between Moody’s middle AAA rate and US Treasury Bill (3 month)</td>
</tr>
<tr>
<td>Consumer Sentiment</td>
<td>CSEN</td>
<td>University of Michigan Consumer Sentiment Index</td>
</tr>
<tr>
<td>Purchasing Manager sentiment</td>
<td>PMSEN</td>
<td>ISM Purchasing Managers Index</td>
</tr>
<tr>
<td>Investor Sentiment</td>
<td>ISEN</td>
<td>Robert Schiller One-Year Index - Institutional sentiment index</td>
</tr>
</tbody>
</table>

**Figure 1. Shows a graph over our sample period of the monthly levels of M&A activity in the United States**

![Number of Monthly M&A Deals in the US](image.png)
6. Methodology

6.1 Testing for Regimes

Some of the biggest challenges regarding the use of regime switching models involve firstly, hypothesizing the number of regimes present in the data and secondly, determining which elements of the model switch over time (mean, variance, autoregressive parameters etc.). We begin our analysis with the assumption that the data is best modeled by a two regime process with regimes characterized by different means and variances. This is in line with the well discussed phenomena of merger waves and also is supported by the evidence of a two regime process in M&A levels found in Town’s 1992 paper.

Regarding testing the non-linear behavior seen in MS models simple likelihood ratio tests fail since the nuisance (switching) parameters are not identified under the null hypothesis of no regime switch (Hansen, 1992). However, Garcia (1992) derives analytically the asymptotic null distribution of the likelihood ratio test for various two regime MS models; we use the critical values derived in this paper to support our assumption of different regimes. Table 2 shows the results of the likelihood ratio test for the null hypothesis of no switching in the M&A time series against three alternative speculations which involve switching between two regimes. The likelihoods are derived using the simplest model with no autoregression so values are drawn from one or more Gaussian distributions with differences in means and/or variances. The 5% and 1% critical values using Garcia’s non-standard distribution are 10.34 and 13.81 for switching
in the means and 13.52 and 17.67 for switching in the mean and variances (Schaller, Huntley and van Norden, 1997).

Table 2. Tests for regime switching in M&A behavior

<table>
<thead>
<tr>
<th></th>
<th>no switching</th>
<th>switching in means</th>
<th>switching in variances</th>
<th>in switching in means and variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log - Likelihood</td>
<td>-1541.335</td>
<td>-1489.214</td>
<td>-1434.96</td>
<td>-1379.878361</td>
</tr>
<tr>
<td>Likelihood Ratio Test</td>
<td>104.242</td>
<td>212.75</td>
<td>322.9132775</td>
<td></td>
</tr>
</tbody>
</table>

Using Garcia’s criteria the conclusion is clear. In all three cases the likelihood ratio tests rejects the hypothesis of no switching in favor of a two regime model. As we move from the first column to the end it is clear that the rejection of the null hypothesis becomes stronger, indicating that the appropriate formulation is mean and variance switching.

Unfortunately Garcia does not provide critical values for tests of models with more then two regimes. We solve this problem by employing some further tests to help determine the number of regimes present in the data. We calculate Smith, Naik and Tsai’s (2005) Markov Switching Criteria (MSC). This test is based on minimization of the Kullback-Leibler (KL) divergence. It is similar to Akaike’s information criteria (AIC) but is specifically adapted for use with regime switching models. In table 3 we estimate the MSC for twelve cases; four different orders of autoregression each for the one, two and three regime cases. Based on our previous test, the two and three regime models have switching in both the means and variances between regimes.
Table 3. MSC criteria.

