Essays on Financial Economics

by

Francesco Franzoni

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2002

© Francesco Franzoni, MMII. All rights reserved.

Author . . . .
Department of Economics
July 15, 2002

Certified by . . . .
Sendhil Mullainathan
Mark Hyman Jr. Career Development Assistant Professor of Economics
Thesis Supervisor

Certified by . . . .
Jonathan Lewellen
Jon D. Gruber Assistant Professor of Finance, Sloan School of Management
Thesis Supervisor

Accepted by . . . .
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Department Committee on Graduate Studies
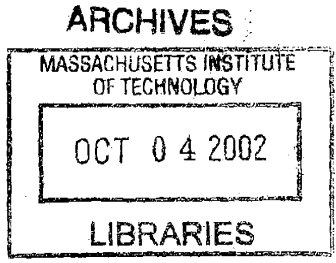
Essays on Financial Economics

by

Francesco Franzoni

Submitted to the Department of Economics
on July 15, 2002, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

The thesis consists of three essay on asset pricing topics.

The first essay finds that the market betas of value and small stocks have decreased by about 75% in the second half of the twentieth century. The decline in beta can be related to a long-term improvement in economic conditions that made these companies less risky. The failure to account for time-series variation of beta in unconditional CAPM regressions can explain as much as 30% of the value premium. In some samples, about 80% of the value premium can be explained by assuming that investors tied their expectations of the riskiness of these stocks to the high values of beta prevailing in the early years.

Moving from these findings, the second essay (co-authored with Tobias Adrian) explores in detail the relation between the 'value premium' and the decrease in value stocks' beta. We develop an equilibrium model of learning on time-varying risk factor loadings. In the model the CAPM holds from investors' ex-ante perspective. However, the econometrician can observe positive mispricing, whenever the expected beta is above the true level. Given the finding of a decreasing beta, it is likely that investors' expectation of the beta of these stocks has been above the actual level. Therefore, our model can provide an explanation for the 'value premium'. We present the results of simulations in which the model accounts for up to 80% of the 'value premium' in the 1963-2000 sample.

The third essay analyzes the response of stock returns to earnings information. First, I test the assumption that market expectations of earnings reflect a seasonal random walk, despite the actual process being autoregressive. This hypothesis is rejected. Second, I test the opposite view that expectations are unbiased. The data rejects this possibility for small firms. On the other hand, large firms' prices provide evidence of efficiency. Finally, I show that in the case of small firms the market understates the autoregression coefficient in the earnings process, and it incorrectly assumes that this coefficient is positive, even when actual earnings are seasonal random walks.

Thesis Supervisor: Sendhil Mullainathan
Title: Mark Hyman Jr. Career Development Assistant Professor of Economics

Thesis Supervisor: Jonathan Lewellen
Title: Jon D. Gruber Assistant Professor of Finance, Sloan School of Management

Acknowledgments

Many are the persons who helped me to accomplish this work, and to whom I am grateful.

First of all I would like to thank my thesis advisors Jonathan Lewellen and Sendhil Mullainathan. They provided guidance and support in all the stages of this process. It was for me a pleasure to be exposed to their knowledge and their method of scientific research. I am also grateful to Dimitri Vayanos, who served as my third advisor. It is thanks to his teaching that I developed an interest in Finance in the first place. Professors Olivier Blanchard and Bengt Holmström guided me through important decisions on my career. I am very thankful and honored for their attention. Other faculty members at MIT have contributed to the development of this thesis with helpful comments and insights. In particular, I would like to thank: Xavier Gabaix, S.P. Kothari, Guido Kuersteiner, and Jiang Wang.

An important part of this thesis is joint work with Tobias Adrian. Tobias, besides being a good friend, is an excellent researcher, and it was a pleasure to work with him. I thank him for always believing in my work, and in the successful realization of our idea.

The staff at MIT has made working here an enjoyable experience on an every day basis. I am especially grateful to Gary King, whose availability, thoroughness, and kindness are crucial ingredients for the success of the Economics Department.

There are several fellow students at MIT, who contributed to this thesis in different ways. Some of them with helpful comments and insights on the papers. Others by explaining economics to me along these years. Still others by playing soccer and partying with me. All of them by simply being my friends, laughing with me, and listening to the intricate stories of my life whenever it was necessary (very often indeed!). Among them I particularly thank: Manuel Amador (and Doireann Fitzgerald, for being SOOO nice),

Roberto Benelli (for his friendly company, and his parents, for the great dinners), Matilde Bombardini (for good counsel in non-economical matters), Toan Do (for his constant and reliable presence), Francesco Franco (for the enjoyable time during dinners, movies, and conversations), Veronica Guerrieri (for our daily talk shows), Ty Harris (for organizing soccer, skiing, dancing, ... etc.), Jonathan Kearns (for always being able to find good opportunities to celebrate), Augustin Landier (for the glimpses I had of his genius), Malte Loos (for sharing time in the office and the dojo), and Daniel Paravisini (for the fun time together).

I am also grateful to my friends in Italy, who were always available to start from where we had left. A special thank to Giuseppina, who walked with me the important steps towards this Ph.D.

Finally, and most importantly, without the help and support of my family this adventure would not have been possible. Therefore, I thank my mother, my sister Giuliana, and my brother in law Osvaldo.

This thesis is dedicated to my father.

Cambridge, July 2002

6

Contents

Chapter 1

Where is Beta Going? The Riskiness of Value and Small Stocks

1.1 Introduction

Since the development of the Capital Asset Pricing Model (CAPM) by Sharpe (1964) and Lintner (1965), beta risk has become an important input into many asset-pricing applications. The market beta of a portfolio plays a central role not only in the academic tests of the CAPM, but also in mutual fund performance evaluation, portfolio optimization and cost of capital estimation. Beta is also of independent interest, as it summarizes some of the relevant characteristics of the firm's fundamentals. The analysis of the direction and the causes of the change in a firm's beta is informative on the relationship between the company's payoffs and general business conditions, as well as on the correct asset-pricing model. Consequently, understanding whether and how the market beta changed for some portfolios is informative on the reasons behind the failure of CAMP in pricing those portfolios For all of these reasons, this paper takes a close look at the evolution of beta for book-to-market (B/M) and size portfolios, specifically those stocks that create major problems for the CAPM (e.g., Fama and French, 1992 and 1993).

Using monthly data from 1926 to 2000, I find a striking decrease in the market beta of value and small stocks. Beta fell by about 75% in sixty years for both these portfolios. In the case of value stocks beta dropped from 2.20 in the early forties to 0.55 in the late nineties. Similarly, small stocks' beta dropped from 2.50 to 0.65. This decline does not seem to follow mechanically from portfolio formation procedures, nor does it depend on a reduction in the leverage of these companies. Also, the magnitude of this decline is by no means solely imputable to the behavior of these portfolios in the late nineties. Moreover, the decrease in beta can be explained by a decline in the volatility of these portfolios relative to the rest of the market, rather than by a drop in their correlation with the market.

The paper also attempts to explain this evolution of beta. The conditional CAPM literature provides one way to go about this task. Following Shanken (1990), I assume a linear relation between beta and some state variables, and estimate the parameters of this function in a conditional CAPM time-series regression. The resulting fitted beta series tracks very closely the original estimated series, and it captures as much as 71% of its variance in the case of value stocks. The variables that I use as instruments (the T-bill rate, the dividend yield, the default spread, the term spread, and the growth rate of industrial production) are tightly linked to general economic conditions (see, e.g., Fama and French, 1989). The general result is that when the state variables predict an improvement in the economy, these stocks' betas become smaller. This result not only applies to the cyclical movements of beta, but also to the long run decrease, because some of these variables display a trending behavior that is believed to be related to a long-term improvement in business conditions. This evidence is consistent with the findings in Lettau and Ludvigson (2001), who show that the returns of value stocks are more highly correlated with fundamental factors when times are bad.

12

The evidence suggests that the decline in beta has to be tied to the effect of better economic conditions on the structure of value and small companies' cashflows. Since these companies are presumably more prone to financial distress (Chan and Chen, 1991, Fama and French, 1995), a general reduction in the likelihood of distress can have made their payoffs less risky. To investigate this explanation, I consider whether the decrease in the beta can be imputed to a decrease in these firms' cashflow sensitivity to the market. After breaking excess returns into components related to news about future dividends, news about future excess returns, and news about future real interest rates, following Campbell and Mei (1993), I express the overall market beta as the sum of the betas of each of these components with the market. Using this approach within a rolling regression framework, I can determine the importance of each component in the observed decrease of the overall beta. The conclusion of this analysis is that the decline occurs because of a fall in the dividend news beta. Overall, the results from the beta decomposition are consistent with the interpretation of the conditional CAPM analysis, because they point in the direction of reduced cashflow riskiness.

The decrease in the beta of value and small stocks is interesting by itself, as it sheds light on the behavior of portfolios widely used in empirical studies and in the asset management industry. However, the fact acquires even more relevance if it can be related to the debate on the CAPM anomalies. This paper establishes a connection between the decrease in the beta of these portfolios and the emergence of a premium in their expected return in two ways.

The first way is suggested by the evidence that conditioning information tracks the variation in beta. Failing to consider the variability of betas causes the constant in the unconditional CAPM time-series regressions to capture some of the effect of the state vari-

ables on the beta. Since the portfolios for which the decrease in beta is bigger (value and small stocks) are the ones that load more heavily on the state variables, these portfolios are more likely to have a high premium. It turns out that for value stocks as much as 30% of the alpha in the time-series regressions can be explained by the time-varying beta.

A connection can also be drawn with the behavioral explanation of the CAPM anomalies. Daniel and Titman (1997), for example, argue that characteristics, rather than risk, are priced in equilibrium. They suspect that investors consider these stocks more risky than they actually are. I argue that the large drop that occurred in the betas of value and small stocks could have been the reason why investors made mistakes in the assessment of risk. The market could have incorrectly tied its expectation of the price for risk to the high levels of beta, which characterized these stocks until the early sixties, even after beta had experienced a major decrease. Support for this conjecture comes from the result that about 80% of the value premium in the second part of the sample (1963-2000) can be explained assuming a beta such as the one estimated in the first part of the sample (1926-1962).

