A RE UNDERWRITING CYCLES REAL AND FORECASTABLE?

M. Martin Boyer
Eric Jacquier
Simon Van Norden

ABSTRACT
Speculative efficiency often requires that future changes in a series cannot be forecast. In contrast, series with a cyclical component would seem to be forecastable with decreases, possibly relative to a trend, during the upper part of the cycle and increases during the lower part. On the basis of autoregressive model (AR) estimates, it is considered that there is strong evidence of cycles in insurance underwriting performance as measured by the premium-to-loss ratio. Indeed, a large literature attempts to explain this documented cyclicality. First, we show that the parameter estimates from AR models do not lead to any such inference and that in the contrary, the evidence in the data is consistent with no cyclicality at all. Second, we show that a number of different filters lead to the same conclusion: that there is no evidence of in-sample or out-of-sample predictability in annual insurance underwriting performance in the United States.

INTRODUCTION
The so-called “liability crisis” of 1985 through 1986 produced a large volume of literature that attempted to explain why insurance markets go through periods of high profitability followed by periods of low profitability. This succession of high- and low-profitability periods is known in the academic and professional insurance literature as an “underwriting cycle.” Since 1950, there have been six or seven such underwriting cycles, lasting between 6 and 7 years on average in the United States (see Venezian, 1985; Cummins and Outreville, 1987). These cycles appear to exist as well in other Organisation for Economic Co-operation and Development (OECD) countries. Cummins and Outreville (1987) document periods ranging from 4.7 years in Australia to 8.2 years in France (see Lamm-Tennant and Weiss, 1997; Chen, Wong, and Lee, 1997; Meier, 2006b, for other international evidence). As Weiss (2007) states, however, cycles seems to be lengthening or vanishing completely. For
instance, Meier (2006a) shows that for the United States, cycle length increases to 10 years if we include different explanatory variables in the regression.

Many theories have been put forward to explain the existence of underwriting cycles in the property and casualty insurance business: Forecasting errors (Venezian, 1985; Boyer, Eisenmann, and Outreville, 2011), insurer moral hazard (Harrington and Danzon, 1994), arbitrage theory (Cummins and Outreville, 1987), risky debt (Cummins and Danzon, 1997), interest rate variations (Doherty and Kang, 1988; Doherty and Garven, 1995), and underwriting capacity constraints (Gron, 1994; Niehaus and Terry, 1993; Winter, 1994). For a more in-depth literature review, see Harrington (2004) and Weiss (2007).

A literature reports evidence of predictability and cyclicality in underwriting cycles. For instance, Haley (1993, 1995) and Choi, Hardigree, and Thistle (2002) find a negative cointegrating relationship between the underwriting margin and the risk-free rate. Such evidence could have an implication on the modeling of the insurance sector. However, one first needs to review the existing evidence. Indeed, the experience from macroeconomic and empirical asset pricing shows us that evidence of predictability can be elusive for two reasons. First, when analyzing in-sample time series properties, cyclicality is typically overstated by standard estimation techniques. Second, the out-of-sample performance of a model is adversely affected by the amount of data mining as well as the instability of the parameters, even in the absence of data mining. The latter typically causes some inevitable degree of model misspecification. With these issues in mind, we reinterpret the predictability and cyclicality of insurance profitability measures for the United States.

Our results are surprising for researchers who have studied insurance underwriting data. Using time series techniques previously used to understand economic cycles (GDP growth and inflation) and stock returns, we find that any evidence of underwriting cycles in the property and casualty insurance market could simply be spurious. Namely, we find that underwriting data in the United States follow time series processes from which no profitable investment strategy or product pricing can be developed. This means that there is no significant evidence against the null hypothesis of speculative efficiency. Therefore, to the extent that cycles may exist, they do not appear to help speculators or sophisticated underwriters forecast movements in underwriting ratios and industry profitability.

Our investigation of insurance cycles is motivated in part by the results in Harrington and Yu (2003) and Jawadi, Catherine, and Sghaier (2009). Harrington and Yu examine loss ratios, expense ratios, combined ratios, and economic loss ratios for many insurance lines and for all lines combined from 1953 through 1998. After careful consideration of multiple unit root tests and their power to detect unit roots, they conclude that most of these series appear to be trend stationary. This in turn implies that deviations from trend are expected to diminish over time and, if the trend is known with certainty, that these deviations will have a predictable component. When these deviations are persistent, such series are commonly characterized as containing a short-term cyclical component and a long-run deterministic trend. Jawadi, Catherine, and Sghaier find that the convergence to equilibrium of property and casualty premiums toward equilibrium is time varying depending on the importance of the disequilibrium. These authors do not study the loss or combined ratio, however.
The remainder of the article is as follows. First, we review the in-sample evidence and properly characterize the times series process of the data. Drawing from the empirical literature on departures from rational expectations, we test for deviations from the random walk. In addition to standard tests, for example, generalized method of moments or variance ratios, we also use a test powerful under cyclical alternatives. We then measure predictability and cyclicality within the framework of a classic linear autoregressive model (AR). Because cyclicality is generally overstated by the classical estimation of these ARs, we compare competing models and hypotheses using posterior analysis and odds ratios, paying special attention to the effect of the prior information. Second, we conduct out-of-sample diagnostics to assess the economic significance of the in-sample cyclicality. This approach allows us to verify whether the best estimated models in-sample produce out-of-sample forecasts that are superior to naïve benchmarks consistent with unpredictability. Finally, we conclude with a discussion of the economic implications of our results.