Using notation and recommendations from Smith, Naik and Tsia, 2005 we set $\lambda = N$ and $\delta = 1$

<table>
<thead>
<tr>
<th></th>
<th>1 Regime</th>
<th>2 Regimes</th>
<th>3 Regimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(0)</td>
<td>3360.014599</td>
<td>3314.029197</td>
<td>3492.043796</td>
</tr>
<tr>
<td>AR(1)</td>
<td>3181.044118</td>
<td>3176.132679</td>
<td>3429.280827</td>
</tr>
<tr>
<td>AR(2)</td>
<td>3158.088889</td>
<td>3148.312769</td>
<td>3404.733822</td>
</tr>
<tr>
<td>AR(3)</td>
<td>3145.149254</td>
<td>3134.571895</td>
<td>3388.41159</td>
</tr>
</tbody>
</table>

Our results show that for each order of autoregression the data favors a two regime model. In the three regime case in line with our merger wave theory what we hypothesize is a case where a third regime exists, with values of the mean and variance higher still than the elevated values in the second regime, possibly at the peak of a wave or existing alone as a regime of extremely high activity. We do not investigate higher then the three regime case for a number of reasons. Firstly, we have no theoretical reason to suspect a four regime case in line with our merger wave theory, secondly, a four regime model is computationally challenging to estimate and finally, the vast existing literature on MS models cases of more then three regimes are extremely rare.

6.2 Principle Component Analysis

Since we aim to incorporate the effects of sixteen different variables into the transition probabilities of our model, we use principle component analysis to ease the computational burden of this task. We create three composite indices from the resultant factor loadings for the first three principle components. The factor loadings for each index are shown in table 4 and the evolution of the indices over time is shown in figure 2. Overall the three factors capture nearly 80% of the total variance shown in the data.
### Table 4. Resultant factor loadings for first three principle components.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor1</th>
<th>Factor2</th>
<th>Factor3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNP Gross National Product</td>
<td>0.99</td>
<td>-0.04</td>
<td>-0.064</td>
</tr>
<tr>
<td>EI Economic Indicators Index</td>
<td>0.989</td>
<td>-0.027</td>
<td>-0.099</td>
</tr>
<tr>
<td>IPROD Industrial Productivity</td>
<td>0.983</td>
<td>0.076</td>
<td>0.092</td>
</tr>
<tr>
<td>BKR Bankruptcy</td>
<td>-0.955</td>
<td>-0.093</td>
<td>-0.007</td>
</tr>
<tr>
<td>LIQ Liquidity</td>
<td>0.922</td>
<td>-0.241</td>
<td>-0.184</td>
</tr>
<tr>
<td>PI Price Index</td>
<td>0.92</td>
<td>0.205</td>
<td>0.164</td>
</tr>
<tr>
<td>CPR Corporate Profits</td>
<td>0.917</td>
<td>0.063</td>
<td>-0.331</td>
</tr>
<tr>
<td>ISEN Investor Sentiment</td>
<td>0.528</td>
<td>-0.536</td>
<td>-0.108</td>
</tr>
<tr>
<td>UNPLY Unemployment</td>
<td>-0.403</td>
<td>-0.741</td>
<td>-0.016</td>
</tr>
<tr>
<td>BLIR Bank Loan Interest Rate</td>
<td>-0.373</td>
<td>0.793</td>
<td>-0.185</td>
</tr>
<tr>
<td>CSEN Consumer Sentiment</td>
<td>0.322</td>
<td>0.581</td>
<td>0.521</td>
</tr>
<tr>
<td>PE Price Earnings Ratio</td>
<td>0.316</td>
<td>-0.322</td>
<td>0.816</td>
</tr>
<tr>
<td>PMSEN Purchasing Managers Sentiment</td>
<td>0.304</td>
<td>0.025</td>
<td>0.026</td>
</tr>
<tr>
<td>SPIR Spread in Interest Rates</td>
<td>-0.186</td>
<td>-0.829</td>
<td>0.365</td>
</tr>
<tr>
<td>COUNT M&amp;A Level</td>
<td>-0.127</td>
<td>0.697</td>
<td>0.4</td>
</tr>
<tr>
<td>PVOL Price Volatility</td>
<td>0.075</td>
<td>-0.014</td>
<td>0.724</td>
</tr>
</tbody>
</table>

| Variance           | 7.3094  | 3.2005  | 1.9937  |
| % Variance         | 0.457   | 0.2     | 0.125   |