This paper is organized as follows. Section 1.2 presents in detail the decrease in the estimated beta of value and small stocks, and tests the robustness of the fact to mechanical explanations. Section 1.3 explains the path of beta using conditional information, as in a conditional CAPM analysis. Section 1.4 implements Campbell and Mei's (1993) beta decomposition in order to identify the sources of the change in beta. Section 1.5 relates the decrease in the beta to the mispricing of value and small stock portfolios. Section 1.6 draws the conclusions of this work.

1.2 The decrease in the beta of value and small stocks

1.2.1 The data

The data come from the merger of three different sources. Monthly return data are taken from the Center for Research in Securities Prices (CRSP) database, which covers NYSE, Amex and Nasdaq stocks between January 1926 and December 2000. Accounting data come from two sources. The Compustat annual research file contains the relevant information for most publicly traded US stocks. This information is supplemented by Moody's book equity information manually collected by Davis, Fama, and French (2000) [1]. Their paper contains a precise definition of the book-value-of-equity variable.

Portfolios are formed according to the procedure described in Fama and French (1993). At the end of June of year t stocks are sorted on either B/M or size. B/M is measured as the ratio of book value of equity at the end of year $t-1$ to market value of equity in December of year $t-1$. Size is market capitalization, i.e. price times shares outstanding, at the end of June of year t. All stocks are assigned to ten deciles for each characteristic using the break-points of the distribution of NYSE stocks. For each decile a portfolio return is computed between July of year t and June of year $t+1$ as the value-weighted return of the stocks in the decile. The excess returns (returns minus the one-month Treasury Bill rate) on these ten B/M and ten size portfolios are the main variables of interest in this paper. From now on, unless otherwise specified, when I refer to 'value stocks' I mean the tenth B/M decile, and by 'small stocks' I mean the first size decile.

Panel A of Table 1.1 provides some summary statistics for the portfolios. Notice the similarities between the small and value stocks portfolios in terms of means and standard

---

[1] I thank Ken French for providing me with the accounting data. The portfolio returns can be downloaded directly from his web-site.

deviations of returns, and the high negative correlation between the B/M and size decile assignments, especially in the first part of the sample. These two categories of stocks become more homogenous to the rest of the market in terms of mean and standard deviation of returns in the second part of the sample, when also the correlation between the decile assignment decreases in absolute value. This last fact is consistent with the results in Fama and French (2001), who show that a large part of newly listed firms tend to be small firms with the glamour characteristic.

1.2.2   The evolution of beta

The first graphical impression of the decrease in the estimated beta of value and small stocks can be obtained from Figures 1-1 and 1-2. The figures plot the series of estimates of beta for these two portfolios. The estimates come from rolling regressions, with five-year estimation windows and one-month increments. The sample goes from July 1926 to December 2000. The tenth B/M decile portfolio (Figure 1-1) displays drastic changes in beta that can be as high as 2.2 between July 1938 and June 1943, and as low as 0.55 between December 1995 and November 2000. Similarly, the beta of the first size decile portfolio (Figure 1-2) peaks at 2.5 between September 1939 and August 1944, and it touches the minimum at 0.65 between April 1991 and March 1996. For both portfolios, betas display an increase at the beginning of the sample, peaking in the early forties. Then the series experience a large decline until the beginning of the sixties, when for both portfolios beta drops below one, this decrease being more pronounced in the case of small stocks. In the sixties the two series rebound above one, being more or less stable through the beginning of the eighties, when they start dropping again. From the mid-eighties through all the nineties the betas stay below one. In spite of the short-term swings, the long-term picture that emerges is the

16

decreasing trend that caused value and small stocks' estimated beta to decrease by 75% in about sixty years.

The magnitude of the standard errors of the betas is such that we can statistically rule out the equality of the estimates from different subsamples. For example, the 2.2 estimate of beta for value stocks in the 7/38-6/43 subsample has a standard error of 0.17, while the standard error for the 0.55 estimate from the 12/95-11/00 interval is 0.07. In the middle of the sample, namely in the first half of the sixties, when beta lingers around 1.2, the standard error is about 0.10. More generally, one can check if beta takes on statistically different values over time by performing tests of structural change. The results of these tests for the tenth B/M and first size decile portfolios (not-reported) reject the equality of the betas over any subsample in which the total 1926-2000 sample can be split. Even in the shorter 1963-1991 sample, which is the one used by Fama and French (1993), the tests reject the equality of the betas between subsamples for many possible splits. The beta in the second subsample is significantly smaller than the one in the first subsample.

In order to compare the time behavior of the beta for the different B/M and size deciles, I regress (the log of) each beta series on a time trend. The results are reported in Table 1.2. The t-statistics are computed using Newey-West estimator of variance which corrects for the autocorrelation due to the use of overlapping windows in estimating beta. In the entire 1926-2000 sample, the trend for B/M portfolios (Panel A) is negative for deciles four through ten, and it decreases uniformly from the first to the tenth decile. The fact that beta increases for the lower deciles is the mirror image of the increase for the highest deciles, and it is consistent with the theoretical constraint that the value-weighted sum of the betas is one. The estimated trend in the beta of the tenth B/M decile portfolio is -0.1% per month (1.1% annually). In the case of size portfolios (Panel B) the trends are negative for all the

deciles but the last one. The trend in the small stock portfolio is -0.08% per month (-0.9% annually).

One might wonder if the responsibility of the negative trend lies with the big drop that the betas experienced in the fifties. In fact, the trend in the beta estimates for the high B/M and low size deciles is still there, even when I let the estimation sample start in July 1963, which is the beginning of Fama and French's (1993) sample. From Table 1.2 one can see that trend coefficients for value and small stocks are actually larger in absolute value in the shorter samples. The trend in beta for the first size decile portfolio in the 1963-2000 sample is twice as much as in the overall sample.

Given the large correlation between the small and value characteristics reported in Table 1.1, the question could rise whether the decrease in the betas is a small stock phenomenon. A first reply to this question can be the fact that the negative trend is actually larger for the tenth B/M decile portfolio than for the first size portfolio. The relevance of the value characteristic also appears from a double sort of stocks by size and B/M. Companies are assigned to five quintiles for each characteristic, and then twenty-five portfolios are formed from the intersection of the two sorts, like in Fama and French (1993). I perform the rolling regressions analysis on these twenty-five portfolios and obtain the series of beta estimates. The estimated trend (not reported) in the portfolio of big high B/M stocks (fifth size quintile and fifth B/M quintile) is still -0.1% per month (t-stat. = -6.54). Moreover, these stocks are on average bigger than the companies in other portfolios that are in lower B/M deciles, and for which the trend is positive. For example, the average size of the companies in this portfolio is over twenty times that of the stocks in the intersection of the second size and first B/M quintiles, for which the trend in the beta is instead 0.03% (t-stat. = 2.26). This evidence confirms that the value characteristic is relevant independently of size. I

18

can infer that also the size characteristic matters by itself from the fact that the beta of the portfolio of small glamour stocks (first size and first B/M quintiles) has a significantly negative trend (-0.04%, t-stat. = -2.69)/footnoteI consider the effect on the observed trend in the market sensitivity of small and value stocks of the introduction of Fama and French's (1993) HML and SMB factors. As one might expect, the coefficient on the market factor is no longer decreasing over time for value and small stocks. In fact, this coefficient captures returns sensitivity to the component of the market that is orthogonal to HML and SMB, and these portfolios mimic the behavior of value and small stocks. Therefore, the trending behavior in the beta that is peculiar of these two categories of stocks is filtered out by the inclusion of HML and SMB..

1.2.3 A different perspective

A different way to look at the decline in beta is asking whether it is imputable to a decrease in the correlation of these portfolios with the rest of the market, or to a drop in their relative volatility. This analysis generates some additional evidence that any explanation of the decrease in beta will have to account for.

We can consider the market index as composed of two portfolios. Portfolio 1 is either the value or small stock portfolio, and portfolio 2 is the rest of the market.

$$R_m = w_1 R_1 + w_2 R_2, \tag{1.1}$$

where $w_1 + w_2 = 1$.

The beta of portfolio 1 can be written as

$$
\begin{aligned}
\beta_1 &= \frac{Cov(R_1, R_m)}{Var(R_m)} \\
&= \frac{Cov(R_1, w_1 R_1 + w_2 R_2)}{Var(w_1 R_1 + w_2 R_2)} \\
&= \frac{w_1 \sigma_1^2 + w_2 \sigma_{12}}{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_{12}} \\
&= \frac{w_1 r + w_2 \rho}{w_1^2 r + w_2^2 \frac{1}{r} + 2w_1 w_2 \rho},
\end{aligned}
\tag{1.2}
$$

where σ_i^2 is the variance of return i, σ_{12} is the covariance between R_1 and R_2, $r = \frac{\sigma_1}{\sigma_2}$ and ρ is the correlation coefficient between R_1 and R_2 [2].

The changes in beta are governed by the changes in its two components r and ρ. We can study the sign of the derivatives of beta with respect to these components:

$$
\frac{\partial \beta_1}{\partial r} > 0, \quad \frac{\partial \beta_1}{\partial \rho} > 0 \Leftrightarrow \frac{w_2}{w_1} > \frac{\sigma_1}{\sigma_2}
$$

The size of both the value and small stock portfolios relative to the rest of the market is so small [3] that the condition for $\frac{\partial \beta_1}{\partial \rho} > 0$ is always respected.

Equation 1.1 implies that

$$
w_1 \beta_1 + w_2 \beta_2 = 1
$$

Therefore β_1 and β_2 mechanically move in opposite directions, if weights are constant.

