

# Forecasting the Size Effect

*It has been known for some time that the returns of small firms often differ from those of large firms, and that asset pricing theories, including the Capital Asset Pricing Model and Arbitrage Pricing Theory, cannot account for the difference. Small-capitalization stocks have provided higher average returns than large-capitalization stocks, and the outperformance has been strongest in the month of January. Various researchers have sought to explain this size effect as the result of differential transaction costs, liquidity, informational uncertainty and year-end tax-loss selling. Others have suggested that small stocks outperform because they tend to have lower P/E ratios.*

*A multifactor analysis "disentangles" the effect of firm size from related factors that may influence return. These factors include cross-sectional effects, such as firm neglect and low P/E, and calendar effects, such as tax-loss selling. Disentangling provides "pure" returns to size that avoid the confounding associated with proxy effects. For instance, disentangling reveals the January small-firm seasonal to be a mere surrogate for the rebound that follows the abatement of tax-loss selling.*

*An analysis of pure returns reveals that the size effect is buffeted by economic forces. There are times when small stocks outperform the market, and other times when they lag. But while the payoffs to the size effect are not at all regular to the naked eye, they are predictable in a broader empirical framework that incorporates macroeconomic drivers such as interest rates and industrial production. An examination of various forecasting models, including univariate and multivariate time-series techniques, indicates that one that imposes a Bayesian random-walk prior belief on the coefficients of a vector autoregressive model provides the best results.*

**S**OME EQUITY RETURN regularities, such as the return reversal effect, produce persistent payoffs; they represent anomalous pockets of market inefficiency. Others, such as the size effect, are predictable only in a broader macroeconomic framework; that is, they are "empirical return regularities" driven by macroeconomic forces.<sup>1</sup>

This article provides an in-depth look at the small-firm effect. We review the controversy surrounding this return effect and examine various methods for forecasting the size effect. Our

findings indicate that macroeconomic "drivers" are essential in predicting returns to size.

## The Size Effect

Gordon showed in 1962 that common stock returns are inversely related to a firm's size.<sup>2</sup> Banz later found that, over the 40 years ending 1975, smaller firms on the New York Stock Exchange (NYSE) had higher average returns than larger firms, when returns were adjusted for risk using the Capital Asset Pricing Model (CAPM). Surprisingly, the strength of the relation between risk-adjusted average return and size was comparable in magnitude to that between average return and systematic risk as measured by beta.<sup>3</sup>

Reinganum examined the small-firm effect across a broader universe including both NYSE

---

1. Footnotes appear at end of article.

Bruce Jacobs and Kenneth Levy are principals of Jacobs Levy Equity Management in Fairfield, New Jersey.

and American Stock Exchange (AMEX) firms.<sup>4</sup> He found superior risk-adjusted returns for small firms over the years 1963 to 1977. Brown, Kleidon and Marsh discerned that average risk-adjusted returns were linearly related to the logarithm of firm size. Moreover, they found the magnitude and sign of the relation between return and size to be unstable. Over the 1969–73 period larger companies outperformed, while smaller companies fared better over the 1974–79 period.<sup>5</sup>

Does the size effect simply reflect transaction costs? Is it risk mismeasurement or the product of deficiencies in asset pricing models? Is it a proxy for other return effects? We review the evidence below.

### Size and Transaction Costs

Blume and Stambaugh found that studies using daily returns tended to overstate the small-firm effect because of the “bid-ask” effect. Reinganum, for example, had compounded arithmetic average daily returns to estimate the size effect. This procedure replicates a portfolio strategy of daily rebalancing to equally weighted positions; stocks closing at the bid are generally purchased and those closing at the ask sold in order to reestablish equal weights. But assuming that purchases can be made at the bid and sales at the ask artificially inflates returns. This overstatement is greater for smaller firms, because they generally have lower prices, hence larger relative bid-ask spreads. Using one-year holding-period returns, Blume and Stambaugh found the size effect to be only half the magnitude of the effect estimated by Reinganum; furthermore, on average, the size effect was confined to one month—January.<sup>6</sup>

Stoll and Whaley assessed the impact of transaction costs on the Banz and Reinganum results, which were based on gross return. They examined NYSE securities and found the size effect eliminated for a three-month horizon, after controlling for the higher bid-ask spreads and brokerage commissions on small stocks. For holding periods of one year, abnormal returns were positive, but only weakly significant.<sup>7</sup>

Shultz extended the analysis to AMEX stocks. Using this broader universe, he found that smaller stocks exhibited significant risk-adjusted returns after transaction costs, even over short holding periods. Shultz also noted that transaction costs cannot explain the periodic sign reversal found by Brown, Kleidon and

Marsh, or the abnormal January behavior of smaller firms.<sup>8</sup>

Amihud and Mendelson hypothesized that investors demand compensation for illiquidity, and that the size effect proxies for an illiquidity premium. They employed the bid-ask spread as a measure of market thinness. This spread is inversely correlated with attributes that reflect liquidity, such as trading volume, number of shareholders, number of dealers making a market and degree of price continuity. They found that returns, both gross and net of trading costs, were an increasing function of the bid-ask spread; the effect of firm size was negligible after controlling for liquidity.<sup>9</sup>

Chiang and Venkatesh maintained that the higher spread for small firms is not due to illiquidity, but results rather from the higher proportion of “insiders” trading these stocks. The presence of such informed traders leads dealers to raise the spread, which in turn causes ordinary investors to require a higher expected return on small stocks.<sup>10</sup>

### Size and Risk Measurement

Roll proposed that a bias in beta estimates leads to an overstatement of the small-firm effect. Because small stocks trade less frequently than large stocks, their risk as estimated from daily returns understates their actual risk. Roll estimated the impact of non-synchronous trading on beta by using Dimson’s aggregated coefficients method. (Dimson betas are formed by regressing security returns on lagged, contemporaneous and leading market returns, then summing the three slope coefficients.) The use of Dimson betas moderated the observed small-firm effect, but Reinganum found it insufficient to explain the full magnitude of the effect.<sup>11</sup>

Hunda, Kothari and Wasley showed that even the betas of small-stock portfolios estimated from monthly returns will be downward biased. For all but the smallest companies, estimated betas approach their true values when quarterly returns are used. To minimize the beta measurement bias for the smallest firms, return intervals as long as one year are needed. Using quarterly returns to estimate beta, Hunda *et al.* found the small firm effect to be insignificant.<sup>12</sup>

Roll demonstrated that performance mismeasurement arises when the selected surrogate market portfolio, or benchmark, is not *ex ante* mean-variance efficient. Banz and Reinganum have acknowledged that their findings could be

due to benchmark error. But Banz used several different surrogates for the market portfolio and found the small-firm effect to be robust in every case. Booth and Smith have concluded that the small-firm effect cannot be explained by measurement error caused by benchmark error or non-synchronous trading. Using an errors-in-variables method, they demonstrated that the small-firm effect is robust over the feasible range of true coefficients.<sup>13</sup>

Mounting evidence supports the proposition that risk, hence expected return, varies over time.<sup>14</sup> Ferson, Kandel and Stambaugh examined the weekly returns on 10 portfolios of NYSE and AMEX securities ranked by firm size over the 1963–82 period. They found a single-premium, time-varying risk model capable of explaining the return differences across size-ranked portfolios.<sup>15</sup>

Chan and Chen used firm size in a novel fashion as an "instrument" to model changing risk premiums.<sup>16</sup> In the original size studies, stock betas were estimated from five years of monthly return data. This is a common choice for beta estimation, representing a compromise between a period long enough to ensure statistical accuracy and short enough to ensure stationarity. If daily or weekly data are used to obtain a greater number of observations, non-synchronous trading becomes an issue; as the length of the estimation period is increased, the stationarity of the data becomes questionable.

To overcome these problems, Chan and Chen formed portfolios based on size rankings and allowed the composition of these portfolios to change over time. They asserted that such portfolios maintain their risk characteristics over long spans of time. These size-ranked portfolios were used to estimate portfolio betas and to test if a size effect exists after controlling for beta. Estimating beta from over 30 years of return data, Chan and Chen found that the small-firm effect is subsumed.