**In-Sample Evidence of Cyclicality**

We first review the evidence in well-known empirical studies. We then conduct certain reality checks and show the limitations of the evidence. The data used in our study were acquired from “Best’s Aggregates and Averages” and spans the years 1967 through 2004. This is the main source of information for property and casualty insurance cycle studies in the United States (see, among many others, Cummins and Outreville, 1987; Lamm-Tennant and Weiss, 1997; Harrington and Niehaus, 2001; Harrington and Yu, 2003; Meier, 2006a, 2006b).

**Background**

The insurance industry goes through periods of high and low premiums relative to losses. These peaks and troughs culminated during the so-called liability crisis of 1985 through 1986. This produced a large literature, which attempted to explain why insurance markets cycle through periods of high and low profitability. The literature in insurance agrees that there have been seven such cycles in the United States since 1950. The typical methodology is twofold. First, it uses econometric models of univariate cycles to document the existence of cycles. Second, it uses regressions to link the variations in premiums to exogenous variables suggested by theoretical models. Herein, we concentrate on the univariate evidence, and will argue that properly used time series methods neither produce any evidence of cyclicality, nor of predictability owing to cyclicality.

Consider the standard univariate AR:

\[
\Phi(B)q_t = \mu + \epsilon_t \quad \epsilon_t \sim i.i.d. N(0, \sigma),
\]

where \(q_t\) is the underwriting ratio, and \(\Phi(B)\) is a polynomial of degree \(p\) in the backshift operator \(B\). When \(p\) is larger than 1, the characteristic equation may have pairs of complex roots, if the corresponding determinant is negative. Then, for each such pair, of norm \(||\lambda||\) and real part \(R(\lambda)\), there is a cycle in the autocorrelation...
function of $R_t$, with period

$$\tau = \frac{2\pi}{\arccos \left( \frac{R(\lambda)}{||\lambda||} \right)}.$$  \hfill (2)

This suggests a maximum likelihood estimator of $\tau$ based upon the maximum likelihood estimator of the complex roots, which itself follows from the maximum likelihood estimators of $\Phi$, the vector of autoregressive parameters in Equation (1).

The insurance literature essentially uses this method, reporting the estimates $\hat{\Phi}$ and the corresponding period of a cycle, when there is one. Typically, standard errors for $\hat{\Phi}$ pointing toward strong statistical significance are used to assess the validity of the model. Table 1 shows the results of some well-known studies.

Cummins and Outreville (1987) analyze annual data for 13 countries from 1957 to 1979. They find $R$-squares between 0.16 and 0.90, and the existence of a cycle for 10 countries. When there is a cycle, the periods range from 4.7 years (Australia) to 11.7 years (France). The cycle period exceeds 6 years for 8 out of these 10 countries.
Given that 23 years of data cover less than four cycles of 6 years, one suspects that these periods are not estimated with much precision. They also study the automobile insurance for six countries, all with statistically significant coefficients, high $R$-squares, and a cycle. Note again that the average cycle period is high, 7.1 years, which amounts to about three cycles for the entire sample of 23 years; the periods range from 5.2 years (Switzerland) to 9.9 years (Italy).

Chen, Wong, and Lee (1997) also report cycles. They analyze four lines of insurance business in five Asian countries from 1970 to 1995 and report a cycle for 10 estimates. For these 10 cycles, the period exceeds 10 years for three series, which is fewer than two cycles over the calendar span. This leaves seven estimates out of 25 for which one can hope for some acceptable degree of precision. Yet the authors conclude:

The results of the second-order autoregressive model largely support the existence of the underwriting cycle in Asia because underwriting cycles are found in at least one line of all five Asian countries tested.

Lamm-Tennant and Weiss (1997), not reported in our Table 1, analyze 23 years of OECD data from 1965 to 1987. They report period estimates for six lines of insurance as well as the overall country cycle for 12 countries. Overall, they estimate AR(2) models on 80 series of insurance loss ratios. They find a cycle for 49 series, that is, 61 percent of the series. On the basis of this evidence they conclude, “This study further substantiates the presence of underwriting cycles in the average loss ratio and by-line loss ratios for a sample of twelve countries.”

Harrington and Niehaus (2001) analyze the combined ratio (loss ratio plus expense ratio) for U.S. data, but different time periods. They find evidence of cycles with periods averaging 8 years. They conclude that:

the evidence of second-order autoregression . . . must be considered anomalous from the perspective of the perfect markets model . . . (and) because there is no reason to expect that shocks are predictable, the evidence of second-order autoregression is combined ratios or the other variables also is not readily explained by shock models.

Finally, Meier (2006a, 2006b) expands the Cummins and Outreville (1987) study by adding exogenous variables to the AR(2) model. The cycle periods increase to approximately 10 years on average in the United States and Switzerland, compared to 5 to 6 years in Cummins and Outreville. The author concludes that “the theory of cycles of about six years in length . . . may no longer be capable of adequately explaining the development of the insurance markets in the 1980s and 1990s.” We report some of these multivariate results with exogenous variables in Table 1, Panel B.