---

[2] A more straightforward way of decomposing beta is:

$$
\beta_1 = \rho_{1,m} \frac{\sigma_1}{\sigma_m},
$$

where $\rho_{1,m}$ is the correlation coefficient between R_1 and the market return, σ_1 is the standard deviation of R_1, and σ_m is the standard deviation of the market return. This decomposition yields the same results as the one in the text. In particular, $\frac{\sigma_1}{\sigma_m}$ and $\rho_{1,m}$ track closely $\frac{\sigma_1}{\sigma_2}$ and ρ, respectively. The advantage of the decomposition in the text is that portfolio 2 does not contain stocks from portfolio 1, which makes the interpretation of the results unambiguous.

[3] The share of the tenth B/M decile portfolio is on average 2% of total market capitalization, and that of the first size decile is on average 1%.

20

Figure 1-3 graphs the estimated beta for the tenth B/M decile portfolio along with its two components: the ratio of the volatility of high B/M stocks (10th decile) to the volatility of the rest of the market (1st to 9th deciles), and the correlation coefficient between these two portfolios. The volatility is estimated as the standard deviation of the portfolio monthly excess returns over a five-year rolling window. Similarly, the correlation is the correlation coefficient between the returns of the two portfolios over five-year rolling windows.

It appears clearly from the figure that the driving force behind the movements in the betas is the ratio of the volatilities. This impression is confirmed by regressing (the log of) the ratio on a time trend. The coefficient is 0.08% per month, very close to the 0.1% of the betas in Table 1.2, while the estimated correlation decreases only by 0.01% per month, and this trend is largely driven by the drop in the nineties. From Figure 1-3 one can notice that the estimated volatility of value stocks was 2.6 times that of the rest of the market between August 1938 and July 1943, and it dropped to 0.7 times in the period between December 1995 and November 2000. A similar picture (not reported) describes small stocks' beta and its components.

One might be concerned that the change in the weights of the portfolios might affect the comparability of the series in Figure 1-3 with the decomposition in Equation (1.2). In fact, I obtain a similar plot when I use portfolios constructed to have constant weights throughout the sample.

To complete the picture one needs to describe the evolution of idiosyncratic risk for these portfolios. For both value and small stocks, idiosyncratic risk, computed as the estimated standard deviation of the residuals from rolling window CAPM regressions, follows broadly the path of market volatility, which is documented in Schwert (1989) and Campbell et al. (2001). Hence, idiosyncratic risk peaks in the years of the Great Depression and World War

21

II, but then it drops drastically, without displaying any trending behavior. However, unlike market volatility, the idiosyncratic volatility of these portfolios is high in the first half of the nineties. The absence of a trend in portfolio idiosyncratic risk is not inconsistent with the finding in Campbell et al. (2001) that individual stocks have become more volatile. In fact their paper also finds that portfolio idiosyncratic volatility (in their case at industry level) is not trending.

From a market model, where portfolio return is broken into market risk and idiosyncratic risk, it follows that the variance of portfolio 1 can be expressed as

$$\sigma_1^2 = \beta^2 \sigma_m^2 + \sigma_{1,\varepsilon}^2, \tag{1.3}$$

where σ_m^2 is the variance of the market return, and $\sigma_{1,\varepsilon}^2$ is the idiosyncratic variance of portfolio 1. Since beta can be expressed as the product of the correlation coefficient between portfolio 1 and the market return ($\rho_{1,m}$) times the ratio of portfolio 1 standard deviation to market standard deviation (see footnote 2), Equation (1.3) can be rearranged to obtain

$$\frac{\sigma_{1,\varepsilon}^2}{\sigma_m^2} = \frac{\sigma_1^2}{\sigma_m^2} \left(1 - \rho_{1,m}\right). \tag{1.4}$$

Given that $\frac{\sigma_1}{\sigma_m}$, like $\frac{\sigma_1}{\sigma_2}$, is decreasing much more strongly than $\rho_{1,m}$, which in turn tracks closely ρ, it has to follow that idiosyncratic risk as a fraction of market volatility has decreased. The data confirm this prediction, and the evolution of the ratio of idiosyncratic volatility to market volatility follows closely the ratio of total portfolio volatility to the volatility of the rest of the market that is plotted in Figure 1-3 (thick solid line).

22

1.2.4 Robustness checks

The decrease in the estimated market beta of value and small stocks might be the artifact of portfolio formation procedures or, more generally, it can be a mechanical result with little economic content. In order to investigate this possibility I perform a number of robustness checks.

The share of value and small stocks' capitalization over total market size has changed over time. This could have caused the decrease in these portfolios' betas, by mechanically reducing their weight in the market. I construct a portfolio that includes the highest B/M stocks up to a certain share of market capitalization, which I keep constant over the entire sample. I try with a market share of 2%, which is the average market share of the tenth decile portfolios, and with other values as well (1%, 3%, 5%, and 10%). In all of these cases the beta of the resulting portfolio, estimated with the rolling regression methodology described above, displays a comparable decrease to the one for the original value portfolio. Similarly, I rank the stocks by size, and construct a portfolio of small stocks that has a constant share of market capitalization. For different market shares (0.1%, 1%, and 5%), the beta of this small stock portfolio is still decreasing.

Another change that occurred in the portfolio composition is the strong increase in the number of stocks included in the portfolios. There were 42 companies in the value portfolio in July 1926 (52 in the small stocks portfolio), while this number was 480 in December 2000 (the number is 2502 in the case of small stocks) [4]. The increased number of included stocks might have affected the portfolio beta if it was combined with some change in the shape of the cross-sectional distribution of betas. Hence, I form portfolios of high B/M and

---

[4]The reason why there are so many stocks in the lower size deciles is that size portfolios are formed using NYSE capitalization break-points, and many Nasdaq stocks are small compared to NYSE stocks.

small stocks that have a constant number of stocks throughout the sample period. These portfolios continue to display a decline in their estimated betas, for all the number of stocks at which the composition is held fixed.

A related fact is the inclusion in the data set of Nasdaq stocks in 1973. This event was relevant especially for the small stock portfolio, since Nasdaq stocks were in general smaller than NYSE stocks. This inclusion could have affected the portfolio beta because the market index is heavily tilted towards NYSE stocks. However, when Nasdaq stocks are excluded from the portfolios, the trend in both the small stocks' beta and the value stocks' beta is unaffected.

Another objection that could be raised against the relevance of the fact under examination, is that the industry composition of the value and small stocks portfolios might have changed over time in such a way that these portfolios are now composed of firms belonging to industries that bear less market risk. The first control that I perform is a within industry analysis. I construct the value and small stock portfolios using only stocks in one industry, and restrict the attention to industries that presumably did not experience major technological changes, so that I control for industry effects. For all the industries I consider (food, consumer products, clothing and oil), the betas of value and small stocks significantly decrease over time. An alternative control for industry effects consists of replacing the return of each stock in the portfolios with the return of the industry portfolio to which the stock belongs [5]. If the trend in beta is due to the B/M or size characteristics, as opposed to industry effects, we should expect that the beta of these new portfolios does not trend down. Consistent with this expectation, the resulting portfolios do not display the same decrease as the original value and small stock portfolios. In the case of value stocks, for

---

[5] I used the 17 industry portfolio classification that can be found on Ken French's web site.

example, the estimated trend in the portfolio constructed with industry returns is -0.03%, compared to the -0.10% of the original value portfolio. I interpret the fact that there is still some decrease in the betas of the new portfolios as due to the correlation between industries and the B/M characteristic.

The decrease in the market beta of value and small stocks could be the result of a decline in the leverage of these companies. Lower leverage should lead to a smaller beta. To assess whether this phenomenon is driving beta, Figure 1-4 plots the leverage series, defined as book value of debt over market value of equity, for the value stock portfolio and for the rest of the market [6]. The figure shows that, if anything, there was an increase in the leverage of value companies over time, so that leverage is not driving the decrease in beta. As leverage in the overall market is increasing, one might be concerned with the evolution of leverage for the portfolios of interest relative to the rest of the market. In fact, not even the ratio of value companies' leverage to the leverage of the rest of the market displays a decreasing trend. Similar results rule out a leverage effect for small companies.

Other possible explanations of the observed decrease in the betas are linked to changes in the informational flows in the market. Lo and MacKinlay (1990) note that the positive autocorrelation of stock indices is mainly determined by cross-autocorrelations. In particular, large stocks tend to lead small stocks, possibly because of non-synchronous trading. As noted by Scholes and Williams (1977), if a stock is infrequently or non-synchronously traded, the standard estimate of beta is not representative of its true sensitivity to the market. Hence, it is possible that changes in the pattern of non-synchronous trading for small and value stocks determined the decrease in their beta. In order to control for this

---

[6]The accounting data come from the Compustat annual dataset and they start in 1950. Debt is defined as the sum of book value of current liabilities, long-term debt, convertible debt and preferred stocks. Portfolio leverage is computed as value weighted average of company leverage.

possibility, I compute a corrected version of the sensitivity to the market as the sum of the beta on the lagged monthly return and the standard beta (as suggested by Scholes and Williams, 1977). This correction does not affect the size of the estimated negative trend in the value and small stock portfolios' market sensitivity.

More generally, every explanation that relates to changes in the informational structure in the market should have different implications at different frequencies of the data. In low frequency data information has had more time to reveal itself than in higher frequency data. Hence, if the decrease in the beta is related to some informational story, it should be less pronounced at lower frequencies. Using quarterly and annual overlapping data, and extending the estimation window to ten years in order to have enough data points, does not seem to give different results from the ones obtained with monthly data. For values stocks, with all three data types the estimated beta drops from about two to below one. Similarly, for small stocks the beta drops as much with annual and quarterly data as with monthly data.

In summary, the drop in beta does not seem to depend on mechanical explanations relating to portfolio formation procedure, nor does it depend on changes in leverage over time. Moreover, changes in the patterns of non-synchronous trading do not seem to be relevant.

The next section, which relates the decrease in beta to macroeconomic conditions, is, I believe, the most convincing reply to most doubts that still linger about the economic relevance of the decline in beta of value and small stocks.