Several studies have asked whether the size effect, or other market anomalies, can be accounted for by Ross' Arbitrage Pricing Theory (APT). Reinganum, controlling for APT risk using factor modeling, concluded that the size effect is not explained. Lehmann and Modest also found APT incapable of explaining the small-firm effect, even after adjusting for the January size seasonal and infrequent trading. Conner and Korajczyk found that APT appears to explain the January size seasonal but does not

explain the size effect in other months. However, Chen reported that size has no power to explain return in an APT framework.<sup>17</sup>

### Size and Risk Premiums

Barry and Brown have proposed that differential information across securities may account for the size effect. That is, there is more risk involved in estimating the valuation parameters for small firms, because there is less information available on small firms than on large ones. As a measure of information availability, Barry and Brown used the period of time a firm had been listed on an exchange. Analyzing beta, size and period-of-listing, as well as the interactions between these variables, they found a period-of-listing effect present for NYSE firms over the 1926–80 period. Unlike the firm-size effect, the period-of-listing effect had no January seasonal. Also, Barry and Brown found the interaction between size and period-of-listing to be more significant than the size effect itself.<sup>18</sup>

Merton developed a model of capital market equilibrium with incomplete information, where each investor has information about only some of the available securities.<sup>19</sup> This appears to be a sensible assumption even for institutional investors, because their "closely-followed" lists are often quite small compared with the universe of all listed securities. In Merton's model, the information available is the same for all stocks; that is, parameter estimation risk does not differ across securities. However, information about a particular stock is not available to *all* investors, but only to some.

Under these assumptions, Merton proved that expected returns will be higher, the smaller a stock's investor base, the larger firm-specific variances and the larger firm size. The positive association between expected return and large firm size appears contrary to the empirical evidence. But Merton compared small and large firms having identical investor bases and firm-specific variances. Smaller firms tend to have less investor recognition and larger specific variances than larger firms; Merton's finding is thus not necessarily inconsistent with the observed higher returns on smaller firms.

Barry and Brown's empirical evidence is consistent with Merton's theory to the extent that period-of-listing is positively associated with investor recognition. Arbel, Carvell and Strelbel's findings on neglected stocks (stocks not well followed by analysts and institutional in-

vestors) also support Merton.<sup>20</sup> Stratifying firms by risk, size and degree of institutional ownership, they found higher returns associated with less institutional following, even after controlling for firm size. They also concluded that the small-firm effect is subsumed by the neglect factor.

Shefrin and Statman have conjectured that the size anomaly proxies for a "responsibility" effect.<sup>21</sup> That is, advocacy of lower-quality stocks carries a higher degree of personal responsibility than advocacy of more reputable stocks. While no one ever second-guesses a recommendation to buy conventional names like IBM, recommending less well known stocks carries greater potential for regret. Low-reputation stocks should thus provide higher expected returns than high-reputation stocks. Shefrin and Statman also suggested that the period-of-listing, neglected-firm and low-P/E anomalies proxy for this responsibility effect.

#### Size and Other Cross-Sectional Effects

The small-firm effect may proxy for effects associated with other equity characteristics, such as low P/E. In 1977, Basu, using the CAPM to adjust for risk, demonstrated that low-P/E stocks on the NYSE provided higher average risk-adjusted returns than high-P/E stocks. Reinganum studied the low-P/E and size effects jointly to determine whether these anomalies are related. A two-way classification of NYSE and AMEX firms by company size and P/E ratios revealed the size effect to subsume the P/E effect. That is, after controlling for company size, there was no remaining P/E effect, while the size effect existed even after controlling for P/E. However, Reinganum did not adjust for risk.<sup>22</sup>

Basu examined NYSE firms over the 1963–80 period; using a randomized design and adjusting for risk, he reached conclusions contrary to those of Reinganum.<sup>23</sup> He found the P/E effect significant even after controlling for company size, but the size effect subsumed after controlling for differences in risk and P/E. The strength of the P/E effect appeared to vary inversely with firm size (stronger for smaller companies), however, so that an interaction may exist.

Cook and Rozeff reexamined the size-P/E controversy using an analysis of variance method.<sup>24</sup> They reported that size does not subsume P/E, nor does P/E subsume size; rather, size and P/E are independent effects.

They did not find an interaction between size and P/E, as claimed by Basu. Furthermore, they replicated Basu's method and found their results unaltered, thus concluding that Basu's findings must be specific to his sample.

Banz and Breen focused on the effect of database biases on the size-P/E connection.<sup>25</sup> They examined two separate databases—one bias-free and the other bias-prone—over the eight years from 1974 to 1981. The bias-free database was collected in real time from sequential Compustat tapes. The bias-prone database was represented by a current version of Compustat.

The Compustat database suffers from *ex post* selection bias, including survivorship bias (merged, bankrupt and liquidated companies are absent) and retrospective inclusion bias (newly included firms are entered with their prior history). It also suffers from look-ahead bias. The historical earnings reported for a given year end, for example, are not actually available until the following year. This bias tends to place companies that experience positive earnings surprises in low-P/E portfolios, and companies that experience negative earnings surprises in high-P/E portfolios. This tendency could magnify or even create a low-P/E effect.

Using the current Compustat database, Banz and Breen found evidence of statistically significant, independent P/E and size effects. Using the bias-free database, they found that the incremental returns accruing to low P/E were insignificant, while the size effect remained.

The small-firm effect may proxy for effects such as neglect, low price or high volatility. Jacobs and Levy fully "disentangled" the return effects associated with 25 different equity attributes, including firm size, in order to distinguish between "naive" and "pure" return effects.<sup>26</sup>

Naive returns to the size effect were calculated using monthly cross-sectional regressions of security returns on a normalized size attribute. Such univariate regressions naively measure the return effects associated with only one attribute at a time; no effort is made to control for related effects. In contrast, pure returns to the size attribute were calculated with monthly cross-sectional regressions of security returns on multiple attributes simultaneously. These multivariate regressions measure all effects jointly, thus purifying each effect so that it is independent of other effects.

Jacobs and Levy measured pure returns to

size and other effects for a bias-free universe of the 1500 largest-capitalization stocks over the 108-month period from January 1978 to December 1986. Noting the return effect to be linear in the logarithm of size, they used the logarithm of market capitalization as their measure of firm size. They found pure returns to size statistically significant, after controlling for all other attributes. They also found small size, low P/E and neglect to be independent.

### Size and Calendar Effects

The size effect appears to have significant interactions with calendar effects.<sup>27</sup> For example, the size effect has been related to the day-of-the-week effect. Among many researchers who have documented a stronger size effect on Fridays, Keim found that 63 per cent of the size effect occurs on that day.<sup>28</sup>

In 1976, Rozeff and Kinney showed that the stock market exhibits higher returns in the month of January.<sup>29</sup> Keim later showed this January return seasonal to be related to the size effect.<sup>30</sup> Analyzing NYSE and AMEX firms over the years 1963 to 1979, he found that one-half of the size effect occurs, on average, in January; moreover, one-quarter of the effect occurs during the first five trading days of the year.

Seasonalities in risk, in the release of information and in insider trading activity have been proposed to explain the January size effect. Rogalski and Tinic found that both the systematic and non-systematic risks of small stocks rise in January, but not by enough to account for the observed return pattern. Arbel cited year-end release of accounting information as a potential explanation of a neglect-driven size seasonal in January. But Chari, Jagannathan and Ofer found no excess return at fiscal year end for companies that do not have a December year end, casting doubt on the informational hypothesis. Seyhun found that small-firm insiders adjust their positions around year end, but that insider trading activity does not explain the January size seasonal.<sup>31</sup>

Roll and Reinganum have examined year-end tax-loss selling as a possible explanation for the small-firm January seasonal. This research was motivated by Branch and Dyl, who had earlier reported evidence of tax-loss selling consistent with the observed January effect. The tax hypothesis maintains that investors establish losses to shelter taxable income prior to the new

year. Once the year-end selling pressure abates around the turn of the year, prices rebound.<sup>32</sup>

Roll regressed securities' turn-of-the-year returns on their returns over the preceding year and found a significant negative relationship. Those securities experiencing negative returns in the preceding year were more likely to be subject to tax-loss selling and to bounce back early in the new year. Roll conjectured that the effect was largest for small firms because they are more likely candidates for tax-loss selling, given their higher volatility and lower relative representation in tax-exempt institutional portfolios. Also, the higher transaction costs for smaller firms inhibit arbitrage of the January seasonal.

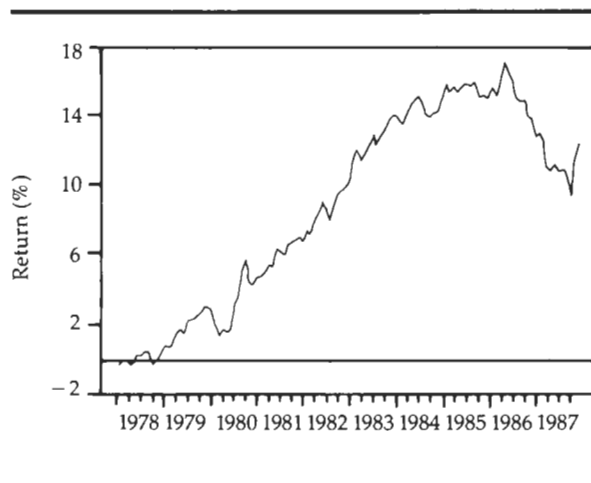
Reinganum constructed a measure of potential tax-loss selling by using the loss suffered from the high price in the previous period, classified for tax purposes as short-term. While he found the size seasonal early in January consistent with tax-loss selling, the entire January size effect could not be fully explained. Inconsistent with the tax-loss hypothesis was the finding that last year's winners outperformed in January.