**Figure 2. Evolution of the three factor weighted indices over time**

![Factor Weighted Indexes](image-url)
The factor weightings give us an idea of the most dominant variables in each index. Index 1 captures the majority of the variance in the data and weights heavily on variables which we classify as indicators of economic conditions such as GNP, industrial productivity, securities prices, bankruptcy etc. Liquidity also weights very highly adding a dimension reflecting the ease of financing to this index. Furthermore the signs of the weights are largely in line with what we expect from the theory; for example the bank loan interest rate, the spread in interest rate, unemployment and bankruptcy, variables where we expect an inverse relation to in M&A activity, weight negatively compared to positively weighted variables such as GNP, industrial productivity, sentiment proxies etc. where we expect a direct relation with M&A levels. The second index does weight heavily on interest rates but unfortunately these show opposing signs and hence make it more difficult to interpret overall. This index also captures the effects of changing levels of M&A activity. Weighing most heavily on price-earnings ratio and price volatility the third index seems to capture the variables reflecting the idea of valuation discrepancies discussed previously.

7. Estimation Results

One of the questions faced when developing the TVTP model is that of how many lagged values of our indices to include as leading indicators for the transition probabilities; i.e. choosing the time period between the effect of our indices on firms incentives to merge and the actual observed rise in announced M&A deals recorded. We roughly gauged this could be anywhere from three to twelve months. The second question
facing us was that of choosing the autoregressive order of the model\(^4\). Town’s 1992 paper sets it at two but this is chosen somewhat arbitrarily. To solve these two problems simultaneously we ran numerous models using zero to three as autoregressive parameters and a three to twelve month lag in the indices. In the end we chose the model where the coefficients of the indices in the TVTP model were most significant. The resulting model is a first order regression incorporating the 9 month lagged value of each of our three indices.

In table 4 we show the results of both the FTP and TVTP model for our two regime model with switching in means and variances.

\textit{Table 4. Estimation results of coefficients from FTP and TVTP models along with standard deviations and significance levels.} \(\sigma\) and \(\mu\) represent the variance and mean respectively in each regime.

\textit{For explanation of other parameters see data description.}

<table>
<thead>
<tr>
<th></th>
<th>FTP</th>
<th></th>
<th></th>
<th></th>
<th>TVTP</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff</td>
<td>Std. Dev</td>
<td>Signif</td>
<td></td>
<td>Coeff</td>
<td>Std. Dev</td>
<td>Signif</td>
</tr>
<tr>
<td>(\mu_1)</td>
<td>171.441</td>
<td>12.902</td>
<td>0.000</td>
<td>153.611</td>
<td>10.332</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(\mu_0)</td>
<td>89.349</td>
<td>4.293</td>
<td>0.000</td>
<td>105.555</td>
<td>3.852</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(\sigma_1)</td>
<td>73.783</td>
<td>6.451</td>
<td>0.000</td>
<td>130.865</td>
<td>10.542</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(\sigma_0)</td>
<td>18.947</td>
<td>1.086</td>
<td>0.000</td>
<td>22.861</td>
<td>1.123</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>(\phi_{11})</td>
<td>-0.008</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td>-0.008</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>(\phi_{01})</td>
<td>-0.009</td>
<td>0.005</td>
<td>0.055</td>
<td></td>
<td>-0.009</td>
<td>0.005</td>
<td>0.055</td>
</tr>
<tr>
<td>(\phi_{12})</td>
<td>0.019</td>
<td>0.007</td>
<td>0.006</td>
<td></td>
<td>0.019</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>(\phi_{02})</td>
<td>-0.003</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td>-0.003</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>(\phi_{13})</td>
<td>0.035</td>
<td>0.002</td>
<td>0.000</td>
<td></td>
<td>0.035</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>(\phi_{03})</td>
<td>0.006</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td>0.006</td>
<td>0.001</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Log Likelihood: -1334.1158
Likelihood Ratio Test for TVTP Chi-Squared(6)= 97.727500 with Significance Level 0.00000000

\(\text{\(4\) We attempted to base this decision on AIC criteria but found that it favored unrealistic orders above 5 which hugely complicated the maximum likelihood of the model along while also causing the probability of switching to become too sensitive to small changes in the level of M&A.}\)
7.1 FTP Model