Despite these authors’ strong conclusions, we do not believe that these previous results are strong enough to conclude, in-sample, that underwriting cycles are anything but spurious. Next, we show that such estimates do not allow one to conclude that there is any evidence that would support the existence of univariate cycles in property and casualty loss ratios. The key to understanding the time series properties of cycles (including underwriting cycles) lies in the nonlinear relationship between the
TABLE 2
Univariate AR(p) Estimation on Underwriting Ratio 1967 to 2004

<table>
<thead>
<tr>
<th>AR(2) With Time Trend: q_t = \phi_0 + \phi_1 q_{t-1} + \phi_2 q_{t-2} + \gamma \times \text{trend}</th>
<th>\phi_1</th>
<th>\phi_2</th>
<th>\phi_3</th>
<th>\gamma</th>
<th>%R^2</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.94</td>
<td>-0.40</td>
<td>-</td>
<td>0.08</td>
<td>58</td>
<td>8.6</td>
<td></td>
</tr>
<tr>
<td>(5.6)</td>
<td>(-2.3)</td>
<td>(1.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AR(3): q_t = \phi_0 + \phi_1 q_{t-1} + \phi_2 q_{t-2} + \phi_3 q_{t-3}</th>
<th>\phi_1</th>
<th>\phi_2</th>
<th>\phi_3</th>
<th>\gamma</th>
<th>%R^2</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.94</td>
<td>-0.27</td>
<td>-0.05</td>
<td>-</td>
<td>56</td>
<td>10.9</td>
<td></td>
</tr>
<tr>
<td>(6.1)</td>
<td>(-1.3)</td>
<td>(-0.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The t-statistics are reported in parentheses.

basic parameter estimate of \( \Phi \) and the cycle period \( \tau \). This relationship appears to have been either misunderstood in the insurance literature, or at least judged to be irrelevant.

Reality Check: Bayesian Inference

This section provides two sets of results based on an in-sample assessment of cycles in the property and casualty insurance industry in the United States. Our first reported results will show that the main results presented by insurance economists, as reported previously, do not constitute evidence of cyclicity. Our second result will show that the period estimates are uninformative.

Aforementioned studies often add a linear time trend to an AR(2) to detrend the data. As shown in Nelson and Plosser (1988), Rudebusch (1993), and Stock and Watson (1988), the detrending method has a strong influence on the estimated deviation from the trend (see Harrington and Yu, 2003, for an application of these techniques to underwriting margins). This means that a simple linear time trend may not be the ideal way to stationarize data, in which case alternative methods are often preferred. For instance, Geweke (1988) argues in favor of using an AR(3) process that allows for a unit root as a preferable alternative to adding a linear trend to the AR(2) process. Table 2 reports the estimation of both the AR(2) process with a time trend and the AR(3) process for our time series data that spans 38 years from 1967 to 2004.

These results look somewhat similar to those obtained in the existing literature, albeit with a longer calendar span. Clearly, the point estimate of the period is sensitive to small changes in \( \phi_2 \) and the presence of the preferred stochastic rather than a deterministic trend. The period point estimate are 8.6 and 10.9 years. Although our argument is not dependent on the AR(2) versus AR(3) issue, the results in the table do raise a first important issue, which is the reliability of the point estimate of the period. This issue is of paramount importance when the point estimate for the period is high relative to the calendar span of the data. Here, an 11-year period means that the data barely cover three cycles. What is clearly missing in the existing literature is a specification of the uncertainty surrounding the point estimate of the period. More fundamental is that the existing results do not produce the probability of the existence of a cycle, which can be inferred from the data. The estimation method in
the empirical insurance literature only computes a point estimate of the period given the point estimates of $\phi_1$.

Estimating the standard error of the point estimate of the cycle period is easily done. A first approach would be to invoke asymptotic theory and the $\Delta$-method, even though an asymptotic framework is not the most appropriate approach for such a short (38 observations) time series. This approach gives an approximate expression for the variance of the estimator of any function of parameters estimated by maximum likelihood. Namely, the asymptotic variance of $\hat{T}$ based upon the covariance matrix of $\hat{\Phi}$, and an expansion of $T(\hat{\Phi})$ around the true value $\Phi$ can be shown to be

$$V(\hat{T}) \approx T' \Sigma_{\phi} T,$$

where $\Sigma_{\phi}$ is the variance–covariance matrix of $(\hat{\phi}_1, \hat{\phi}_2)$. An estimate of $V(\hat{T})$ is obtained by substituting for the estimate of $\Sigma_{\phi}$. Such estimates reveal very high standard errors, basically implying that the point estimates of cycle periods in the literature are not informative.

As it results from asymptotic theory, the delta method is possibly misleading in small samples. An exact, small-sample approach should be used to calculate the cycle period by simple Monte Carlo estimation as in Geweke (1988) and Jacquier (1991). Namely, one draws from the posterior distribution $\Phi | D$, with mean $\hat{\Phi}$ and covariance matrix $\Sigma_{\phi}$. One then computes the determinant and the roots of the characteristic equation for each draw. Clearly, even if $\hat{\Phi}$ lies in the cyclical region, many draws do not due to the uncertainty of $\Phi$. The mere counting of the draws that lie in the cyclical region yields the posterior probability of the existence of a cycle, which is not reported in the insurance literature. In turn, each draw resulting a cycle yields a draw for the period of this cycle. These draws describe the posterior density of the cycle’s period given that there is a cycle.$^1$

So what does one expect the distribution of the period to look like? The draws can be represented on a histogram, and quantiles can be computed. Figure 1 shows such a histogram constructed with 50,000 draws. The median of the period in Figure 1 is 3.8 years; 55 percent of the probabilities lie between 3 and 6 years. This appears quite informative and, as intuition would have it, constitutes a precise inference on the period. However, Figure 1 was not constructed with draws from the parameters in the literature such as in Tables 1 or 2. In fact, we assumed no knowledge of $\phi_1, \phi_2$ other than imposing stationarity on the AR(2) process. We drew from a uniform distribution in the well-known triangle of stationarity of an AR(2) (see Zellner, 1971, p. 196; Sargent, 1986). Specifically, we imposed $\phi_2 > -1, \phi_2 > -\phi_1 + 1$, and $\phi_2 < \phi_1 + 1$. Figure 1 represents the prior density of the period, the density representing uncertainty before looking at the data.