Several studies provide an international perspective on the January-size connection. Brown, Keim, Kleidon and Marsh examined stock returns in Australia, where the tax year ends June 30th. They found that market returns exhibit July and also January seasonals; they also found a year-round size effect, but no size seasonals. They interpreted these results as inconsistent with the tax-loss hypothesis. However, Gultekin and Gultekin found seasonality in market returns consistent with a tax explanation in 13 countries (not including Australia).<sup>33</sup>

Tinic, Barone-Adesi and West have reported that, while taxes are not the sole explanation for the January size effect in Canada, the 1972 imposition of a capital gains tax did affect the behavior of returns. Also consistent with the tax-loss selling hypothesis, Shultz found no evidence of a U.S. January seasonal prior to the levy of personal income taxes in 1917. But Jones, Pearce and Wilson analyzed U.S. stock market returns extending back to 1871 and found evidence of a January seasonal prior to the advent of the income tax code.<sup>34</sup>

Constantinides' model of optimal tax trading implies that there should be no relation between the January seasonal and rational tax trading; the optimal time to recognize losses should be

**Figure A** Cumulative Return to Small Size



just prior to their becoming long-term, not year end. Chan has noted that there is little reason to realize a long-term loss near year end, presumably at depressed prices. Employing two measures of potential short-term and long-term tax-loss selling, he found the January seasonal to be as strongly related to long-term as to short-term losses. Chan concluded that the January seasonal is not explained by optimal tax trading.<sup>35</sup>

Jacobs and Levy examined the January size seasonal after controlling for potential short-term and long-term tax-loss selling pressure, as well as the return effects associated with 22 other equity attributes.<sup>36</sup> While they corroborated the existence of a January size seasonal for returns measured in naive form, they found it fully dissipated in pure form. This evidence is consistent with the size seasonal being a mere proxy for tax-related selling.

Furthermore, Jacobs and Levy found the long-term tax-loss measure to have a rebound about twice the magnitude of the short-term measure. While this result is inconsistent with rational tax-trading strategies, it may be explained by investor psychology. Shefrin and Statman have noted that investors tend to ride losers too long in an effort to break even.<sup>37</sup> Thus tax-loss selling is stronger for stocks suffering long-term losses.

### Modeling the Size Effect

Figure A graphs cumulative pure returns to size, as measured by Jacobs and Levy.<sup>38</sup> Below, we develop some models for forecasting these returns. First, however, we discuss the criteria

we used to assess the accuracy of alternative forecasting methods.

A commonly used measure in portfolio management is the information coefficient (IC), defined as the correlation between forecast and actual returns. One drawback of the IC as a forecast evaluation tool is its independence of both origin and scale. The measured correlation between forecast and actual returns is unaffected by adding a constant to each forecast or multiplying each forecast by a positive constant. But while IC is invariant to these transformations, the forecast errors—the difference between forecast and actual returns—clearly are not. Thus the IC cannot differentiate between a perfectly accurate set of forecasts and an alternative set that consistently overestimates by 50 per cent.

The criteria we use assess the accuracy of alternative forecasting methods more directly than the IC. They are defined in the appendix. The first criterion is mean error (ME), which is a simple average of forecast errors. If the ME is positive (negative), the model tends to underestimate (overestimate) actual returns. The ME is likely to be small, because positive and negative errors tend to offset each other. The mean absolute error (MAE), which is a simple average of the absolute value of forecast errors, avoids this problem. The MAE is an appropriate criterion if the cost of erring is proportionate to the size of the forecast error.

Root mean squared error (RMSE) is the square root of a simple average of squared forecast errors. The RMSE is consistent with a squared loss function, where the pain of erring grows with the square of the forecast error. Simply stated, the RMSE counts a larger error more heavily than a collection of smaller errors having similar aggregate magnitude. The RMSE is larger than the MAE unless all errors are of the same size, in which case the two measures are identical.

The three criteria reviewed thus far are all absolute measures. They do not facilitate comparisons across different variables, or across different time periods. Relative measures can be constructed on the basis of percentage error. Still more useful is a criterion that measures gains in accuracy compared with a naive benchmark forecast. The Theil U statistic compares the RMSE of the forecast model with that of a naive forecast using last period's actual return.<sup>39</sup>

This naive benchmark forecast explicitly assumes no change in the next period.

A Theil U less than (in excess of) one indicates that the forecast method is better (worse) than a naive forecast. A Theil U less than one should be interpreted with caution, however, because a no-change forecast may be a bad benchmark. For example, the level of the Consumer Price Index (CPI) generally rises over time; a forecast of the CPI will thus be more accurate if it adds a trend component, rather than just assuming that last period's level will continue unchanged. In this instance, a no-change benchmark is a poor one, and a Theil U less than one is relatively easy to achieve.

We apply these forecast accuracy criteria "out of sample." That is, our models are estimated over a portion of the historical time series and their forecast accuracies tested over a more recent "hold-out" sample. This is fundamentally different from "in-sample" fitting. For example, regression analysis locates the line of best fit to historical data. In ordinary-least-squares (OLS) regression, squared errors are minimized. The focus is on "goodness-of-fit" measures such as R, the correlation coefficient between actual and fitted values or, more commonly, R-squared, the proportion of the variance in the variable of interest explained by the regression.

But obtaining a good fit to historical data does not necessarily imply a good forecast model. In fact, a perfect fit can be obtained by using a polynomial of sufficiently high order. A polynomial of degree  $n-1$ , for example can be specified to pass through any  $n$  data points. The resulting model will match the data exactly, but it will tell us nothing about the process generating the data. Such a model has little or no forecasting power, even though the fit is perfect.

Out-of-sample tests, by contrast, focus on the predictive ability of a model. Rather than fitting a historical series, the goal is to improve the model's predictive ability. A particularly useful out-of-sample measure is a t-statistic on the economic insight provided by the model forecasts.

We use these criteria to compare the forecast models discussed below.

### Simple Extrapolation Techniques

A simple extrapolation technique may use averages or moving averages to forecast. Such an approach is generally regarded as "determin-

istic," because it ignores the underlying randomness in the series. As a result, expected forecast accuracy cannot be measured. Nevertheless, these simple approaches sometimes work well.

We looked at a simple average of past observations. This approach assumes that the process generating the data is in equilibrium around a constant value—the underlying mean. The process is subject to random error, or "noise," which causes the observed returns to be scattered about the mean.

First we calculated the average monthly pure return to size over the base period January 1978 to December 1981, then used this average value as the return forecast for January 1982. In each successive month, the average return was updated by adding one additional observation, and the revised average became the return forecast for the next month. The last forecast month was December 1987. Table I gives statistics on the accuracy of forecasts for the next one to six months.

The simple average tended to overestimate actual returns slightly, as indicated by the MEs, which averaged  $-0.09$  per cent across forecast steps. The MAEs ranged from 0.43 to 0.44 per cent, while the RMSEs ranged from 0.55 to 0.57 per cent. The "goodness" of these absolute statistics can be assessed only by comparison with other forecast models.

The Theil Us indicate that the RMSE ranged from 73 to 87 per cent of the RMSE associated with the naive benchmark, which was simply last month's return. Thus a simple historical average provided a much better return forecast than a no-change benchmark. The Theil values associated with the simple average became our standard for comparing the forecasting models.

Simple averaging equally weights all observations; with exponential smoothing weights decay in an exponential manner, so that recent observations are weighted more heavily than older ones.<sup>40</sup> The rate of decay is determined by a parameter that ranges from zero to one. A parameter close to one places most of the weight on recent observations, while a parameter close to zero distributes the weight more evenly across all observations. While exponential smoothing adapts faster than simple averaging to recent changes, it will necessarily trail any trend in the data.