This model provides the benchmark for our study. Our results show support for the existence of the proposed two regime model. Estimates of the means and variance indicate two distinctly different regimes, one characterized by a high mean/high variance and another by low mean/low variance. The mean in the high regime is almost double that in the low regime ($\mu_1 / \mu_0 \approx 1.9$) and the high regime variance is almost four times that of the low regime ($\sigma_1 / \sigma_0 \approx 3.9$), this supports our hypothesis of a two regime system. Figure 3 plots the inferred probabilities of being in regime one over time. When close to one this indicates that the time series is the high mean/variance regime. From this figure we can see one long lasting period where the probability remains almost at one from 1994 to 2000. This coincides with the well documented fifth US merger wave (Black 2000, Lipton 2001) which began around 1993 following the economic recession of 1990-91 in the US and ended with the collapse of the millennium bubble in 2000.
Figure 3. Inferred probability of being in regime one in the FTP model. 
Graphed here with the level of M&A activity for comparative purposes.
7.2 The TVTP Model

The TVTP results in table 4 for the mean and variance again support the high and low regime classification we assume ($\mu_1 / \mu_0 \approx 1.4$, $\sigma_1 / \sigma_0 \approx 5.6$). We note that while the extension to time variance does alter the regime dependent values of the mean and variance the change is not large and the underlying hypothesis of high and low mean/variance regime classification is still strongly supported. Figure 4 plots the smoothed inferred probability of being in the high regime state over time.

*Figure 4. Inferred probability of being in regime one in the TVTP model.*

*Graphed here with the level of M&A activity for comparative purposes.*
We conduct a simple likelihood ratio test for the TVTP model under the null hypothesis of no time variation (see Table 4). The result concludes that the TVTP model is superior to the FTP case. Compared with the FTP case, the extension to time variance and dependence on the conditioning variables has the overall effect of increasing the variance in the inferred probabilities. This can be useful so far as it can give further insight into the structure of the two regime process over the time period of a perceived “wave”. For example, analysis of the case of the fifth wave between 1993 and 2000 from our model indicates that the wave was strongest from about 1998 to mid 1999. Also the small peaks in the first quarter of 1995 are the first real indication of a regime switch, this indication is absent in the results from the FTP model.

The $\phi_{xy}$ coefficients measure the effect of each of our indices ($y = 1$ to 3) on each of our time varying transition probabilities ($x=0$ for $p$, and $x=1$ for $q$). For our first index both coefficients are negative hence increases in our index with increase the probability of
a regime switch in each case (1-p_t, 1-q_t increase). This is consistent with theory that improving economic conditions can signal a merger wave. Liquidity also weights highly in this index so this finding is also consistent with the argument that a liquidity component is necessary to facilitate a merger wave (Harford 2005). Index two has a negative coefficient for p_t and a positive coefficient for q_t. This inverse relation is easier to interpret; increases in the index increase the probability of switching from low to high regime as well as increase the probability of staying in the higher regime. Overall increases favor the higher regime. Looking at the components of our index we can draw several conclusions. The level of merger activity has a large positive factor weight which implies that high levels of merger activity are a signal of high levels in the future. Unemployment and the spread in interest rates weight negatively in this index, indicating that decreases in these variables have the same effect, overall signaling that the time series will be in the high regime in the next time step. Finally the third index which represents our valuation discrepancies index has positive coefficients in both cases. For q_t this coefficient is the largest of the six. We interpret this as follows; increasing price volatility and price earnings ratios prolong the duration of the higher regime. While these variables do not directly indicate a regime switch from low to high, they provide momentum to a merger wave if they are increasing. This supports the idea that valuation discrepancies, over or under valued stocks, have an effect in merger waves.