What Figure 1 shows is that, due to the nonlinearity of $T$ with respect to $\Phi$, diffuse information on $\Phi$ implies a quite informative-looking prior on the period $T$. As for the probability of observing a cycle given the random distribution of $\phi_1$ and $\phi_2$, it is

$^1$ For a standard analysis, with diffuse or conjugate prior, the posterior density of the AR parameters is a multivariate Student-$t$. See Zellner (1971) and many others.
easy to show by analytical integration that it is two-thirds, a probability disturbingly higher than 50 percent. It can be computed analytically as the area below the parabola $\phi_1^2 + 4\phi_2 < 0$ and above the base of the base of the triangle $\phi_2 > -1$, where the AR(2) has two complex roots (see Zellner, 1971, p. 196). For an AR(3), it can easily be shown that this prior probability of the existence of a cycle jumps to 93 percent.

So any posterior estimate of the probability of a cycle should be higher than 66 percent; otherwise, it means that the data brought evidence against cyclicality. Posterior densities for the period of the cycle should look tighter than in Figure 1. Looking back at the period estimates in Tables 1 and 2, they now appear remarkably uninformative. Namely, their posterior distribution is no more precise than inferred by knowing nothing about the parameters. Consider the evidence of the very existence of a cycle; this simple output of estimation is not provided by the empirical studies in insurance. We can, however, compare the 66 percent probability of a cycle implied by knowing nothing about the parameters to the frequency of cycles reported in the literature: 10 out of 13 in Cummins and Outreville (1987), 10 out of 25 in Chen, Wong, and Lee (1997), 49 out of 80 in Lamn-Tennant and Weiss (1997), and 11 out of 16 in Meier (2006b) for a total of 80 out of 134, or 60 percent. It appears that the literature has inadvertently started the analysis with strong priors in favor of the existence of a cycle. To remedy this, especially given the short samples available, one needs to
compare the posterior density of the period with the prior density implied by the
priors on the parameters.\(^2\) One can compute odds ratios for the competing models
and alternatives so as to take into account the short sample size. Odds ratios allow a
“horse race” between multiple competing alternatives.

This Bayesian exercise allows us to suggest that underwriting cycles in the property
and casualty insurance industry exist only because we expect them to exist. Similar
to the debate surrounding Rorschach ink blot tests, where an objective meaning
is thought to be extracted from responses to meaningless blots of ink, our results
suggest that insurance economists studying underwriting cycles attempted to extract
a cyclical pattern from nothing more than a random walk.

**Out-of-Sample Evidence of Cyclicality**

Speculative efficiency requires that future changes in a series cannot be forecast.
In contrast, series with a cyclical component would seem to be forecastable with
predicted decreases (perhaps relative to a trend) during the upper part of the cy-
ble and increases during the lower part. It is therefore crucial to determine the ro-
ustness of the evidence of cyclicality and predictability. This is the purpose of this
article.

The analysis presented in the previous section of the article examined whether cycles
can be identified *ex post*. For questions of speculative efficiency however, the key issue
is whether cycles can help forecast *ex ante* future changes in levels. Once we allow for
measurement error in our estimates of the cycle, the former need not imply the latter.
To understand why, it may be useful to consider the literature on business cycles.

**Business Cycles: Ex Post and Ex Ante**

It is often observed that the estimates of business cycles undergo large revisions. For
example, Orphanides (2003) analyzes the successive estimates of the cycle in U.S.
Real Output at the end of 1974. He reports that in 1976, economists estimated that
the output was at 14 percent below trend, but this estimate was revised to 9 percent
by 1979 and to 4 percent by 1994. Although extreme, this episode highlights the fact
that the subsequent behavior of a time series may have an important influence on
our perception of cycles. Orphanides and van Norden (2002) show that, as a result,
most common methods of estimating business cycles are subject to large revisions
over time.

The existence of such revisions means that speculative efficiency could be compati-
ble with the presence of some types of cycles. As discussed previously, speculative
efficiency simply requires that changes in the level of the series are not forecastable.
More precisely, this means that contemporary, or end-of-sample, estimates of the cycle
cannot help predict future changes in levels. It is noteworthy to mention, however,
that the analysis presented earlier focused on historical, or in-sample, estimates of the
cycle. It should not be assumed that these are similar to estimates economic agents
would have formed based only on the information available to them at the time. In
the case of business cycles for example, Orphanides and van Norden (2005) find that

---

\(^2\) One can use modified prior distributions, that is, diffuse priors on the roots themselves, to
alleviate this problem.
although in-sample estimates of the business cycle were very useful in predicting inflation, end-of-sample estimates were not.

To better understand the relationship between speculative efficiency and insurance cycles therefore, we now turn to consider the behavior of end-of-sample estimates. Specifically, we consider whether contemporary estimates of cycles are sufficiently accurate to enable underwriters to forecast changes in the loss ratios. To do so, we follow the experimental design along the lines of that in Orphanides and van Norden (2005). We begin by reviewing, in the next subsection, a variety of different estimators of cycles and thereafter discuss tests of forecast accuracy.

Estimating Cycles

The estimation of the cycle involves a detrending method that decomposes the underwriting ratio, \( q_t \), into a trend component, \( \mu_t \), and a cycle component, \( z_t \).

\[
q_t = \mu_t + z_t. \tag{4}
\]

Some methods use the data to estimate the trend, \( \mu_t \), and define the cyclical component as the residual. Others specify a dynamic structure for both the trend and cycle components and estimate them jointly. We examine detrending methods that fall into both categories.

**Deterministic Trends.** The first set of detrending methods we consider assumes that the trend in the variable of interest (output or loss ratio) is well approximated as a simple deterministic function of time. The linear trend is the oldest and simplest of these models and has been used in the insurance literature as mentioned previously. The quadratic trend is a popular alternative in the output forecasting literature.