1.3 Relating beta to macroeconomic conditions

1.3.1 Time-varying betas and conditioning information

Several studies have produced evidence of time-varying betas for single stocks and for portfolios (e.g., Ferson and Harvey, 1991, Ferson and Korajczyk, 1995, Braun, Nelson, and Sunier, 1995). Shanken (1990) models the time variation of conditional betas as a linear function of predetermined state variables. Later studies apply this approach to testing multi-factor pricing models (Ferson and Korajczyk, 1995, Ferson and Harvey, 1999, Lewellen, 1999), and mutual fund performance evaluation (Ferson and Schadt, 1996).

In the context of this paper, modeling conditional betas as a function of state variables can help identify the macroeconomic factors, if any, that are driving the decrease in the beta of value and small stocks.

The rationale to believe that some economic state variables are related to the decrease in betas is that the value and small characteristics supposedly denote companies that are in a condition of relative distress [7]. Hence, it is reasonable to believe that changing macroeconomic conditions affect the severity of this condition of distress, and consequently the riskiness of these stocks's payoffs, as summarized by their market beta.

Suppose the following conditional one-factor model describes the excess portfolio return $R_{i,t+1}$

$$R_{i,t+1} = \alpha_i + \beta_{i,t} R_{m,t+1} + \varepsilon_{i,t+1}, \qquad (1.5)$$

where $R_{m,t+1}$ is the market excess return, and $E_t(\varepsilon_{t+1}) = E_t(\varepsilon_{t+1} R_{m,t+1}) = 0$, which implies that the unconditional expectations of the same expressions are also zero.

---

[7]Chan and Chen (1991) show that small firms are more likely to have higher leverage, lower Returns-On-Equity, and have cut dividends in the recent past.

Following Shanken (1990), portfolio's betas are assumed to be a linear function of a vector of k state variables z_t

$$\beta_{i,t} = b_{0,i} + b'_{1,i} z_t + \eta_{i,t}. \tag{1.6}$$

While not imposing any constraint on the process of the market factor, the assumption that conditional betas depend linearly on some lagged variables allows the second moments of the conditional distribution of portfolio and market returns to change over time in a simple way. The variables used to predict conditional betas are public information, and summarize the state of the macroeconomy.

Using Equation (1.6) to replace for $\beta_{i,t+1}$, Equation (1.5) can be rewritten as

$$R_{i,t+1} = \alpha_i + b_{0,i} R_{m,t+1} + (b'_{1,i} z_t) R_{m,t+1} + \eta_{i,t} R_{m,t+1} + \varepsilon_{i,t+1}. \tag{1.7}$$

Providing that $\eta_{i,t}$ is regressively independent of all the information at time t, the sum $\eta_{i,t} R_{m,t+1} + \varepsilon_{i,t+1}$ can be considered as an orthogonal error term u_{t+1}, and the regression in Equation (1.7) yields consistent estimates.

The estimates of b_0 and b_1 from the time-series regression in Equation (1.7) allow us to obtain a fitted value for $\beta_{i,t}$

$$\widehat{\beta}_{i,t} = \widehat{b}_{0,i} + \widehat{b}'_{1,i} z_t, \tag{1.8}$$

which gives the benchmark series to which compare the observed decrease in the estimated betas of value and small stocks.

1.3.2 Empirical implementation

The state variables that I use in the analysis are the ones that in previous studies proved to be good predictors for expected returns and betas. They are: (1) the dividend yield on the

S&P Composite Index (see, e.g., Fama and French, 1988, Ferson and Harvey, 1999); (2) the one-month T-bill rate (see, e.g., Shanken, 1990); (3) the growth rate of industrial production, computed as the first difference in the logarithm of the monthly seasonally adjusted index of industrial production provided by the Federal Reserve Board (see, e.g., Campbell and Mei, 1993); (4) the term spread defined as the end-of-month difference between the yield on Aaa corporate bonds and the annualized one-month T-bill rate (see, e.g., Fama and French, 1989, Ferson and Harvey, 1999); (5) the default spread, defined as the end-of-month difference between the yields on Baa and Aaa corporate bonds (see, e.g., Fama and French, 1989, Ferson and Harvey, 1999).

Panel B of Table 1.1 provides summary statistics for the state variables. Figures 1-5 and 1-6 graph them, along with NBER business cycle dates.

Fama and French (1989) give a thorough discussion of the cyclical behavior of the state variables. Here I summarize the main points. The default spread, although showing some negative correlation with the business cycle, displays major swings that go beyond the economic cycle (Figure 1-5). The spread is high during the thirties and the early years of World War II, a period characterized by major economic uncertainty. In the rest of the sample it is lower except for some blips in the periods of recession during the seventies and early eighties. A similar behavior characterizes the dividend yield (Figure 1-5), which is highly correlated with the default spread. What is peculiar about the dividend yield is the drop that occurred during the bull market of the second half of the nineties. The T-bill rate gravitates around zero in the 1933-1951 period that covers much of the Great Depression and the period after World War II, when the Fed fixed T-bill rates. Outside that interval the T-bill rate comes close to defining the business peaks and troughs identified by the NBER (Figure 1-6). Since the Aaa yield does not track the business cycle as closely as the T-bill

rate, the term spread, except for the 1933-1951 period, follows more closely the business cycle (Figure 1-6) [8]. It is low at peaks, predicting recessions, and high at troughs, predicting recoveries. Finally, the growth rate of industrial production is strongly mean-reverting, so that high growth rates are soon followed by negative growth.

The estimation of the regression in Equation (1.7) for B/M and size portfolios produces estimates of b_0 and b_1 that can be replaced in Equation (1.8) along with the series of the state variables to fit the path of these stocks' beta. Figure 1-7 graphs the fitted beta for value stocks, along with the series resulting from the rolling regressions estimation. The fitted series in the graph has been constructed using two sets of estimates of b_0 and b_1 coming from the 1926-1962 and the 1963-2000 subsamples. The series of the estimated beta is aligned with the end date of the five-year estimation window. The tracking ability of the fitted beta (solid line) is striking. The estimated beta series appears smoother than the fitted series, because the effect of one month of data is not relevant over a five-year estimation horizon. However, the fitted beta follows closely all the main swings in the estimated beta. The correlation coefficient between the estimated and the fitted series is 0.84. The reader may be concerned that this high level of correlation is affected by a 'spurious regression' type of problem. To tackle this concern, I perform a test of unit root on the difference between the two series. In other words, I test whether the estimated and fitted beta are cointegrated with a (1 -1) cointegration vector. If the high correlation is spurious, the test should detect a unit root in the difference. In fact the correlation is authentic, as a Dickey-Fuller test on the difference in the two series produces a test statistic of -7.2, which rejects the null hypothesis of unit root at the 1% confidence level [9]. The picture is very similar if b_0 and b_1

---

[8]The T-bill series and the term series in Figure 1-6 appear to have different volatility. This is a result of using different scales for the two series. In reality they move together, being the term spread mainly driven by the T-bill rate component.

[9]Notice that the Dickey-Fuller test does not reject the null hypothesis that the estimated beta series has

are estimated over the whole 1926-2000 sample. The correlation is 0.78, and again the two series are cointegrated at the 1% confidence level.

Figure 1-8 plots the estimated and fitted beta series for small stocks. Although the fitted series does not track so closely the estimated one as in the case of value stocks, still it captures the major drop in the beta that occurred in the twenty years between the 1940 and 1960. The correlation coefficient is in this case 0.63, suggesting that perhaps some relevant state variable has been left out from the information set.

Table 1.3 reports the coefficients from the estimation of Equation (1.7) in the case of B/M portfolios. Looking at the column for the tenth decile, we notice that in the whole 1926-2000 sample (Panel A) the risk free interest rate has the highest predictive ability: a one-standard deviation increase (0.25%) in the monthly T-bill rate would cause a decrease of about 0.25 in the conditional beta. This coefficient decreases in absolute value as we move towards lower B/M deciles, consistent with the theoretical constraint that the weighted sum of the b_1 coefficients is zero (while the weighted sum of the b_0 coefficients is one). The default spread and the dividend yield have a similar predictive power for the conditional beta of value stocks: an increase in both variables causes the conditional beta to go up. For example, an increase of one percentage point in the annualized default spread causes the beta of the tenth decile portfolio to rise by 0.1. In the whole sample the term spread is generally not significant, while the growth rate of industrial production is significantly positively related to conditional betas only for the higher deciles.

The analysis by subsamples (Panels B and C of Table 1.3) helps to further clarify the effect of each conditioning variable. As far as the tenth B/M decile portfolio is concerned, the default spread and the dividend yield take turns in explaining the conditional beta. The

---

a unit root

31

first variable is significant only between 1926 and 1962, while the second one is significant only between 1963 and 2000. The T-bill rate is always negative and significant, although more so in the first subsample. The term spread is significant with a negative coefficient in both subsamples. The impact of the growth rate of industrial production on the conditional beta changes from positive to negative. In general, as we move towards lower B/M deciles the predictive power of the state variables drops, suggesting that it is correct to focus the attention on the changes that affected value stocks.

The regression results for small stocks (first size portfolio in Table 1.4) are similar to the case of value stocks. However, the predictive power of the state variables tends to drop in the second subsample. This fact is in line with the reduced tracking ability of the fitted series in Figure 1-8. Nevertheless, the major drop in small stocks' beta occurs before 1960, and that is mostly captured by the state variables.

Finally, other state variables turn out to be significant predictors of beta for both B/M and size portfolios. They have not been used for the plots in Figures 1-7 and 1-8, because their inclusion would have increased the high frequency volatility of the fitted series, and decreased its ability to track the smooth estimated series. The most important of these variables are the lagged excess market return and the volatility of the T-bill rate, constructed like in Shanken (1990). A positive market return predicts an increase in the beta of value and small stocks, whereas the effect of interest rate volatility is positive in the first subsample and negative in the second one. The purpose of my analysis was tracking the long run trend in the estimated beta series, but if the goal is predicting the future evolution of beta, then one may want to include these instruments, which capture the high frequency movements in the series of interest.