Another advantage of the TVTP model is that we can examine changes in the transition probabilities around the times of regime switches. Figures 5 and 6 and show the time variation in p_t and q_t.
Figure 5. Time evolution of $p_t$

Figure 6. Time evolution of $q_t$
Comparing the graphs of the time evolution of the transition probabilities shows that the effect of the conditioning variables is much more significant on the evolution of the higher state i.e. the probability of remaining in the higher state or dropping from the high to low state. This is evident also from the sizes of the estimated coefficients $\varphi_{xy}$. Figures 7 and 8 and show $p_t$ and $q_t$ graphed along with the inferred probability of being in the higher state.

*Figure 7. Time evolution of $p_t$ graphed with the inferred probability of being in the higher state*

*Figure 8. Time evolution of $q_t$ graphed with the inferred probability of being in the higher state*
Focusing on the regime switch from 1994 to 2000 we can see some interesting behavior; $q_t$ begins to rise and becomes more volatile about 12 months before the inferred regime switch. It then continues to increase rapidly indicating persistence in the regime switch. While $p_t$ shows much less volatile behavior it does show a marked decrease before the end of the regime switch in 2000 as well as a prolonged decrease before the 2007 regime switch just captured. Overall extending the regime switching probabilities to incorporate time variance gives us a better insight into both the nature of the time series as well as the theorized relations between the macro economic climate and levels of merger and acquisitions activity.
8. Forecasting

Recent empirical studies examining the forecasting power of Markov Switching models repeatedly report that while the non-linear models may have a superior in sample fit, the out of sample forecasts are often inferior compared to linear models such as the random walk (Clements and Krolzig 1997, Engle 1992). However Engle does find that a Markov Switching model outperforms several generalized versions of the random walk model in predicting the direction of change of the exchange rate. Also studies have shown the relative forecast performance to be dependent on the regime present at the time of the forecast (Krolzig, 2000). These issues are well established regarding MS models with fixed transition probabilities but less work has been done examining forecasts from MS model with time varying transition probabilities.

We aim to forecast one step ahead predictions of the level of merger activity using the calculation in (15) given in Simpson, Osborn and Sensier (2001). This is specific to a first order autoregression such as ours.

\[
\hat{y}_{t+1|t} = \sum_{S_{t+1}=0}^{1} \sum_{S_t=0}^{1} \hat{\mu}(S_{t+1}) + \phi_1[y_t - \hat{\mu}(S_t)] \times Pr[S_{t+1} = s_{t+1}, S_t = s_t = S_t|Y_t, Z_t; \hat{\theta}]
\]

The second term denotes the one step ahead predicted probability distribution of the two regimes conditional. This is conditional on \(Y_t\) and \(Z_t\) which are the information available regarding the dependant variable and the vector of leading indicator variables respectively. The forecast is also conditional on the estimated parameters of the model.
contained in $\theta$. Using this equation subsequent forecasts are calculated by recursively re-estimating the model and computing the one step ahead forecast.

Results to be added.

9. Conclusions

We draw several conclusions about the behavior of the M&A time series from the results presented in this article. Firstly, our preliminary tests give substantial support to the idea that the M&A time series follows a regime switching process. Furthermore, in both models we saw very strong evidence of a high-mean-variance in wave regime and a low-mean-variance off wave regime. Both models were successful in capturing the well-noted fifth merger wave of the late 1990’s. This high/low regime classification supports the idea that merger waves do exist in the time series and furthermore characterizes a merger wave as a significant increase in both the mean and variance of the time series over a prolonged period of time.

Based on a simple likelihood ratio test we conclude that the TVTP model is superior to the FTP model in capturing the behavior of this data. Furthermore, from the estimation results of the coefficients of the conditioning variables we gained several insights into the relationship between our chosen variables and their effect on the level of M&A activity. Our results support the idea that improving economic conditions can prompt a merger wave as well as supporting the idea that a liquidity component is a very
important influencing factor of merger waves. We find that while valuation discrepancies do not seem to precipitate a merger wave; they do appear to be an important factor contributing to the persistence of the wave regime. We find that falling levels of unemployment and decreases in the spread in interest rates can facilitate a merger wave. Finally we find that previous levels of M&A activity are a good leading indicator of levels in the future.

Thus we conclude that this article presents some solid evidence for the application of Markov regime switching models with time varying transition probabilities to the dynamics of merger markets.
References