**Mechanical Filters.** Dissatisfaction with the assumption of deterministic trends has led many to prefer cycle models that allow for stochastic trends. Harvey (1989) popularized the structural time series model, which attempts to separately characterize the dynamics of the trend and the cycle. A simpler (but related) approach prefers to remain agnostic about their dynamics, or to simply model their joint dynamics as some autoregressive integrated moving average (ARIMA) process. We examine three examples of this approach: the Hodrick–Prescott filter, the band-pass filter, and the Beveridge–Nelson decomposition. The distinction between these approaches and the structural time series approach is somewhat arbitrary as any of these models can be related to a particular unobserved components model. This is rarely done in practice however.

The earliest approach was to model potential output with a smoothing spline, as shown in Henderson (1924). A popular smoothing spline nowadays is obtained using the filter proposed by Hodrick and Prescott (1997), and known as the HP filter. We apply the HP filter with a smoothing parameter of 400 for annual data.
**Band-Pass Filter.** Another approach to cycle-trend decomposition is via the use of band-pass filters in the frequency domain. The clearest exponent of this approach is Baxter and King (1999), who suggest the use of truncated versions of the ideal (and therefore infinitely long) band-pass filter with a band passing fluctuations with durations between 6 and 32 quarters in length. Stock and Watson (1998) adapt this for use at the end of data samples by padding the available observations with forecasts from a low-order AR model fit to the data series. We use a filter tuned to capture fluctuations between 2 and 8 years in length and pad using an AR(2) forecast.

**Beveridge–Nelson Decomposition.** A very different approach is that of Beveridge and Nelson (1981), who consider the case of an ARIMA(p,1,q) series \( y_t \), which is to be decomposed into a trend and a cyclical component. For simplicity, we can assume that all deterministic components belong to the trend component and have already been removed from the series. Since the first difference of the series is stationary, it has an infinite-order MA representation of the form

\[
\Delta y_{t} = \epsilon_{t} + \beta_{1} \cdot \epsilon_{t-1} + \beta_{2} \cdot \epsilon_{t-2} + \cdots = \epsilon_{t}, \tag{5}
\]

where \( \epsilon \) is assumed to be an innovations sequence. The change in the series over the next \( s \) periods is simply

\[
y_{t+s} - y_{t} = \sum_{j=1}^{s} \Delta y_{t+j} = \sum_{j=1}^{s} \epsilon_{t+j}. \tag{6}
\]

The trend is defined to be

\[
\lim_{s \to \infty} E_{t}(y_{t+s}) = y_{t} + \lim_{s \to \infty} E_{t} \left( \sum_{j=1}^{s} \epsilon_{t+j} \right), \tag{7}
\]

From Equation (5), we can see that

\[
E_{t}(\epsilon_{t+j}) = E_{t}(\epsilon_{t+j} + \beta_{1} \cdot \epsilon_{t+j-1} + \beta_{2} \cdot \epsilon_{t+j-2} + \cdots) = \sum_{i=0}^{\infty} \beta_{j+i} \cdot \epsilon_{t-i}. \tag{8}
\]

Since changes in the trend are therefore not forecastable, this has the effect of decomposing the series into a random walk and a cyclical component, so that

\[
y_{t} = \tau_{t} + c_{t}, \tag{9}
\]

where the trend is

\[
\tau_{t} = \tau_{t-1} + e_{t}, \tag{10}
\]
and $e_t$ is white noise. Morley, Nelson and Zivot (2003) discuss the relationship between this model of the cycle and the unobserved components models, which we discuss next. To use the Beveridge–Nelson decomposition one must therefore

1. Identify $p$ and $q$ in the ARIMA($p,1,q$) model.
2. Identify the $\{\beta_j\}$ in Equation (5).
3. Choose some large enough but finite value of $s$ to approximate the limit in Equation (7).
4. For all $t$ and for $j = 1, \ldots, s$, calculate $E_t(e_{t+j})$ from Equation (8).
5. Calculate the trend at time $t$ as $y_t + E_t(\sum_{j=1}^{s} e_{t+j})$ and the cycle as $y_t$ minus the trend.

Based on results for the full sample, we use an ARIMA(1,1,2), with parameters re-estimated by maximum likelihood methods before each recomputation of the trend.

**Unobserved Component Models.** Unobserved component models offer a general framework for decomposing output into an unobserved trend and a cycle, allowing for an explicit dynamic structure for these components. We examine three such alternatives, by Watson (1986), by Harvey (1985) and Clark (1987), and by Harvey and Jaeger (1993). All are estimated by maximum likelihood.

The Watson model modifies Harvey’s linear level model to allow for greater business cycle persistence. Specifically, it models the trend as a random walk with drift and the cycle as an AR(2) process:

$$
\mu_t = \delta + \mu_{t-1} + \eta_t, \quad (11) 
$$

$$
z_t = \rho_1 \cdot z_{t-1} + \rho_2 \cdot z_{t-2} + \epsilon_t. \quad (12)
$$

Here $\epsilon_t$ and $\eta_t$ are assumed to be i.i.d. mean-zero, Gaussian and mutually uncorrelated and $\delta$, $\rho_1$, and $\rho_2$, and the variances of the two shocks are parameters to be estimated (five in total).