1.3.3 Discussion

As mentioned above, the connection between variables that summarize the state of the macroeconomy and the beta of value and small stocks can be drawn because these companies are more likely to be in a situation of relative distress. The likelihood with which they actually are in distress can reasonably depend on the general state of the economy. Hence, their riskiness, as summarized by the market beta, can vary as a function of the business cycle and the general economic conditions.

In more detail, one can think of a model where distressed firms approach default, or move away from it, depending on the evolution of economic conditions. When a company is closer to the earnings cutoff point below which it defaults, the firm's payoff distribution can become more volatile, and so can its stock returns. Then, this model can yield the prediction that the beta of distressed firms decreases when economic conditions improve, and vice versa.

The relationship between the macro variables and the beta of value and small stocks, that was found in the previous subsection, seems to be in line with this model. Notice that the improvement in business conditions that I refer to, can take place both along the business cycle, and over the long run. In the first case the model describes the high frequency variation in beta. In the second case, the result is the long-term decrease in beta that is the main focus of the paper.

Fama and French (1989) interpret the power of the dividend yield and the default spread to predict increases in the expected return of stocks and bonds as related to the long-term evolution of business risk. These two variables track some components of expected returns that are high during periods like the Great Depression when business is persistently poor, and low otherwise. Consistent with their interpretation, I find that the measure of riskiness

of some companies that are a priori believed to be more exposed to changes in business conditions, their beta, follows closely the evolution of these two variables. In particular, the decrease in the default spread and the dividend yield that occurred after the war explains a large part of the decline in beta in that period. Similarly, the drop in the dividend yield in the late nineties is responsible for the plunge in beta over those years.

The term spread tracks more closely the business cycle. It is low at the top of the expansion, and high at the end of a recession. I find that an increase in the lagged term premium predicts a decrease in the beta of value and small stocks. This result is in line with the above interpretation, because a high term premium predicts a recovery from a recession, from which distressed firms should benefit. The significance of the term spread in the subsamples, and not in the whole period, suggests that this variable captures the high frequency variation of beta, rather than the long run trend.

The interest rate is highly pro-cyclical. However, like the default spread and the dividend yield, it also tracks the long-term changes in business conditions that occurred starting from the early fifties. Therefore, its relevance as a predictor of beta is due to both its long-term swings and its cyclical movements. This fact is confirmed by the significance of the T-bill rate in whole sample and in the subsamples.

The reason behind the predictive power of the growth rate of industrial production is more dubious since its coefficient changes from positive to negative. The positive coefficient, that prevails when the estimation is performed on the whole sample, would suggest that due to the strong mean reversion of growth, a positive growth rate predicts a worsening of economic conditions in the future.

The interpretation of the relation of these macro variables to the decline in beta that is proposed here is also consistent with the evidence from the beta decomposition in Sec-

tion 1.2.3. The general improvement of macroeconomic conditions can have made these stocks, normally prone to distress, less risky, and their volatility smaller relative to the volatility of the rest of the market.

The above analysis suggests a relation between a conditional CAPM approach and the Fama-French three-factor model. The explanation of the link between these macroeconomic variables and the beta of value and small stocks involves changes in the sensitivity to the overall discount factor. The effect of these changes shows up in the beta of these companies because they are presumably more sensitive to business conditions. If this is the case, the loading on the market factor seems to capture some of the risk sources, like distress risk, that Fama and French (1993) use to justify the introduction of additional factors in the pricing model. Consequently, a conditional CAPM can be more appropriate than a three-factor model in pricing portfolios other than value and small stocks, for which the HML and SMB factors are bound to perform well.

Notice that the argument in favor of a conditional pricing model is consistent with the results in Lettau and Ludvigson (2001). Similar to the evidence in my paper, they find that the correlation value stocks' returns with fundamental factors increases when risk or risk aversion is high. This situation in turn occurs when economic conditions are poor, as signalled by their *cay* state variable.

Finally, it has to be acknowledged that the behavior of value stocks' beta in the second half of the nineties lends itself to a different interpretation from the one proposed so far. The drop in beta that occurs in that period is entirely driven by the decline in the dividend yield, which in turn depends on the surge of the price level during the bull market of the nineties. If a speculative bubble was behind that price increase, then the explanation for the decline in beta cannot hinge on the evolution of the stochastic discount factor. Hence,

35

one may want to invoke a style investing argument (Barberis and Shleifer, 2001). In such a scenario, beta could have dropped because the returns of value and glamour stocks have become delinked, as a result of flows of funds moving from one style of investment to the other. The investigation of this explanation is left for future research.

1.4 A decomposition of market betas

1.4.1 Theoretical framework

If the interpretation of the link found in Section 1.3 between beta and the conditioning variables is correct, the decrease in the beta of value and small stocks should be associated with a reduction in the sensitivity of these companies' cashflows to the factors that cause movements in the market. The reason behind this prediction is that, according to the above interpretation, an improvement in the distress condition of these companies makes their cashflows less volatile in response to shocks. I use Campbell and Mei's (1993) beta decomposition to address this issue.

Campbell and Shiller's (1988) log-linearized present value relationship allows one to express unexpected excess returns, or excess return innovations, in terms of news about dividends, news about excess returns and news about real interest rates. Following Campbell (1991), $e_{i,t+1}$ is the (continuously compounded) excess return on portfolio i over the (continuously compounded) real return $r_{i,t+1}$ on a one-month T-bill, and $d_{i,t+1}$ is the (log) real dividend. Then, portfolio i's unexpected excess return $\tilde{e}_{i,t+1}$ can be expressed as

$$
\begin{aligned}
\tilde{e}_{i,t+1} &\simeq (E_{t+1} - E_t) \left\{ \sum_{j=0}^{\infty} \rho^j \Delta d_{i,t+1+j} - \sum_{j=0}^{\infty} \rho^j r_{t+1+j} - \sum_{j=1}^{\infty} \rho^j e_{i,t+1+j} \right\} \\
&= \tilde{e}_{di,t+1} - \tilde{e}_{r,t+1} - \tilde{e}_{ei,t+1}.
\end{aligned} \tag{1.9}
$$

36

The notation $(E_{t+1} - E_t)$ indicates a revision in the conditional expectation between times t and $t+1$. The constant ρ comes from the linearization process, and can be interpreted as a discount factor. The value of ρ is assumed to be the same for all portfolios [10]. The second equality in (1.9) introduces simpler notation for dividend news $\tilde{e}_{di,t+1}$, real interest rate news $\tilde{e}_{r,t+1}$ and excess return news $\tilde{e}_{ei,t+1}$.

Equation (1.9) follows from an approximation of a present value identity after ruling out explosive behavior of stock prices, and can be thought of as a consistency condition for expectations. It simply states that, if unexpected returns are high today, then either there has been an upward revision in the expectation of future dividends, or a downward revision in the returns that the stock is expected to pay in the future, or both. The effect of future real returns is similar to that of future excess returns.

The Appendix describes in detail how to obtain each component of the return innovations. Briefly, the expected return on each portfolio is assumed to be a linear function of a vector of predetermined state variables, one of which is the real interest rate. The residuals in these predictive regressions represent the return innovations. The state variables are assumed to follow a VAR process. It is therefore possible to compute the revision in the expectation of every future value of the state variables, and take the discounted sum of these terms, which, combined with the parameters in the predictive regression, gives the expected return news component of returns. The real interest news component is also obtained from the parameters of the VAR. The cashflow news component is obtained residually using Equation (1.9) and the other two components.

As far as the cross-sectional aspects of the analysis are concerned, Campbell and Mei

---

[10]I refer to the Appendix in Campbell (1991) for the derivation of Equation (1.9) and to Campbell and Shiller (1988) for a discussion of its approximation accuracy. It turns out that $\rho = \frac{1}{1+e^{d-p}}$, where $d - p$ is set to the average log dividend price ratio. Campbell and Mei (1993) argue that the assumption of a unique value of ρ across all portfolios does not affect the results for plausible variations in ρ.

(1993) define a portfolio beta using unconditional variances and covariances of portfolio and market innovations. That is, beta is the unconditional covariance of the excess return innovation $\tilde{e}_i$ with the market innovation $\tilde{e}_m$, divided by the unconditional variance of the market innovation

$$\beta_{i,m} \equiv \frac{Cov(\tilde{e}_i, \tilde{e}_m)}{Var(\tilde{e}_m)}. \tag{1.10}$$

This beta is neither an unconditional beta (which would use returns themselves rather than innovations) nor a conditional beta (which would use conditional moments). However, it would coincide with a conditional beta if the conditional variance-covariance matrix of innovations had constant elements, or at least elements that changed in proportion to one another.

The definition in (1.10) has the advantage that the portfolio beta can be expressed as the sum of the market betas of the three news components. From Equations (1.9) and (1.10), it follows directly that

$$\begin{aligned}
\beta_{i,m} &= \frac{Cov(\tilde{e}_{di}, \tilde{e}_m)}{Var(\tilde{e}_m)} - \frac{Cov(\tilde{e}_r, \tilde{e}_m)}{Var(\tilde{e}_m)} - \frac{Cov(\tilde{e}_{ei}, \tilde{e}_m)}{Var(\tilde{e}_m)} \\
&= \beta_{di,m} - \beta_{r,m} - \beta_{ei,m},
\end{aligned} \tag{1.11}$$

where $\beta_{di,m}$ is the market beta of news about portfolio i's cashflows, $\beta_{r,m}$ is the market beta of news about future real interest rates, and $\beta_{ei,m}$ is the market beta of news about portfolio i's future excess returns.

The assumption behind this analysis is that the vector of state variables represents entirely investors' information set. If this was not the case, the estimated dividend component of returns would contain a reaction to changes in expected returns, which would undermine the interpretation of the results.

1.4.2 Empirical Implementation

The first step in implementing the beta decomposition developed in Section 1.4.1 is to estimate the return components. To this purpose one needs to estimate the VAR system (A-5), along with the predictive regressions in (1.18), and to replace the estimated parameters in the expressions given in (1.21).