The second panel of Table I displays forecast statistics for exponential smoothing based on a

**Table I** Forecast Statistics

| Constant Model |        |         |          |         |     |
|----------------|--------|---------|----------|---------|-----|
| Forecast Step  | ME (%) | MAE (%) | RMSE (%) | Theil U | T   |
| 1              | -0.08  | 0.43    | 0.55     | 0.74    | 1.3 |
| 2              | -0.08  | 0.43    | 0.56     | 0.73    | 1.3 |
| 3              | -0.09  | 0.43    | 0.56     | 0.78    | 1.2 |
| 4              | -0.08  | 0.43    | 0.56     | 0.77    | 1.2 |
| 5              | -0.09  | 0.43    | 0.56     | 0.87    | 1.1 |
| 6              | -0.10  | 0.44    | 0.57     | 0.80    | 0.9 |

| Exponential Smoothing Model |        |         |          |         |      |
|-----------------------------|--------|---------|----------|---------|------|
| Forecast Step               | ME (%) | MAE (%) | RMSE (%) | Theil U | T    |
| 1                           | 0.07   | 0.45    | 0.55     | 0.74    | 0.8  |
| 2                           | 0.07   | 0.45    | 0.56     | 0.73    | 0.9  |
| 3                           | 0.07   | 0.45    | 0.56     | 0.78    | 1.3  |
| 4                           | 0.07   | 0.45    | 0.56     | 0.77    | 1.6  |
| 5                           | 0.06   | 0.45    | 0.56     | 0.87    | 1.8* |
| 6                           | 0.05   | 0.45    | 0.56     | 0.79    | 0.7  |

| Vector Time-Series Model |        |         |          |         |      |
|--------------------------|--------|---------|----------|---------|------|
| Forecast Step            | ME (%) | MAE (%) | RMSE (%) | Theil U | T    |
| 1                        | -0.18  | 0.51    | 0.67     | 0.90    | 0.5  |
| 2                        | 0.00   | 0.48    | 0.63     | 0.83    | 1.0  |
| 3                        | 0.02   | 0.56    | 0.69     | 0.96    | -1.1 |
| 4                        | 0.89   | 0.96    | 1.15     | 1.58    | -0.3 |
| 5                        | -1.16  | 1.22    | 1.42     | 2.20    | 0.0  |
| 6                        | 0.41   | 0.58    | 0.77     | 1.09    | 0.6  |

| Bayesian Model |        |         |          |         |      |
|----------------|--------|---------|----------|---------|------|
| Forecast Step  | ME (%) | MAE (%) | RMSE (%) | Theil U | T    |
| 1              | -0.08  | 0.41    | 0.53     | 0.71    | 3.1* |
| 2              | -0.06  | 0.42    | 0.54     | 0.71    | 3.0* |
| 3              | -0.07  | 0.42    | 0.54     | 0.76    | 2.8* |
| 4              | -0.07  | 0.41    | 0.55     | 0.76    | 2.5* |
| 5              | -0.09  | 0.41    | 0.56     | 0.86    | 2.3* |
| 6              | -0.10  | 0.42    | 0.57     | 0.80    | 2.0* |

\* Significant at the 5 per cent level.

parameter of 0.3. The results are generally similar to those for the constant process. The biggest difference is that the MEs are on average slightly positive, indicating a tendency to underestimate actual returns. Also, the MAEs are somewhat higher. (The forecast statistics obtained from an array of decay parameter settings were similar.)

### Time-Series Techniques

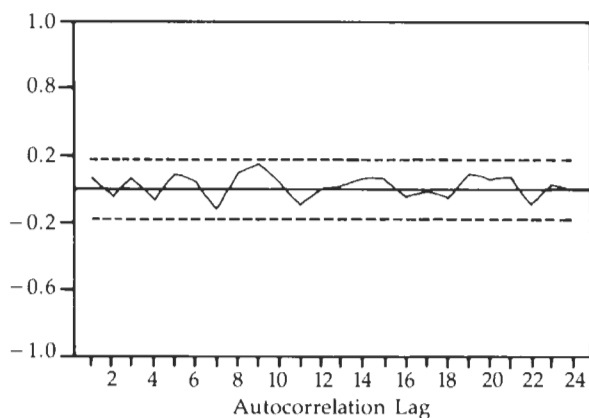
The objective of time-series analysis is to identify and model patterns in the historical data. This assumes that a time series is generated by a "stochastic," or random, process which can be described and replicated by a model. Time-series approaches include autoregressive models, which depend on a weighted sum of past

values, and moving-average models, which depend on a weighted sum of past errors. Some stochastic processes exhibit both autoregressive and moving-average characteristics, and can be modeled with mixed autoregressive/moving-average processes.<sup>41</sup>

Autocorrelations are used to measure the "memory" in the time series—that is, whether past values can predict future values.<sup>42</sup> Morgan and Morgan studied autocorrelation patterns in the returns of small-firm portfolios and found positive autocorrelation of monthly returns, six and 12 months apart, which was *not* specific to January. Lo and MacKinley found positive autocorrelation for weekly and monthly returns, independent of the effects of infrequent trading. Levis found significant autocorrelation of quar-



**Figure B** Autocorrelations of Return



terly returns on the London Stock Exchange.<sup>43</sup> But these autocorrelation patterns are for naive returns to size, hence may incorporate related return effects.

Figure B, a correlogram, shows the autocorrelation pattern for pure returns to small size. The corridor defined by the dotted lines represents a 95 per cent confidence band.<sup>44</sup> As all the autocorrelation coefficients lie within this confidence band, none is statistically significant at the 5 per cent level.

The hypothesis that all the autocorrelations jointly are insignificant can be tested using the Portmanteau Q statistic.<sup>45</sup> The Q statistic is not significantly different from zero. Pure returns to small size, at least over this time period, are indistinguishable from "white noise." That is, the time-series of returns does not differ from that of a random variable independently distributed across time.<sup>46</sup>

### Transfer Functions

The analysis of any single series excludes information that may be contained in related series. Regression analysis allows one to ascertain cause-and-effect relations between one or more independent variables and the dependent variable being forecast. Regression methods are causal, or explanatory, in nature.<sup>47</sup>

Transfer functions blend time-series analysis with explanatory variables.<sup>48</sup> The dependent variable being forecast is related to a weighted sum of lagged values of itself, current and lagged values of one or more independent variables, and an error term. The error term is

modeled with time-series techniques. Potentially useful explanatory variables may be chosen on the basis of economic theory, but the precise form the relationship takes will depend on the data, including the autocorrelation function of each series and contemporaneous and lagged cross-correlations between series.

The transfer function approach has been applied to forecasting stock market returns, using the Composite Index of Leading Indicators as an explanatory variable.<sup>49</sup> Levis used transfer functions to examine the relation between the size effect in the U.K. and institutional trading patterns. He found the size effect on the London Stock Exchange unrelated to institutional acquisitions or dispositions.<sup>50</sup> On the contrary, he found institutional trading follows, rather than leads, market behavior.

Transfer functions can only model one-way causality. Consider, for example, a time-series, Y, to be forecast, and a related series, X. Information contained in time-series X may be useful for forecasting series Y. X is said to "cause" Y if using past values of X, in addition to past values of Y, yields a more accurate prediction of Y than past values of Y alone.<sup>51</sup>

This is a useful concept, which lends itself to statistical testing, but it does not necessarily correspond to the commonsense meaning of causality. Causality between time-series can run in both directions: X can cause Y, and Y can cause X. In this instance, X and Y cause each other. This is referred to as "feedback." When series exhibit feedback, transfer-function modeling is not appropriate; vector time-series models should be considered.

### Vector Time-Series

Vector time-series approaches model a group, or vector, of related variables.<sup>52</sup> In vector autoregression (VAR), each variable is regressed on its own historical values as well as on past values of other explanatory variables. Such joint modeling of time series permits an understanding of the dynamic relationships among the series: Series may be contemporaneously related, some may lead others, or there may be feedback present. By incorporating the information contained in the multiple time series, the accuracy of forecasts can be improved.

The differential returns to smaller firms may derive from their greater sensitivity to certain pervasive, economy-wide factors. Chan, Chen and Hsieh found, for example, that the return

differential between corporate and government bonds explains much of the cross-sectional variation in return to firm size.<sup>53</sup> Small firms fluctuate more with the business cycle than large firms. Their greater sensitivity to economic conditions arises from their more marginal nature.

Chan and Chen found that small firms have often operated inefficiently, have recently decreased in size, have higher financial leverage and have less access to external financing than larger firms.<sup>54</sup> Consistent with their higher risk, small firms tend to perform well when the default spread is narrowing and risk-aversion is abating. The greater sensitivity of smaller firms to default spreads is a risk for which investors may demand compensation.

The perspective of Chan *et al.* derives from equilibrium pricing theory. Our perspective is quite different. We want to know whether differential returns to smaller firms are forecastable. Rather than examining contemporaneous pricing relationships, we want to predict future returns by employing macroeconomic drivers in a time-series framework.