The Harvey–Clark model (CL) similarly modifies Harvey’s local linear trend model:

$$
\mu_t = g_{t-1} + \mu_{t-1} + \eta_t, \quad (13) 
$$

$$
g_t = g_{t-1} + \nu_t, \quad (14) 
$$

$$
z_t = \rho_1 \cdot z_{t-1} + \rho_2 \cdot z_{t-2} + \epsilon_t. \quad (15)
$$

Here $\eta_t$, $\nu_t$, and $\epsilon_t$ are assumed to be independently and identically Gaussian distributed random variables with mean-zero and mutually uncorrelated processes, whereas $\rho_1$, $\rho_2$, and the variances of the three shocks are parameters to be estimated (five in total).
TABLE 3
Root Mean Squared Error of Forecasts for Changes in the Underwriting Ratio, 1967 to 2004

<table>
<thead>
<tr>
<th>Model</th>
<th>$h = 1$ Year</th>
<th>$h = 2$ Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear trend</td>
<td>5.25</td>
<td>6.31</td>
</tr>
<tr>
<td>Quadratic trend</td>
<td>4.88</td>
<td>7.50</td>
</tr>
<tr>
<td>Hodrick–Prescott</td>
<td>4.89</td>
<td>5.69</td>
</tr>
<tr>
<td>Band-pass filter</td>
<td>4.75</td>
<td>6.99</td>
</tr>
<tr>
<td>Beveridge–Nelson</td>
<td>5.37</td>
<td>7.00</td>
</tr>
<tr>
<td>Watson</td>
<td>4.38</td>
<td>5.92</td>
</tr>
<tr>
<td>Harvey–Clark</td>
<td>4.32</td>
<td>6.12</td>
</tr>
<tr>
<td>Harvey–Jaeger</td>
<td>4.42</td>
<td>5.71</td>
</tr>
<tr>
<td>Benchmark</td>
<td>4.42</td>
<td>5.82</td>
</tr>
</tbody>
</table>

Forecast Evaluation

Constructing the Forecasts. Let $\Delta q^h_t = q_t - q_{t-h}$ denote the change in the underwriting ratio over $h$ years ending in $t$. We will examine forecasts of this change over horizons of 1 and 2 years. Note that we assume reporting lags imply that the underwriting ratio for period $t$; that is $q_t$, is not first reported until period $t+1$. Thus, a 1-year-ahead forecast is a forecast of the change between the first and second years after the last period for which data are available.

We examine simple linear forecasting models of the form:

$$\Delta q^h_{t+h} = \alpha + \sum_{i=1}^{m} \gamma_i \cdot z_{t-i} + e_{t+h},$$

where $m$ denotes the number of lags of the cycle included in the equation. To provide a benchmark for comparison, we estimate the same equation after replacing lags of the estimated cycle $z_{t-i}$ with lags of the underwriting ratio $q_{t-i}$.

We use this equation to construct feasible forecasts that closely mirror the forecasts speculators could construct. All parameters are estimated by ordinary least squares. For each period in which a forecast is made, the parameters of the forecasting equation (and the underlying model of the cycle $z_t$) are reestimated using the latest available data. In addition, the lag lengths for the explanatory variables are reestimated each period using the Bayes information criterion. The fitted value of the equation using the last available data point then becomes the forecast change in the underwriting ratio. Table 3 summarizes the results of the forecasting experiment.

3 For the Watson, Harvey–Clark, and Harvey–Jaeger models, we use smoothed rather than filtered estimates of the cycle. This reflects the common practice of practitioners, which is to use the most accurate possible estimate of the cycle in estimating their forecast equations. Orphanides and van Norden (2005) find that their results were insensitive to this distinction. As the cycle estimates appear to exhibit relatively smaller revisions, we expect the results here to be even less sensitive to this distinction.
Table 3 reports the root mean squared (RMS) forecast errors for each of the measures of the cycle, as well as the benchmark which simply uses the underwriting ratio. The smaller is the RMS error, the better is the forecasting model. We calibrated the 10 models using data up to 1980, so that the first forecast period is 1981. We see that the benchmark model produces forecasts more accurate than most of the cyclical models at both the $h = 1$-year and $h = 2$-year forecast horizons. The best-performing cyclic models produced forecasts only slightly more accurate than those of the benchmark model. It would therefore be interesting to determine whether any of the cyclic models produce forecasts that are significantly more reliable, or whether this result could simply be due to chance.

**Evaluating Forecast Performance.** Figure 2 shows the estimated cycles for each method. We wish to test the null hypothesis that forecasts from a given cyclic model have the same RMS forecast error as that of the benchmark model. Various tests of equality of forecasting accuracy have been proposed, notably asymptotic tests by Diebold and Mariano (1995) and finite-sample refinements by Harvey, Leybourne, and Newbold (1998). We calculated both. As they gave the same conclusions; we only report the Diebold and Mariano statistics in Table 4.

The first and third columns show the Diebold–Mariano (DM) statistics for a test of the null hypothesis that the indicated model has the same RMS forecast error as that of the benchmark model. The second and fourth columns show the associated $p$-values. Note that this is a two-tailed test; negative DM statistics indicate that the cyclical model performed worse than the benchmark model. Evidence against market efficiency requires a positive statistic, indicating that the cyclical model forecasts better than the benchmark model. We see that none of the models have a forecast performance significantly better than that of the benchmark model for any reasonable significance level; $p$-values are notoriously hard to interpret due to the Lindley–Smith paradox. Furthermore, when multiple $p$-values are produced, their naïve use (rather than a Bonferroni adjustment) overstates the evidence against the null. There is no such problem here as we report no evidence against the null. Moreover, over a 2-year forecast, only one model performs better than the benchmark, as we see from the negative DM statistic for all cyclical models except the Harvey–Jeager model.

Given these results, we can only conclude that there appears to be no evidence that conventional end-of-sample measures of cycles imply a violation of speculative efficiency. Put differently, and given our current knowledge of time series econometrics, our results imply that it is impossible to reliably forecast loss ratios for the property and casualty insurance industry using only their past values.