The VAR and the predictive regressions are estimated using OLS on each equation. In this case, where all the equations have the same right-hand-side variables, the estimates coincide with the ones obtained with a GMM procedure. Then, I combine the sample variances and covariances of the estimated return components to obtain the betas.

The purpose of this analysis is to find out the source of the observed decrease in market beta for value and small stock portfolios. Thus, I insert the beta decomposition methodology into the rolling regression framework that I adopted to document the fact under consideration. The use of annual data allows the state variables to have considerably higher predictive power in both the VAR equations and the portfolio predictive regressions. This finding is consistent, for example, with the results in Fama (1990), who argues that the predictive variables contain information that pertains to several months of return data, creating an error in variable problem that is attenuated in annual data.

However, the use of annual data reduces the number of available data points and this creates a problem for the convergence of the estimated parameters. Therefore, I construct the annual data from monthly data, so that two consecutive observations have a three quarter overlap. Each annual observation spans the period up to the end of a quarter. Moreover, I use twenty-five-year estimation windows, so that each regression is estimated using one-hundred overlapping annual data points. The estimation window advances by one data point at a time, which means that one quarter of new data is added to the right

39

of the sample, and one quarter is lost on the left.

At this point a caveat is necessary. The Campbell and Mei (1993) procedure makes an implicit stationarity assumption. The parameters of the VAR, as well as the variance-covariance structure of the portfolio returns are assumed to be constant, not only throughout the estimation period, but, as far as the VAR is concerned, over an infinite horizon. The facts that beta changes over time, and that the estimated parameters of the VAR are not constant, might cast some doubt on the validity of the analysis. However, the assumption of a constant beta can be considered as a descriptive shortcut to look at the average beta over the estimation window, as it was the case in Section 1.2, when the rolling window procedure was first introduced. Moreover, the instability of the VAR should not considerably affect the results as long as the VAR coefficients vary at low frequencies, because in that case discounting should reduce the importance of the terms of the present value formulas that are distant in the future.

The returns of interest are the continuously compounded returns on B/M and size portfolios, from July 1926 to December 2000. For the definition of the vector of state variables I follow the previous literature. As said before, the first two variables have to be the market return and the real interest rate. Therefore I use the return on the CRSP value-weighted index, and the continuously compounded return on the one-month T-bill, deflated by the change in the (log of the) CPI index. Campbell and Mei (1993) also include the aggregate dividend yield, the inflation rate, and the growth rate of industrial production, defined in Section 1.3. I use the dividend yield on the S&P Composite Index, which turns out to have slightly more predictive power than the one constructed from the NYSE universe. The inflation rate is the change in the log of the CPI index. In addition to these variables I include the term spread, which is defined as in Section 1.3 (Table 1.1 provides summary

40

statistics for portfolio returns and the state variables). For this sample, consistent with previous studies, a reasonable value for ρ is 0.96 in annual data and 0.9962 in monthly data.

I can provide some evidence on the increased predictive power allowed by the use of annual data rather than monthly data. In my sample the average R^2 over the ten predictive regressions for B/M portfolios in the whole 1926-2000 period is only 1.7% with monthly data (892 observations per regression). Instead it is considerably higher, 10.6%, with overlapping annual data (291 observations per regression).

The evidence on the sources of the decrease in beta comes from both a split of the main sample into two major subsamples, and a rolling regression analysis.

The analysis by subsamples is presented in Table 1.5 for B/M portfolios. Each beta estimate and the corresponding standard error have been obtained from OLS regressions of the appropriate return component on the market innovation.

In Panel A the estimation period coincides with the entire 1926-2000 sample. The plausible values of market betas in the first row testify that using innovations to define betas, as opposed to returns themselves, does not significantly affect the results. The betas of high decile portfolios (value stocks) are in general higher than those of low B/M deciles. This ranking differs from what reported in previous studies, such as Fama and French (1993), because my sample starts much earlier, in a time when value stocks used to have higher betas. The ranking in the overall betas seems to be determined entirely by the market sensitivity of the excess return news component $\beta_{ie,m}$, which, especially for value stocks, is the most sizeable part.

In Panel B the sample covers the 1926-1962 period. The spread in the betas of value over the beta of glamour stocks is even more pronounced than the one in the overall sample.

41

Again, the responsibility of the difference in the overall beta lies with the excess return betas, rather than the cashflow betas.

The results in Panel C (1963-2000 sample), where the ranking in the overall betas is inverted with respect to Panels A and B, reflect the empirical fact that has inspired this paper. The beta of value stocks has decreased considerably over the years, while the opposite happened for glamour stocks [11]. From the comparison of Panels B and C it is clear that the source of the decrease in the overall beta of value stocks is the decrease in $\beta_{id,m}$, namely the cashflow beta. The rolling regression analysis will make it even more clear that the reduction in cashflow beta plays the dominant role. The beta of real interest rates, which was not significant in Panels A and B, becomes negative and highly significant in Panel C.

In order to further define the source of the decreasing trend in the value stock beta, I now turn to the rolling regression analysis. A first graphic impression of the results is provided by Figure 1-9, that graphs the evolution of the estimated beta of the tenth decile, along with its cashflow component and the negative of its excess return component. It is evident from the picture that the responsibility of the decreasing trend in the overall beta lies with the cashflow beta, which drops from 1.11 in the twenty-five-year estimation window ending in June 1961, to -0.65 in the window ending in December 1995. The excess return beta does not seem to have any apparent trend. This impression is confirmed by the statistical tests that follow.

To quantify the evolution of the different beta components, I fit a trend line through each of the beta series obtained from the rolling regressions. The results are reported in Table 1.7, Panel A. As said above, I use overlapping annual data, and the estimation window advances by one quarter per observation, so that the trend coefficients measure the

---

[11]For market betas of innovations, as for normal betas, it holds that the weighted average of the population betas of the different partitions of the market (e.g., B/M or size portfolios) is equal to one.

change in the beta per quarter. The t-statistics are computed using Newey-West estimator of variance which corrects for the autocorrelation due to the use of overlapping windows in estimating beta. Equation (1.11) implies that the trend in $\beta_{id,m}$, minus the trend in $\beta_{ie,m}$, minus the trend in $\beta_{r,m}$ (-0.001), equals the trend in the overall market $\beta_{i,m}$. The trend line is negatively sloped for the highest B/M deciles, and the slope gradually becomes positive as we approach the portfolios of glamour stocks.

Table 1.7 provides detailed evidence on the sources of the decrease in the overall betas. The cashflow betas of all portfolios trend down, and this effect is stronger for values stocks [12]. However, in the case of glamour stocks this trend does not affect the overall betas, given that the negative estimates of $\beta_{ie,m}$ become progressively bigger in absolute value. Instead, in the case of value stocks the negative trend in the cashflow beta is much stronger (-0.007 per quarter for decile ten), and it is not counterbalanced by a decrease in the excess return beta. Hence, I can conclude that the decrease in cashflow beta is the source of the observed decrease in the overall beta.

The fact that the results presented in this section hinge on the cashflow news component, which is estimated residually, might raise some concerns, which would be justified in case the VAR process was misspecified. To tackle this issue, I follow Campbell and Mei (1993) and form a direct measure of cashflow news by regressing annual log real dividend growth on the state variables, and using the VAR process for the state variables to form revisions in expectations of future dividends (details in the Appendix). Then, I can compute the cashflow beta series in the usual way, and fit a trend line through the series of estimates. The results, not reported, confirm the picture presented above. If anything, the decreasing trend in the cashflow beta for the value stock portfolio is stronger (-0.012, with t-stat. = -

---

[12]Notice that the beta components do not have to satisfy the constraint of adding to one across portfolios. Therefore it is possible that the cashflow betas of all portfolios decrease over time.

3.36). Hence, the use of residual cashflow news does not seem to affect the significance of the evidence presented in this section.

The picture for size portfolios largely resembles that for B/M portfolios, and small stocks play the role of value stocks. Over the entire sample the major drop in the overall beta is again imputable to a drop in the cashflow beta. However, after the beginning of the eighties the high frequency movements in the overall beta depend on the decrease in the excess return beta.

Table 1.6 confirms the decline in the beta of small stocks and the corresponding increase in the beta of large stocks. The market beta of the first decile portfolio, for example, is 1.60 between 1926 and 1962, and it drops to 1.36 in the 1963-2000 subsample.

The big drop in $\beta_{ie,m}$ that appears from the table, is not indicative of the global source of decrease in the overall beta, because it is affected by the last years of data. In fact Figure 1-10 shows that β_{ie} becomes very small in absolute value only around the end of the sample. Instead, the graph indicates that over the entire 1926-2000 period the drop in the cashflow beta is the main source of the decline in the market sensitivity of small stocks. Moreover, it is also evident from Figure 1-10 that after the beginning of the eighties the overall beta follows the path of $\beta_{ie,m}$, which declines in absolute value.

The analysis of the linear trends in the estimates from the rolling regressions procedure confirms that the source of the drop in the overall beta of small stocks is the decrease in the cashflow beta. Panel B of Table 1.7 shows that the overall beta of decile one portfolio, for example, has a linear trend of -0.005, which is entirely imputable to the -0.008 trend in $\beta_{id,m}$. Notice, however, that if the trend is computed using only the betas whose estimation window ends after January 1980, then the -0.004 (t-stat. = -10.96) trend in the overall beta depends on the decrease in the absolute value of the excess return beta (trend in $\beta_{ie,m}$ equal

44

to 0.010, with t-stat. = 8.64). The use of direct, rather than residual, cashflow news, in the way described above, confirms entirely the results presented so far.

So, as in the case of value portfolios, the evidence suggests that the observed instability in the overall market betas of small stocks can be imputed to changes in the sensitivity of cashflow news to market returns.