We constructed a monthly VAR model of the size effect using the following set of economic measures as explanatory variables—(1) low-quality (BAA) corporate bond rate, (2) long-term (10-year) Treasury bond rate, (3) Treasury bill (90-day) rate, (4) S&P 500 total return, (5) Industrial Production Index (logarithmic) and (6) Consumer Price Index (logarithmic).<sup>55</sup> These macro drivers were chosen because of their importance in security valuation. (Of course, considerations other than value may be helpful in modeling the size effect.<sup>56</sup>)

We used three autoregressive, or lag, terms for each of the six macroeconomic variables.<sup>57</sup> The model was first estimated over a base period from January 1978 to December 1981, then revised monthly through the end of 1987 using a Kalman filter update.<sup>58</sup> The third panel of Table I displays the forecast statistics.

A glance at the statistics reveals this model to be inferior to the constant model. The MEs, MAEs and RMSEs are substantially larger, and the differences become especially pronounced at forecast steps four, five and six. The Theil U exceeds one for the last three forecast steps, indicating that the forecasts are less accurate than no-change forecasts. The t-statistics are insignificant and sometimes negative.

The poor forecasting power of the model results from its overparameterization—a typical

problem with VAR models. While VAR models have the virtue of allowing the data to “speak for themselves,” there are rarely enough historical data available to allow the modeling of more than a few related series, because the number of coefficients to be estimated grows with the square of the number of variables. As a consequence, an unrestricted VAR model tends to become overparameterized; that is, it lacks sufficient data relative to the number of coefficients to be estimated.

Because it has a large number of coefficients available to explain a small number of observations, a VAR model can explain historical data well. But it is likely to “overfit” the data. That is, it will fit not only systematic, or stable, relationships, but also random, or merely circumstantial, ones. The latter are of no use in forecasting, and may be misleading. Thus, while the model provides a good in-sample fit, it is likely to forecast poorly.<sup>59</sup>

Some improvement in forecasting ability may be gained by aiming for a more parsimonious, or simpler, representation of the return-generating process. This can be done by introducing moving-average (MA) terms, which should improve the efficiency of the parameter estimation process without assuming away important interactions among the variables. The disadvantage of introducing MA terms is that the identification of the order of the moving-average and autoregressive lag lengths is difficult, particularly in multivariate applications. Moreover, such vector autoregressive/moving-average (VARMA) models cannot cope with as many explanatory variables as we consider here.<sup>60</sup>

There are two traditional ways of reducing the dimensionality of vector models. One is simply to use univariate time-series methods. While this approach dramatically reduces dimensionality by excluding all cross effects, it is severely restrictive because interactions among variables are assumed to be nonexistent. The other approach is structural modeling, which relies on restrictions suggested by economic theory. In fact, there is no evidence to date that VARMA modeling can perform as well as commercially available structural models of the economy.<sup>61</sup>

### Structural Macroeconomic Models

The overparameterization of vector models has traditionally been resolved by incorporating economic theory. Such “econometric” models

include only those variables and lags suggested by theory. As a result, the models require substantially fewer variables and lags than an unrestricted VAR.<sup>62</sup>

Econometric models are referred to as structural models because they explicitly incorporate theories about economic structure. There are, needless to say, many conflicting schools of economic thought. Although each theory undoubtedly contains some elements of truth, none is fully descriptive of reality. Thus, while structural models may avoid the overfitting problem of VAR, they also incorporate rigid beliefs, some of which may be unfounded. Furthermore, because theory forces the exclusion of a large number of variables at the outset, these variables never have the opportunity to refute the theory, no matter how strong the evidence of the data.

Structural models use prior beliefs based on economic theory to impose restrictions, but these restrictions are often too severe. As a result, the modeler's confidence in the theory may be overstated.<sup>63</sup> Unrestricted VAR models, by contrast, may understate the modeler's knowledge, because the data alone determine the values of the coefficients. An alternative approach is to represent statistically the modeler's uncertainty regarding the merits of alternative theories and to allow the data to revise the theory. Bayesian methods are designed to accommodate such uncertainty.<sup>64</sup>

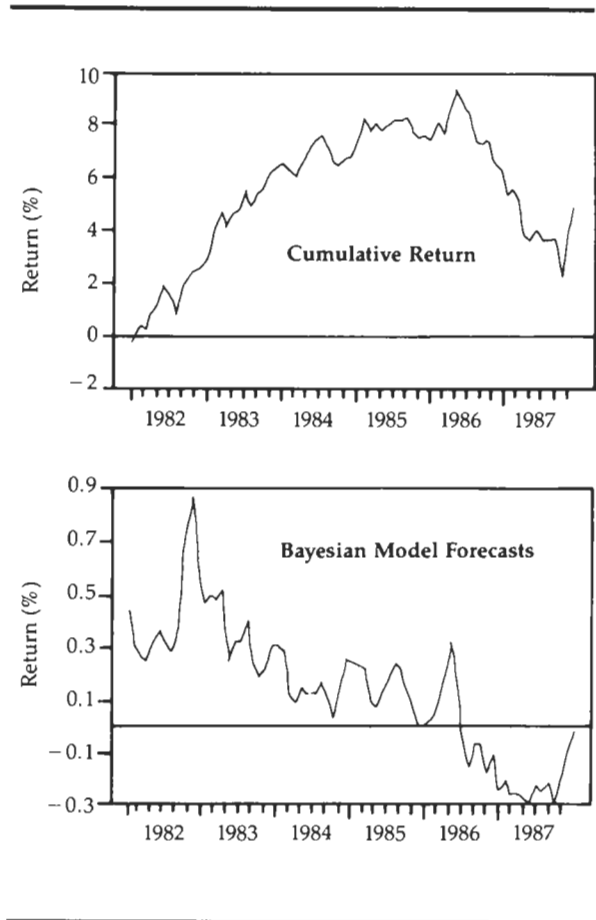
### Bayesian Vector Time-Series

Many economic measures are difficult to predict, but their behavior can often be approximated by a random walk. A random-walk model for interest rates assumes it is equally likely that rates will rise or fall. A random-walk forecast of next month's rate would simply be this month's rate of interest.

There are many ways to specify prior beliefs about the coefficients of a forecast model. One Bayesian specification imposes a random-walk prior on the coefficients of a VAR model. This results in a powerful forecasting tool that provides a viable alternative to structural econometric modeling. In fact, it has been demonstrated that this approach "can produce economic forecasts that are at least competitive with the best forecasts commercially available."<sup>65</sup> Further, there is evidence that these methods can forecast economic turning points.<sup>66</sup>

The difficulty of predicting stock returns is no

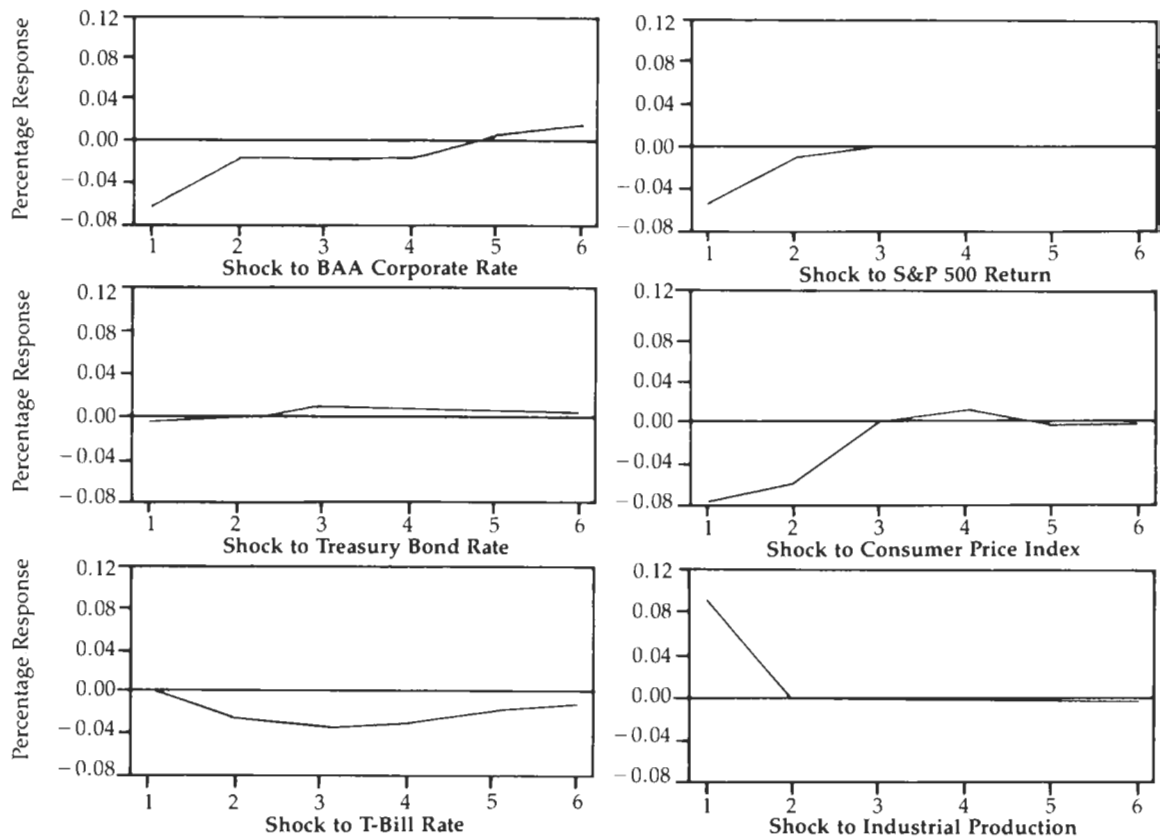
Figure C Forecasting Returns to Small Size