Could the lack of evidence of predictable variation in loss ratios be due to the lack of power in the methods used to detect cycles? In the Appendix we report the results of a simulation experiment that investigates the power of these methods to detect cycles when they exist. These results show that when cycles are sufficiently important, these methods detect statistically significant evidence of cycles in the majority of cases,

---

4 Many authors have noted that such tests are not appropriate for forecasts from estimated models when the models are nested. Because our models have different explanatory variables, they are not nested.
FIGURE 2
Estimates of Cycles for Several Methods
Table 4
Diebold–Mariano Test of Equality of Predictive Accuracy for Changes in the Underwriting Ratio

<table>
<thead>
<tr>
<th>Model</th>
<th>1 Year</th>
<th></th>
<th>2 Years</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DM</td>
<td>p-value</td>
<td>DM</td>
<td>p-value</td>
</tr>
<tr>
<td>Linear trend</td>
<td>-1.68</td>
<td>0.09</td>
<td>-1.79</td>
<td>0.07</td>
</tr>
<tr>
<td>Quadratic trend</td>
<td>-1.66</td>
<td>0.09</td>
<td>-1.63</td>
<td>0.10</td>
</tr>
<tr>
<td>Hodrick-Prescott</td>
<td>-1.16</td>
<td>0.24</td>
<td>-0.35</td>
<td>0.73</td>
</tr>
<tr>
<td>Band-pass filter</td>
<td>-0.62</td>
<td>0.53</td>
<td>-1.41</td>
<td>0.16</td>
</tr>
<tr>
<td>Beveridge–Nelson</td>
<td>-1.78</td>
<td>0.07</td>
<td>-1.13</td>
<td>0.26</td>
</tr>
<tr>
<td>Watson</td>
<td>0.09</td>
<td>0.92</td>
<td>-0.12</td>
<td>0.90</td>
</tr>
<tr>
<td>Harvey–Clark</td>
<td>0.17</td>
<td>0.86</td>
<td>-0.34</td>
<td>0.73</td>
</tr>
<tr>
<td>Harvey–Jaeger</td>
<td>0.00</td>
<td>0.99</td>
<td>0.15</td>
<td>0.88</td>
</tr>
</tbody>
</table>

even with rather short calendar spans for the data. The absence of similar results for existing loss ratios suggests that underwriting cycles are either absent or relatively small.\(^5\)

Conclusion

Our investigation of insurance cycles was motivated in part by the results of Harrington and Yu (2003), who conclude that most time series of insurer profitability appear to be trend stationary. This implies that deviations from trend are expected to diminish over time and, if the trend is known with certainty, that these deviations will have a predictable component. When these deviations are persistent, such series are commonly characterized as containing a short-term cyclic component and a long-run deterministic trend. Our results provide an updated assessment of the existence of cycles in the property and liability insurance industry in the United States. In as much as the existence of underwriting cycles in property and liability insurance is well established in the insurance economic literature, there is little evidence that insurers are able to forecast these cycles to make a profit. There are two reasons that explain this forecasting inability. The first is that standard estimation techniques overstate the likelihood of having a cycle in that the tests are biased in favor of finding one. The second reason that explains cycle unpredictability is that out-of-sample tests are not powerful enough to yield any robust economic inference.

\(^5\) A referee wondered whether the tests might lack power because “the statistical models to capture the cyclical pattern...are...outdated” and therefore the article might lack a foundation to dismiss the presence of cycles. The plethora of cyclical models used here covers the spectrum of modern time series analysis. Therefore, there is little risk of lack of power due to a misspecification of any one cyclical model. Further, we are unaware of any model of cyclical behavior that has been posited for the insurance industry and cannot be detected by one or more of the models used here. Finally, recall that Bayesian methods are somewhat immune to this notion of the power of the test; they basically report the evidence in the data without favoring an alternative or the other.
Using the same traditional time series as in previous studies on underwriting cycles, which is the annual loss ratio in the U.S. property and liability insurance market, we first apply in-sample tests of the cyclicality of time series data. Our data span the years 1967 through 2004. We show that the results from AR estimation as used in Cummins and Outreville (1987), Lamm-Tennant and Weiss (1997), and others must be interpreted with care. Namely, although we find AR(2) and AR(3) parameters similar to those found in these studies, we show that naïve inference on the existence and the period of a cycle based on the point estimates is misleading. Namely, it is strongly biased in favor of finding a cycle. When we correct for this bias, we no longer find any evidence of a cycle.

We then analyze the existence of underwriting cycles by looking at the in-sample and out-of-sample predictability of these time series with several different filters often used in macroeconomic forecasting. Using eight different filters that attempt to test for the presence of cycles, we find that none yield any forecasting power that would be profitable for insurance companies.

Our results have the interesting economic implication that underwriting cycles in the property and casualty insurance market are normal time series processes from which no profitable investment strategy or product pricing can be developed. We fail to find any significant evidence against the null hypothesis of (weak form) speculative efficiency. Therefore, to the extent that cycles may exist, they do not appear to help speculators forecast movements in underwriting ratios.

**APPENDIX: POWER OF FORECAST TESTS FOR CYCLES**

An alternative explanation for the lack of evidence of predictable variation in loss ratios is that our forecast tests lack power to detect cycles. Here, we investigate this possibility with a simulation experiment. The results show that our methods detect cycles with high probability when the cycles are sufficiently important.