Overall, the evidence from the beta decomposition is consistent with the conclusions from Section 1.3, which used conditional information to track the evolution of beta. That analysis suggested that the decrease in the beta occurred because the general improvement in the economic conditions has made the activities of these companies less risky. This section showed that, indeed, the decrease in the beta is imputable to the fact that the cashflows of these firms are less sensitive to market news.

1.5   Relation to mispricing

The previous sections have established that the beta of value and small stocks has experienced a major decline over the past sixty years. A question that arises naturally is how this fact relates to the debate surrounding the failure of the CAPM to price correctly these categories of stocks.

Figure 2-1 plots the beta and the intercept (i.e. the alpha) from time-series CAPM regressions for value stocks [13]. A one-tailed t-tests rejects the hypothesis that the intercept is equal to zero at 5% level in most of the estimation windows from the early seventies to the mid-nineties. It is evident from the picture that the occurrence of mispricing starting from the seventies goes side by side with the decrease in beta. Alpha started to rise in the early

---

[13]The series are produced using rolling regressions with a ten-year estimation window. I extended the estimation window from five to ten years to obtain smoother series, and to have more power in the t-tests on the intercept.

sixties, when beta was experiencing a major drop, and its growth is always accompanied by a decline in beta. The correlation between the two series is about -50%. The graph for small stocks (not reported) also shows that alpha rises when beta declines, although for this portfolio the mispricing disappears in the late eighties.

The decrease in beta can be related to mispricing in at least two ways. First, if the correct pricing model is a conditional CAPM, and the estimated model is an unconditional CAPM, the premium estimated by the econometrician contains a bias due to correlated omitted variables. Secondly, the drastic changes in beta can have caused investors to formulate the wrong expectations on the riskiness of these stocks. The next subsections examine these two explanations.

1.5.1   The omitted variable bias

The time-series tests of the CAPM, like the ones in Fama and French (1993), which rely on the significance of the intercept to decide if the market value-weighted portfolio is on the mean-variance efficient frontier, fail to incorporate the time variation in beta. So, for example, Fama and French (1993, Table 9a) find that the portfolio of small high B/M stocks has a monthly premium of 0.54% (t-stat. = 2.53) in the 1963-1991 sample.

However, even if the CAPM holds conditionally, it does not necessarily hold uncondi-tionally. For it to be the case the relevant moments of the joint conditional distribution of returns would have to be constant over time, or change proportionally. The evidence of a decreasing beta does not depose in favor of this possibility. In fact, the analysis of Section 1.3 showed that beta can be closely tracked by a number of state variables sum-marizing the state of the economy. Hence, there are reasons to believe that part of the premium found for these categories of stocks in the time-series tests can be explained by

46

time variation in beta, in the form of a correlated omitted variable bias.

Suppose one estimates the following unconditional CAPM regression

$$R_{i,t+1} = \alpha_i + \beta_i R_{m,t+1} + u_{i,t+1}, \tag{1.12}$$

but the correct model is a conditional CAPM

$$E_t R_{i,t+1} = \beta_{i,t} E_t R_{m,t+1}, \tag{1.13}$$

where the conditional beta is a linear function of some state variables, as in Equation (1.6). Then, solely because of the omission of the time variation in beta from the unconditional regression, the estimated intercept turns out to be different from zero. In particular, the probability limit of $\widehat{\alpha}_i$ is

$$\text{Plim } \widehat{\alpha}_i = b'_{1,i}(E\widetilde{z} - \gamma ER_m), \tag{1.14}$$

where $\widetilde{z} = z_t R_{m,t+1}$, and γ is the linear projection coefficient of $\widetilde{z}_t$ on the market return $R_{m,t+1}$. So, when beta is a non-trivial function of the state variables, and the term in parenthesis is also not zero [14], part of the premium in the time-series unconditional CAPM regressions can be explained on the basis of Equation (1.14).

From Panel C of Tables 1.3 and 1.4 it appears that the tenth B/M portfolio and the first size portfolios are those that in the 1963-2000 subsample have the highest absolute values of the coefficients in the vector $b_{1,i}$. This fact makes it more likely for the estimated intercept from an unconditional CAPM regression to be different from zero, as it can be seen from Equation (1.14).

---

[14]Notice that the term in parentheses in Equation (1.14) is trivially zero if the state variables are constant over time. It is also zero if the market return is independent of the state variables.

In order to assess what part of the premium of value and small stocks is accounted for by the omission of the time variation in beta, one can compute the sample equivalent of the expression in Equation (1.14) and compare it with the intercept from the regression in Equation (1.12). An equivalent, and more simple, way to do that is estimating Equation (1.7), and comparing the intercept from that regression with the intercept from the unconditional regression in Equation (1.12). In the case of value stocks, about 30% of the 0.44% monthly premium in the 1963-2000 sample can be accounted for when beta is allowed to vary. For small stocks this share is 66%, but the premium in the 1963-2000 sample is just 0.10%, and it is not significant. In shorter samples, like the 1963-1980 one, this procedure does not seem to account for any sizeable part of the premium to small stocks.

Besides shedding some light on the sources of the value premium, the results from this analysis suggest that the conditional version of the CAPM should be preferred to the unconditional one in most applications. The time variation in portfolio betas and the ability to track it with state variables, which have been documented in this paper, make the stationarity assumption behind the unconditional CAPM not realistic.

1.5.2 Exaggerated perception of risk

Daniel and Titman (1997) argue that characteristics, as opposed to covariation with risk factors, generate the observed premium of small and value stocks. In their view, the market dislikes these categories of stocks, so that a premium is required for investors to hold them. This negative attitude towards small and value stocks might depend on the fact that investors overestimated their systematic risk. As more powerful computing resources become publicly available, these anomalies should disappear. The size effect actually disappeared in the early 80's, as well as the B/M effect in the late nineties.

Consistent with this argument, I argue that the large decline that the beta of value and small stocks experienced starting from the early forties can be the reason why investors made mistakes in the assessment of risk. The observation that these assets used to bear a great deal of market risk in the years of the Great Depression and World War II can have convinced investors that the value and small characteristics were associated with higher risk. Consequently, these stocks had to pay a premium, even when the amount of systematic risk they were bearing had declined.

Suppose the return that the market expects in the next period for the portfolios of interest is based on the unconditional CAPM

$$ER_{i,t+1} = \beta_i^e ER_{m,t+1}, \tag{1.15}$$

where β_i^e is the expectation of beta. This assumption implies that the probability limit of the estimated intercept in a time-series unconditional CAPM regression is

$$\alpha_i = (\beta_i^e - \beta_i)ER_{m,t+1}. \tag{1.16}$$

To keep things simple, suppose for now that the expectation of beta is identically equal to the value of beta estimated from all past return realizations. More complicated setups in which the market learns from the path of realized betas will be discussed later.

From Equation (1.16) it is evident that whenever the expectation of beta exceeds the true beta, the portfolio pays a premium relative to the CAPM, and the estimate of alpha from the time-series regression tends to a positive value. Hence, in the case of value and small stocks, the decreasing path of estimated betas and the assumption of adaptive expectations can actually explain part of the premium.

To implement the model empirically, I split the sample of realized returns in two sub-samples, from 1926 to 1962 and from 1963 to 2000. The second subsample more or less coincides with the time period when the CAPM was known to financial markets. Further, I assume that the market expected return for the portfolio of interest in the second subsample is formed according to Equation (1.15), and the expectation of beta is equal to the estimate of beta from the first subsample. Finally, I assume that these expectations are not revised until the end of the second subsample. Although these assumptions are obviously unrealistic, they capture the idea that the high level of market risk born by value and small stocks in the years of the Great Depression and World War II affected the market expectations of how risky these stocks would be later on.

By replacing sample estimates in Equation (1.16), I can compute the fraction of the estimated intercept in the CAPM regression in the 1963-2000 subsample that is explained by the model above. For value stocks the intercept is 0.44% (t-stat. = 2.9) between January 1963 and December 2000, while beta is 1.64 in the first subsample, and 0.97 in the second subsample. The product of the difference in these betas and the mean excess market return between 1963 and 2000 (0.52% monthly) is equal to 0.35%, which is the sample equivalent of the expression in Equation (1.16). Hence, under these assumptions, the misperception of the riskiness of value stocks can have caused about 80% of the premium that they paid in the 1963-2000 subsample.

Small stocks did not have much of a premium in the whole 1963-2000 subsample (only 0.10%). Therefore, the above method would imply an intercept that is almost three times as big as the actual intercept. Instead, when I restrict the second subsample to end in 1980, which was approximately when the size effect disappeared, the estimated premium is 0.59% (t-stat. = 2.1). Beta in the 1926-1962 subsample is 1.65, and in the 1963-1980 subsample

50

is 1.29, while in the same period the mean monthly market excess return is 0.34%. Hence, misperception of risk, as defined above, generates a 0.12% monthly premium, which is about 20% of the realized premium for small stocks.

More complicated setups than the one described above can be thought, where the market revises its adaptive expectations on the basis of realized returns. Suppose, for example, that investors form their expectations of beta by combining the corresponding expectation in the previous period and the last estimate of beta, as in the following equation

$$\beta_t^e = \delta\widehat{\beta}_t + (1 - \delta)\beta_{t-1}^e, \tag{1.17}$$

where β_t^e is the beta expected to apply to time $t+1$, $\widehat{\beta}_t$ is the estimated beta at time t, and δ is a weighting coefficient.

To make this model operational, I insert it in a ten-year rolling regression framework, starting from 1963. So, $\widehat{\beta}_t$ corresponds to the estimate of beta in the last ten-year window. In some estimation windows the mean value of the market excess return is negative and I cannot use it as an estimate of the market premium, because it would be inconsistent with CAPM. Hence, I estimate $E_t R_{m,t+1}$ using the mean excess market return from the beginning of month $t+1$ to December 2000. The estimates obtained with this procedure are consistently positive. So, for each window I can compute an estimate of the α in Equation (1.16) using the expectation of beta from Equation (1.17), the estimate of beta in the window, and the estimate of the market excess return. I calibrate δ to minimize the sum of the squared differences between the estimates of α and the estimated intercepts in the CAPM rolling-regressions. Finally, I set the initial condition for β_t^e to equal the estimate of beta in the 1926-1962 subsample.