secret. But stock prices, like other economic data, can be approximated by a random walk. As early as 1900, Bachelier articulated a theory of random walks in security prices.<sup>67</sup> A random-walk model implies that successive price changes are independent draws from the same probability distribution. That is, the series of price changes has no memory and appears unpredictable. Using a random-walk prior to model security returns is thus eminently sensible.<sup>68</sup>

We modeled the size effect using a Bayesian random-walk prior and the same six macroeconomic drivers discussed earlier. The top part of Figure C displays cumulative returns to small size for the period January 1982 to December 1987. The lower part displays "out-of-sample" return forecasts for one month ahead. The forecasts for small stocks were positive during the early years; they gradually declined and turned negative during the last two years, implying

**Figure D** Forecast Response of Small Size to Macroeconomic Shocks



that small stocks were expected to underperform large stocks.

The last panel of Table I displays forecast statistics for the Bayesian model. The MAEs, RMSEs and Theil Us are substantially better than those of the unrestricted VAR. These statistics also show an improvement over the constant model. For instance, the Theil U for one-month forecasts is 0.71 for the Bayesian model and 0.74 for the constant model, a relative improvement of about 5 per cent. The margin of improvement declines with the forecast horizon. Similar results hold for the MAE and RMSE. The t-statistics are significant for all six forecast steps, and decline gradually from 3.1 at step one to 2.0 at step six. It is highly unlikely that the economic insight associated with this approach occurred as a result of chance.

We used an "impulse" analysis to estimate the impact of each macro driver on forecast returns. To consider the effect of an unexpected increase in the corporate bond rate on forecast

returns to size, for example, we defined the magnitude of the rate increase to be one standard deviation, or one unit, of historical BAA interest rate volatility and applied this one-unit BAA rate "shock" to the model.<sup>69</sup>

Figure D graphs the forecast return response of small size to one-unit shocks in each of the six macro drivers. A shock in the BAA rate produces forecast changes of about -6 basis points one month ahead, -2 basis points two, three and four months ahead, and +1 basis point five and six months ahead. The response to a shock in long government rates is negligible. The response to a Treasury bill shock is negative, ranging between 0 and -4 basis points. The response to a shock in the S&P 500 is -5 basis points one month ahead and negligible thereafter. A shock in the inflation rate lowers the forecast return by 8 basis points one month ahead and 6 basis points two months out. A shock in industrial production raises the forecast return by 9 basis points one month ahead.

The negative impact on returns to small size of an unexpected increase in BAA corporate interest rates and the positive impact of an unexpected increase in industrial production are consistent with the fragility of smaller firms. The negative response to an unexpectedly positive S&P 500 return, however, suggests that more is at work than risk considerations alone. The negative response to an increase in Treasury bill rates is consistent with the greater capital constraints on smaller firms. Also, smaller stocks fare less well during periods of unexpected inflation. ■

## Appendix

The following criteria were used to assess the accuracy of the various forecasting methods: mean error (ME), defined as

$$\left[ \sum_{t=1}^n (A_t - F_t) \right] / n;$$

mean absolute error (MAE), defined as

$$\left[ \sum_{t=1}^n |A_t - F_t| \right] / n;$$

root mean squared error (RMSE), defined as

$$\sqrt{\left[ \sum_{t=1}^n (A_t - F_t)^2 \right] / n}; \text{ and}$$

Theil U, defined as

$$\frac{\sqrt{\left[ \sum_{t=0}^{n-1} (A_{t+1} - F_{t+1})^2 \right] / n}}{\sqrt{\left[ \sum_{t=0}^{n-1} (A_{t+1} - A_t)^2 \right] / n}}.$$

In the above,  $A_t$  equals the actual value at time  $t$ ,  $F_t$  equals the forecast value for time  $t$ , and  $n$  equals the number of observations.

## Footnotes

1. See B. Jacobs and K. Levy, "Disentangling Equity Return Regularities: New Insights and Investment Opportunities," *Financial Analysts Journal*, May/June 1988.
2. M. Gordon, *The Investment, Financing and Valuation of the Corporation* (Homewood, IL: Richard D. Irwin, 1962), pp. 81-82, 195.
3. R. Banz, "The Relationship Between Return and Market Value of Common Stocks," *Journal of*

- Financial Economics* 9 (1981), pp. 3-18. See G. Schwert, "Size and Stock Returns, and Other Empirical Regularities," *Journal of Financial Economics* 12 (1983), pp. 3-12, for a review of the early size-effect literature.
4. M. Reinganum, "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values," *Journal of Financial Economics* 9 (1981), pp. 19-46. Also see M. Reinganum, "Abnormal Returns in Small Firm Portfolios," *Financial Analysts Journal*, March/April 1981, and M. Reinganum, "Portfolio Strategies Based on Market Capitalization," *Journal of Portfolio Management*, Winter 1983.
5. P. Brown, A. Kleidon and T. Marsh, "New Evidence on the Nature of Size-Related Anomalies in Stock Prices," *Journal of Financial Economics* 12 (1983), pp. 33-56.
6. M. Blume and R. Stambaugh, "Biases in Computed Returns," *Journal of Financial Economics* 12 (1983), pp. 387-404; R. Roll, "On Computing Mean Returns and the Small Firm Premium," *Journal of Financial Economics* 12 (1983), pp. 371-386, reports similar results.
7. H. Stoll and R. Whaley, "Transaction Costs and the Small Firm Effect," *Journal of Financial Economics* 12 (1983), pp. 57-79.
8. P. Shultz, "Transaction Costs and the Small Firm Effect: A Comment," *Journal of Financial Economics* 12 (1983), pp. 81-88.
9. Y. Amihud and H. Mendelson, "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics* 17 (1986), pp. 223-249, and Amihud and Mendelson, "Liquidity and Stock Returns," *Financial Analysts Journal*, May/June 1986.
10. R. Chiang and P. Venkatesh, "Insider Holdings and Perceptions of Information Asymmetry: A Note," *Journal of Finance*, September 1988.
11. R. Roll, "A Possible Explanation of the Small Firm Effect," *Journal of Finance*, September 1981; E. Dimson, "Risk Measurement When Shares are Subject to Infrequent Trading," *Journal of Financial Economics* 7 (1979), pp. 197-226; and M. Reinganum, "A Direct Test of Roll's Conjecture on the Firm Size Effect," *Journal of Finance*, March 1982.
12. P. Hunda, S. Kothari and C. Wasley, "Bias in Estimation of Systematic Risk and Its Implications for Tests of the CAPM" (NYU working paper #404, January 1987).
13. R. Roll, "A Critique of the Capital Asset Pricing Theory's Test: Part I," *Journal of Financial Economics* 4 (1977), pp. 120-176; Roll, "Performance Evaluation and Benchmark Errors I/II," *Journal of Portfolio Management*, Summer 1980/Fall 1981; and J. Booth and R. Smith, "The Application of Errors-in-Variables Methodology to Capital Market Research: Evidence on the Small-Firm Effect,"