We begin by choosing a stochastic process to simulate, which is capable of capturing both cyclical and acyclical behavior. We selected the stochastic cycle model of Harvey and Jaeger (1993), which takes the form

\[
\begin{align*}
    y_t &= \mu_t + \psi_t + \varepsilon_t \\
    \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t \\
    \beta_t &= \beta_{t-1} + \xi_t \\
    \psi_t &= \rho \cdot \cos(\lambda) \cdot \psi_{t-1} + \rho \cdot \sin(\lambda) \cdot \psi^*_t + \zeta_t \\
    \psi^*_t &= -\rho \cdot \sin(\lambda) \cdot \psi_{t-1} + \rho \cdot \cos(\lambda) \cdot \psi^*_{t-1} + \zeta^*_t,
\end{align*}
\]

where \(\{\varepsilon_t, \eta_t, \xi_t, \zeta_t, \zeta^*_t\}\) are all mean-zero Gaussian i.i.d. errors that are mutually uncorrelated at all leads and lags; \(y_t\) is the observed series (the underwriting loss ratio in our case), which is composed of a cyclical component \(\psi_t\), a noise term \(\varepsilon_t\), and a “trend” \(\mu_t\). The trend is sufficiently general to allow for highly nonstationary, I(2), processes, but also allow for a simple random walk or a stationary process as special cases.
The cycles are completely characterized by (1) \( \rho \), which controls the cycle damping, \( 0 \leq \rho \leq 1 \); (2) \( \lambda \), the frequency of the cycles (in radians per period, not cycles per period); and (3) \( \sigma_{\kappa} \), which controls the noise in the cycle. This allows for both regular cyclical behavior (with a period of \( 2\pi/\lambda \)) and random shocks to the cycle. High values of \( \rho \) imply persistent cyclical behavior whereas low values of \( \rho \) allow the random shocks to dominate the cyclical component.

The maximum likelihood estimation of the parameters of the Harvey–Jaeger model on the available loss ratio data yielded the values shown in the table below. We then chose three alternative parameterizations (Cases 1 to 3) of this process to reflect differing amounts of cyclical behavior. Case 1 should produce cycles of importance similar to that found in the loss ratio data. Cases 2 and 3 accentuate these cycles by making them much more persistent without changing their frequency. Increasing the persistence of these regular cyclical fluctuations should make them easier to distinguish from random shocks.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MLE</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho )</td>
<td>0.818</td>
<td>0.75</td>
<td>0.9</td>
<td>0.99</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.777</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>( \sigma_{\epsilon} )</td>
<td>0.0025</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>( \sigma_{\eta} )</td>
<td>0.0100</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( \sigma_{\zeta} )</td>
<td>0.224</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>( \sigma_{\kappa} )</td>
<td>2.64</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

For each of these three cases, we simulated 5,000 samples, each with the same number of observations as the original loss ratio data. For each of these 15,000 simulated samples, we

1. estimated the cyclical component using the linear trend (LT), quadratic trend (QT), Hodrick–Prescott (HP), and band-pass (BP) methods;
2. examined the ability of each of these 15,000 \( \times 4 = 60,000 \) estimated cyclical components to forecast both the loss ratio and the change in the loss ratio out of sample using the same rolling estimation technique described in the body of the article;
3. evaluated each of these \( 2 \times 60,000 = 120,000 \) series of forecasts using both the Diebold–Mariano and the modified Diebold–Mariano statistics described in the body of the article;
4. compared each of the \( 2 \times 120,000 = 240,000 \) test statistics to their 5 percent asymptotic critical values.

Before turning to results, one should note that none of the methods used to estimate the cyclical component in this experiment are ideal as none of them use the true model generating the data in order to estimate the cycle. This makes the experiment more realistic as it avoids the assumption that one knows the true process driving industry loss ratios. Also note that we use only the four simplest methods to estimate cycles in this experiment, as the computational costs of the other models made their
simulation cumbersome. To the extent that these other models could have greater power to detect cycles, our results may understate the power of the methods that we used.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Loss Ratio</th>
<th>( \rho = 0.75 )</th>
<th>( \rho = 0.90 )</th>
<th>( \rho = 0.99 )</th>
<th>\Delta Loss Ratio</th>
<th>( \rho = 0.75 )</th>
<th>( \rho = 0.90 )</th>
<th>( \rho = 0.99 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>LT</td>
<td>0.04</td>
<td>0.11</td>
<td>0.36</td>
<td></td>
<td>0.10</td>
<td>0.10</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>QT</td>
<td>0.05</td>
<td>0.20</td>
<td>0.36</td>
<td></td>
<td>0.05</td>
<td>0.04</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>HP</td>
<td>0.04</td>
<td>0.15</td>
<td>0.90</td>
<td></td>
<td>0.11</td>
<td>0.08</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>BP</td>
<td>0.11</td>
<td>0.29</td>
<td>0.57</td>
<td></td>
<td>0.04</td>
<td>0.06</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

The table above shows the results of our simulation experiments. The numbers shown are the fraction of outcomes in which we were able to reject the null hypothesis of no forecastable variation in the loss ratio (or change in the loss ratio) at the 5 percent significance level. All results shown are based on the modified Diebold–Mariano statistics. The Diebold–Mariano statistics were less conservative and gave higher numbers in every case.

Looking at the first column of results, we see that we are able to reject the null hypothesis of no forecastability at the 5 percent significance level only about 5 percent of the time. This is precisely what we would expect when the loss ratio is not forecastable. However, as we increase the persistence of the cyclical component (without changing its frequency) in columns 3 and 4, the frequency with which we reject the null hypothesis steadily increases. Depending on the detrending method used, the chances of detecting the cycle range from roughly 33 to 90 percent. The final three columns show the results for forecasts of the change in the loss ratio. Although the numbers are generally lower than before, when cycles are sufficiently persistent, the HP filter is still capable of detecting its presence in more than 80 percent of simulated trials.

We therefore conclude that, taken together, these forecasting tests have high power to detect cyclical behavior when the cycles are sufficiently persistent.

REFERENCES