- Journal of Financial and Quantitative Analysis*, December 1985.
14. M. Rozeff, "Dividend Yields and Equity Risk Premiums," *Journal of Portfolio Management*, Fall 1984; D. Keim and R. Stambaugh, "Predicting Returns in the Bond and Stock Markets," *Journal of Financial Economics*, December 1986; E. Fama and K. French, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy* 96-2 (1988), pp. 246-273; and Fama and French, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, October 1988.
  15. W. Ferson, S. Kandel and R. Stambaugh, "Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas," *Journal of Finance*, June 1987.
  16. K. Chan and N. Chen, "An Unconditional Asset Pricing Test and the Role of Firm Size as an Instrumental Variable for Risk," *Journal of Finance*, June 1988.
  17. S. Ross, "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory* 13 (1976), pp. 341-360; M. Reinganum, "The Arbitrage Pricing Theory: Some Empirical Results," *Journal of Finance*, May 1981; B. Lehmann and D. Modest, "The Empirical Foundations of the Arbitrage Pricing Theory," *Journal of Financial Economics*, September 1988; G. Connor and R. Korajczyk, "Risk and Return in an Equilibrium APT," *Journal of Financial Economics*, September 1988; and N. Chen, "Some Empirical Tests of the Theory of Arbitrage Pricing," *Journal of Finance*, December 1983.
  18. C. Barry and S. Brown, "Differential Information and the Small Firm Effect," *Journal of Financial Economics*, July 1984, "Differential Information and Security Market Equilibrium," *Journal of Financial and Quantitative Analysis*, December 1985, and "Limited Information as a Source of Risk," *Journal of Portfolio Management*, Winter 1986.
  19. R. Merton, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, July 1987. For some empirical results, see Y. Amihud and H. Mendelson, "The Effects of Beta, Bid-Ask Spread, Residual Risk and Size on Stock Returns" (New York U. working paper, 1988), and R. Gilmer and H. Swanson, "An Empirical Test of Merton's Incomplete Information Model" (Mississippi State U. working paper, September 1988).
  20. A. Arbel, S. Carvell and P. Strebel, "Giraffes, Institutions and Neglected Firms," *Financial Analysts Journal*, May/June 1983; A. Arbel, "Generic Stocks: An Old Product in a New Package," *Journal of Portfolio Management*, Summer 1985; and S. Carvell and P. Strebel, "Is There A Neglected Firm Effect?" *Journal of Business Finance & Accounting*, Summer 1987.
  21. H. Shefrin and M. Statman, "A Behavioral Finance Solution to the Noise Trading Puzzle" (Santa Clara U. working paper, December 1987).
  22. S. Basu, "Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, June 1977; and M. Reinganum, "Misspecification of Capital Asset Pricing," *op. cit.*
  23. S. Basu, "The Relationship Between Earnings' Yield, Market Value and Return for NYSE Common Stocks: Further Evidence," *Journal of Financial Economics* 12 (1983), pp. 129-156.
  24. T. Cook and M. Rozeff, "Size and Earnings/Price Ratio Anomalies: One Effect or Two?" *Journal of Financial and Quantitative Analysis*, December 1984.
  25. R. Banz and W. Breen, "Sample-Dependent Results Using Accounting and Market Data: Some Evidence," *Journal of Finance*, September 1986.
  26. See Table I in Jacobs and Levy, "Disentangling Equity Return Regularities," *op. cit.*, for references to earlier studies. The 25 measures used were low P/E, small size, yield, zero yield, neglect, low price, book/price, sales/price, cash/price, sigma, beta, coskewness, controversy, three measures of trends in analysts' earnings estimates, three measures of earnings surprise, earnings torpedo, relative strength, two measures of return reversal, and two measures of potential tax-loss selling. Also, 38 industry measures were utilized to purify returns further.
  27. For a synthesis of the calendar literature, see B. Jacobs and K. Levy, "Calendar Anomalies," *Financial Analysts Journal*, November/December 1988.
  28. M. Gibbons and P. Hess, "Day of the Week Effects and Asset Returns," *Journal of Business*, October 1981; D. Keim and R. Stambaugh, "A Further Investigation of the Weekend Effect in Stock Returns," *Journal of Finance*, July 1984; L. Harris, "A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns," *Journal of Financial Economics* 16 (1986), pp. 99-117; and D. Keim, "Daily Returns and Size-Related Premiums: One More Time," *Journal of Portfolio Management*, Winter 1987. However, E. Miller ("Why a Weekend Effect?" *Journal of Portfolio Management*, Summer 1988) argues that intraweek patterns in returns to size are unrelated to the long-run size effect.
  29. M. Rozeff and W. Kinney, "Capital Market Seasonality: The Case of Stock Returns," *Journal of Financial Economics* 3 (1976), pp. 379-402.
  30. D. Keim, "Size-Related Anomalies and Stock

- Return Seasonality," *Journal of Financial Economics* 12 (1983), pp. 13–32.
31. R. Rogalski and S. Tinic, "The January Size Effect: Anomaly or Risk Mismeasurement?" *Financial Analysts Journal*, November/December 1986; Arbel, "Generic Stocks," *op. cit.*; V. Chari, R. Jagannathan and A. Ofer, "Fiscal Year End and the January Effect" (Kellogg Graduate School of Management working paper #20, July 1986); and H. Seyhun, "The January Effect and Aggregate Insider Trading," *Journal of Finance*, March 1988.
  32. R. Roll, "Vas Ist Das? The Turn-of-the-Year Effect and the Return Premia of Small Firms," *Journal of Portfolio Management*, Winter 1983; M. Reinganum, "The Anomalous Stock Market Behavior of Small Firms in January," *Journal of Financial Economics* 12 (1983), pp. 89–104; B. Branch, "A Tax Loss Trading Rule," *Journal of Business*, April 1977; and E. Dyl, "Capital Gains Taxation and Year-End Stock Market Behavior," *Journal of Finance*, March 1977.
  33. P. Brown, D. Keim, A. Kleidon and T. Marsh, "Stock Return Seasonality and the Tax-Loss Selling Hypothesis," *Journal of Financial Economics* 12 (1983), pp. 105–127, and M. Gultekin and B. Gultekin, "Stock Market Seasonality: International Evidence," *Journal of Financial Economics* 12 (1983), pp. 469–481.
  34. S. Tinic, G. Barone-Adesi and R. West, "Seasonality in Canadian Stock Prices: A Test of the Tax-Loss-Selling Hypothesis," *Journal of Financial and Quantitative Analysis*, March 1987; P. Shultz, "Personal Income Taxes and the January Effect: Small Firm Stock Returns Before the War Revenue Act of 1917: A Note," *Journal of Finance*, March 1985; and C. Jones, D. Pearce and J. Wilson, "Can Tax-Loss Selling Explain the January Effect? A Note," *Journal of Finance*, June 1987.
  35. G. Constantinides, "Optimal Stock Trading With Personal Taxes: Implications for Prices and the Abnormal January Returns," *Journal of Financial Economics*, March 1984, and K. Chan, "Can Tax-Loss Selling Explain the January Seasonal in Stock Returns?" *Journal of Finance*, December 1986.
  36. Jacobs and Levy, "Disentangling Equity Return Regularities," *op. cit.*
  37. H. Shefrin and M. Statman, "The Disposition to Sell Winners Too Early and Ride Losers Too Long," *Journal of Finance*, July 1985.
  38. The payoffs shown are for an exposure of one cross-sectional standard deviation to the small-size attribute. For details, see Jacobs and Levy, "Disentangling Equity Return Regularities," *op. cit.* Here their results have been extended through the end of 1987.
  39. See H. Theil, *Applied Economic Forecasting* (Amsterdam: North-Holland, 1966), pp. 26–32.
  40. See E. Gardner, "Exponential Smoothing: The State of the Art," *Journal of Forecasting* 4 (1985), pp. 1–38.
  41. The current observation in an autoregressive (AR) process of order  $p$  is generated by a weighted average of  $p$  lagged observations. Similarly, the current observation in a moving-average (MA) process of order  $q$  is generated by a weighted average of  $q$  lagged errors. Methods for identifying and fitting time-series models are provided in G. Box and G. Jenkins, *Time-Series Analysis: Forecasting and Control*, rev. ed. (San Francisco: Holden-Day, 1976).
  42. The autocorrelation function is used to determine the order of the stochastic process. It provides a measure of the correlation between sequential data points. The  $n$ th-order autocorrelation is defined as the covariance between each observation and that of  $n$  periods earlier, divided by the variance of the process.
  43. A. Morgan and I. Morgan, "Measurement of Abnormal Returns From Small Firms," *Journal of Business and Economic Statistics*, January 1987; A. Lo and C. MacKinlay, "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test," *The Review of Financial Studies*, Spring 1988; and M. Levis, "Size Related Anomalies and Trading Activity of UK Institutional Investors," in E. Dimson, ed. *Stock Market Anomalies* (Cambridge: Cambridge University Press, 1988). While these studies find short-term autocorrelation patterns, J. Grant, "Long-term Dependence in Small Firm Returns" (Boston College working paper #84-10, March 1984), documents long-term cycles in the daily returns of small-firm portfolios, and Fama and French ("Permanent and Temporary Components of Stock Prices," *op. cit.*) find strong negative serial correlation for long-horizon returns of duration three to five years, especially for smaller firms.
  44. See M. Bartlett, "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series," *Journal of the Royal Statistical Society*, serial B8, 27 (1946), pp. 27–41, and Box and Jenkins, *Time Series Analysis*, *op. cit.*, pp. 34–36.
  45. See G. Ljung and G. Box, "On a Measure of Lack of Fit in Time Series Models," *Biometrika* 65 (1978), pp. 297–303, and G. Box and D. Pierce, "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models," *Journal of American Statistical Association* 65 (1970), pp. 1509–1526.
  46. Further confirming evidence is provided by various criteria that determine the order of a stochastic process. Akaike's AIC criterion determines the order of autoregressive processes. The Hannan-Rissanen criterion determines the order of autoregressive/moving-average (ARMA) processes. Both indicate that this return series cannot be

- successfully modeled as an ARMA process. H. Akaike, "Fitting Autoregressive Models for Prediction," *Ann. Inst. Stat. Math.* 21 (1969), pp. 243-247, and E. Hannan and J. Rissanen, "Recursive Estimation of Mixed Autoregressive-Moving Average Order," *Biometrika* 69 (1982), pp. 81-94 and "Correction" 70 (1983), p. 303. Significant autocorrelation patterns exist for pure returns to some other equity attributes. See Jacobs and Levy, "Disentangling Equity Return Regularities," *op. cit.*, pp. 43-44.
47. S. Makridakis, S. Wheelwright and V. McGee observe: "Unlike explanatory forecasting, time-series forecasting treats the system as a black box and makes no attempt to discover the factors affecting its behavior." *Forecasting: Methods and Applications*, 2nd ed. (NY: John Wiley & Sons, 1983), p. 18.
  48. See chapters 10 and 11 in Box and Jenkins, *Time Series Analysis*, *op. cit.*
  49. D. Umstead, "Forecasting Stock Market Prices," *Journal of Finance*, May 1977.
  50. Levis, "Size Related Anomalies and Trading Activity of UK Institutional Investors," *op. cit.*
  51. C. Granger, "Investigating Causal Relations By Econometric Models and Cross-Spectral Methods," *Econometrica*, July 1969.
  52. M. Quenouille, *The Analysis of Multiple Time-Series* (London: Griffon, 1957), pioneered the analysis of multivariate stochastic models.
  53. K. Chan, N. Chen and D. Hsieh, "An Exploratory Investigation of the Firm Size Effect," *Journal of Financial Economics* 14 (1985), pp. 451-471.
  54. K. Chan and N. Chen, "Business Cycles and the Returns of Small and Large Firms" (U. of Chicago working paper #229, January 1988). They form two mimicking portfolios—one of firms in distress, as measured by reductions in dividend payments, and another of highly levered smaller firms. They find smaller firms have higher sensitivities to the mimicking portfolios than do larger firms, even after controlling for firm size.
  55. Similar variables were used by K. Chan, N. Chen and D. Hsieh ("An Exploratory Investigation of the Firm Size Effect," *op. cit.*) to investigate links with the size effect, by N. Chen, R. Roll and S. Ross ("Economic Forces and the Stock Market," *Journal of Business* 59 (1986), pp. 383-403) to investigate links with stock returns, and by E. Fama and K. French ("Forecasting Returns on Corporate Bonds and Common Stocks" (U. of Chicago working paper #220, December 1987)) to investigate links with stock and bond returns. These studies transform the variables; for instance, default spread measures are formed from the difference between yields, or returns, on low-quality corporate and government bonds. The default spread, and other spreads, are implicitly incorporated in our approach as differences between the independent variables.
  56. Jacobs and Levy have demonstrated, for example, that value considerations alone are insufficient to explain security pricing. For instance, the effectiveness of the dividend discount model is dependent on market psychology (see Jacobs and Levy, "On the Value of 'Value'," *Financial Analysts Journal*, July/August 1988). And Japanese investments in U.S. stocks, which are generally concentrated in larger companies, are influenced by the dollar/yen exchange rate. For expository purposes, however, we limited our investigation to the six valuation variables.
  57. Although standard VAR models use a uniform lag length for all variables, C. Hsiao ("Autoregressive Modelling and Money-Income Causality Detection," *Journal of Monetary Economics* 7 (1981), pp. 85-106) has proposed a stopping-point criterion for choosing the optimal lag length for each variable in each equation.
  58. See R. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, March 1960, pp. 35-44; and R. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, March 1961, pp. 95-107.
  59. For a critique of the usefulness of VAR models for understanding macroeconomic relationships, see D. Runkle, "Vector Autoregressions and Reality," *Journal of Business & Economic Statistics*, October 1987. VAR modeling for small-scale applications was first proposed by C. Sims, "Macroeconomics and Reality," *Econometrica*, January 1980.
  60. As the number of variables increases, VARMA models face what G. Jenkins and A. Alavi ("Some Aspects of Modelling and Forecasting Multivariate Time Series," *Journal of Time Series Analysis* 2 (1981), pp. 1-47) call the "curse of higher dimensionality." C. Granger and P. Newbold (*Forecasting Economic Time Series* 2nd ed. (Orlando: Academic Press, 1986), p. 257) assert that: "even if we had any confidence in our ability to identify a VARMA model relating say seven or eight time series, the full parameterization would involve a huge number of unknown parameters. Not only is model estimation extremely expensive in this case, it is also rather foolhardy, since though such a nonparsimonious structure may fit an observed data set well, it is likely to prove very disappointing when extrapolated forward for forecasting purposes."
  61. R. Litterman, "Forecasting with Bayesian Vector Autoregression—Five Years of Experience," *Journal of Business and Economic Statistics*, January 1986.
  62. A. Zellner and F. Palm ("Time Series Analysis and Simultaneous Equation Econometric Models," *Journal of Econometrics*, May 1974) show that



any structural econometric model may be viewed as a restricted vector time-series model.

63. C. Sims ("Macroeconomics and Reality," *op.cit.*) argues that "existing large models contain too many incredible restrictions" and concludes that "the style in which their builders construct claims for a connection between these [structural] models and reality . . . is inappropriate, to the point at which claims for identification in these cannot be taken seriously." Also, C. Granger ("The Comparison of Time Series and Econometric Forecasting Strategies," in J. Kmenta and J. Ramsey, eds., *Large-Scale Macroeconomic Forecasting* (Amsterdam: North-Holland Publishing Company, 1981)) has asserted that "a 'moderate'-size econometric model of 400 or so equations is beyond the scope of current macroeconomic theory. The theory is hardly capable of specifying all of these equations in any kind of detail."
64. Bayes' theorem is the basis for combining "prior" beliefs with sample information to form a "posterior" distribution. See T. Bayes, "Essay Towards Solving a Problem in the Doctrine of Chances," *Phil. Trans. Royal Society* 53 (1763), pp. 370-418 (reprinted in *Biometrika* 45 (1958), pp. 293-315), and A. Zellner, *An Introduction to Bayesian Analysis in Econometrics* (NY: John Wiley, 1971).
65. Litterman, "Forecasting with Bayesian Vector Autoregression," *op.cit.*, p. 35. For a discussion of the merits of Bayesian VAR modeling, see R. Litterman, "A Statistical Approach to Economic Forecasting," *Journal of Business & Economic Statistics*, January 1987, and S. McNees, "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts," *Journal of Business & Economic Statistics*, January 1986. For an elaboration on the technique, see T. Doan, R. Litterman and C. Sims, "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Review* 3 (1984), pp. 1-144.
66. Antecedents in the literature include ridge regression, smoothness priors and Stein-James shrinkage estimators. For applications in finance, see C. Martin, "Ridge Regression Estimates of the Ex Post Risk-Return Trade-Off on Common Stocks," *Review of Business and Economic Research*, Spring 1978; R. Shiller, "A Distributed Lag Estimator Derived From Smoothness Priors," *Econometrica*, July 1973; and Jacobs and Levy, "Disentangling Equity Return Regularities," *op.cit.*, p. 32.
67. See Litterman, "Forecasting with Bayesian Vector Autoregression," *op. cit.*, p. 33, and J. Kling, "Predicting the Turning Points of Business and Economic Time Series," *Journal of Business*, April 1987.
68. L. Bachelier, *Theorie de la Speculation* (Paris: Gauthier-Villars, 1900).
69. While short-run stock returns are approximated well by a random walk, there is evidence of a mean reversion tendency for longer-run returns. See E. Fama and K. French, "Permanent and Temporary Components of Stock Prices," *op.cit.*
69. The shocks are also referred to as "innovations," because they represent surprises not predicted from past data. They are constructed to be orthogonal by taking into account contemporaneous correlations with the other macroeconomic variables.