When this model is applied to the value portfolio, the optimal δ is around 0.002. This small number is consistent with the idea that one additional month of data is not very informative about beta. The distribution of the share of the premia in the CAPM rolling regressions explained by the estimates of α has mean equal to 104%, median equal to 70%, maximum equal to 487%, minimum equal to 32%, and standard deviation equal to 94%.

Again in the case of small stocks I restrict the sample to the years up to 1980. Further, I consider only the estimation windows for which the intercept in the CAPM regressions is positive, namely the ones in which there is actually a premium. The optimal value of δ is then 0.06. For these windows, the distribution of the share of the premia that is explained by the estimate of α has mean equal to 75%, median equal to 66%, maximum equal to 331%, minimum equal to 22%, and standard deviation equal to 54%.

Notice that these exercises do not account for the effect on returns of the revision in the expectation of future discount rates. Investors' realization that beta is lower than they thought causes a positive surprise in returns, because future payoffs are discounted at a lower rate. To the extent that these surprises are correlated across periods, this effect shows up as a positive premium in the CAPM regression. Therefore, the distributions of premia that were computed above represent a lower bound on the premia that can be generated under the assumption that investors slowly learn about beta. A non-arbitrary assessment of the impact of the revision in discount rates on the premium can be provided within an equilibrium model of learning on beta, which is developed in Adrian and Franzoni (2002).

Although the assumptions underlying the above models are certainly restrictive, the results presented in this subsection suggest that misperception of risk based on the higher values of beta characterizing value and small stocks in the early years of the sample can play a big role in explaining the premia of these portfolios. This hypothesis is pursued in

further research (Adrian and Franzoni, 2002), which develops a formal models of learning on beta, and assesses the relevance of the decrease in beta for the estimated premium of value and small stocks.

1.6 Conclusions

This paper found a striking decrease in the market beta for portfolios of value and small stocks. In the course of sixty years the beta of each of these portfolios decreased by about 75%. The maximum estimated beta for value stocks was 2.2 in the early forties, and the minimum was 0.55 in the late nineties. The situation for small stocks evolved similarly.

The fact that the path of beta can be closely tracked using state variables that are related to business conditions, such as the nominal interest rate, the default spread, the term spread, the dividend yield, and the growth rate of industrial production, suggested the interpretation that a change occurred in the structure of these companies' cashflows. According to this argument, the improvement of business conditions reduced payoff uncertainty for all firms, and especially for value and small companies, which are believed to be more prone to distress, and therefore more sensitive to the status of the economy. Support for this intuition came from the result that the source of the decrease in the beta is the decline in the sensitivity of cashflow news to market news. This interpretation generates the out-of-sample prediction that we should observe a rise in the beta of these portfolios in conjunction we the recent economic downturn and the interest rate cuts.

This evidence is relevant for the debate on the premium to value and small stocks in the CAPM regressions. The amount of market risk born by these stocks went down contemporaneously to the appearance of the premium. The paper found some evidence that a part of the premium, which can be as high as 80% for value stocks, originates from

investor beliefs that a high level of market risk would characterize these stocks even when, due to the decrease in beta, this was no longer the case. This argument is in line with some behavioral explanations of the anomaly (e.g., Daniel and Titman, 1997) that suggest that the market may have misperceived the riskiness of these portfolios.

Finally, the paper provided further evidence in favor of the adoption of a conditional version of the CAPM in place of the unconditional one. The sizeable changes in the beta of these portfolios, the close relationship between these changes and the evolution of some state variable, the ability to explain 30% of the value premium once the time variation of beta is taken into account, are all elements that motivate a model that lets beta depend on some conditioning information.

These findings can have applications in the areas of portfolio selection and performance evaluation. Mutual funds that adopt value or small cap strategies should know that the amount of market risk born by their portfolios is subject to variation over time. For example, the fund manager should expect value and small stock portfolios to become more risky during recessions. The good news is that this evolution can be predicted on the basis of information that is easily available, and using a simple linear relationship. Also, when evaluating performance of mutual fund managers, one needs to take into account the fact that the manager can predict changes in the portfolio beta, especially for these categories of stocks. Hence, the manager should not be rewarded for portfolio returns that can be anticipated using conditioning information. Ferson and Schadt (1996) develop this idea by assessing what part of mutual fund performance is imputable to changes in beta that were predictable using a conditional CAPM.

Further research should extend different aspects of this paper. The interpretation of the link between beta and the state variables was not tested directly, and it was not supported

by some economic model. Possibly, one should look into the evolution of the accounting and financial figures of the companies in the portfolios of interest, and assess whether they conform with the change-in-distress interpretation that was provided in the paper. Preliminary evidence suggests that a measure of distress, such as the ratio of interest expenses to operating income, is not only higher in the case of value and small companies, but it is also more positively related to recessions. Moreover, the gap between the level of this ratio for the companies of interest and for the rest of the market has been shrinking, consistent with the negative trend in the beta.

Another promising direction of research concerns the implications of the decrease in beta for the misperception of the riskiness of these stocks, and the premium that they pay. More sophisticated models of learning on beta than the ones used in the paper can be developed to fit these data. In order to feature a slow learning process, necessary to generate the premium, these models can either assume that investors learn about the default risk of these stocks from infrequent events like recessions, or they can be founded on some psychological bias in probability assessment and expectation formation.

Appendix

Obtaining the news components of returns

This Appendix describes how the different components of excess return innovations in Equation (1.9) are obtained. In this I follow, with some minor modifications, Campbell and Mei (1993), who assume that expectations of portfolio excess returns are linear in a vector x_t of state variables $x_{t,l}$, $l = 1, ..., L$, which represent investor information set. The first element of this vector is the excess return on the market, and the second one is the real return on a one-month T-bill, while the other elements are variables known at the end of period t. Thus, the excess return on portfolio i can be written as

$$e_{i,t+1} = a_i' x_t + \tilde{e}_{i,t+1} \tag{1.18}$$

for some portfolio specific L-element column vector of coefficients. The expected return on the portfolio is thus $a_i' x_t$, while the unexpected return is $\tilde{e}_{i,t+1} = e_{i,t+1} - a_i' x_t$.

Next it is assumed that the state vector follows a first-order VAR process

$$x_{t+1} = \Pi x_t + \tilde{x}_{t+1}, \tag{1.19}$$

where the $\tilde{x}_{t+1}$ denotes the innovation in the state variables. The assumption of a first-order VAR is not restrictive since a higher-order VAR can always be rewritten in first-order form (see, e.g., Campbell and Shiller, 1988). The $L \times L$ matrix Π is known as the companion matrix of the VAR. Given the VAR model, revisions in long-horizon expectations of x_{t+1} are

$$(E_{t+1} - E_t)x_{t+1+j} = \Pi^j \tilde{x}_{t+1}. \tag{1.20}$$

Further, I define the L-elements vectors ι_1 and ι_2 to have a one in first and second position, respectively, and zeros in all other positions. These vectors pick the excess return on the market and the real interest rate out of the VAR. Thus, Equation (1.20) and the definitions of $\tilde{e}_{di,t+1}$ and $\tilde{e}_{ei,t+1}$ in (1.9) imply that the components of portfolio i return and of the market return can be written as follows

$$
\begin{aligned}
\bar{e}_{ei,t+1} &= \rho a_i'(I - \rho\Pi)^{-1}\tilde{x}_{t+1}, \\
\tilde{e}_{r,t+1} &= \iota_2'(I - \rho\Pi)^{-1}\tilde{x}_{t+1}, \\
\tilde{e}_{di,t+1} &= \tilde{e}_{i,t+1} + (\iota_2' + \rho a_i')(I - \rho\Pi)^{-1}\tilde{x}_{t+1}, \\
\tilde{e}_{em,t+1} &= \rho\iota_1'\Pi(I - \rho\Pi)^{-1}\tilde{x}_{t+1}, \\
\tilde{e}_{dm,t+1} &= \tilde{e}_{m,t+1} + (\iota_2' + \rho\iota_1'\Pi)(I - \rho\Pi)^{-1}\tilde{x}_{t+1}.
\end{aligned}
\tag{1.21}
$$

The discounted sum of the revisions in expectations for the vector of state variables at $t+1$ is $(I - \rho\Pi)^{-1}\tilde{x}_{t+1}$, where I is the $L \times L$ identity matrix. This term is translated into revisions of forecast portfolio returns through the vector a_i, which links the state variables to expected returns. Finally, the revision in expected cashflows is determined residually from Equation (1.9).

The cashflow news component can also be computed directly under the assumption that the state variables predict dividends as well. In particular, if c is the vector of regression coefficients of dividend growth on the state variables, and μ_{t+1} are the residuals from this regression, the direct cashflow news is

$$
\tilde{e}_{id,t+1} = \mu_{t+1} + \rho c'(I - \rho\Pi)^{-1}\tilde{x}_{t+1}.
\tag{1.22}
$$

57

Table 1.1: Summary statistics. Panel A reports means and standard deviations of selected portfolios returns, and the average share of market capitalization for these portfolios Panel A also reports the correlation coefficient computed over all the stocks and years between the size and B/M decile assignments. Panel B reports mean, standard deviations, and sample autocorrelation coefficients (ρ) for: the one-month T-bill rate, the dividend yield on the S&P Composite Index, the growth rate of industrial production, the term spread, and the default spread. Panel C reports the correlation matrix for the state variables.

Panel A: Portfolio Returns			
	Mean	St. Dev.	Avg. mkt share
Sample: 1926-2000			
BM decile 1	.93	5.82	.26
BM decile 10	1.37	9.35	.02
Size decile 1	1.44	10.49	.01
Size decile 10	.96	5.18	.62
Market	.99	5.49	
Avg. corr. of size and BM deciles	-.33		
Sample: 1926-1962			
BM decile 1	.91	6.42	.27
BM decile 10	1.28	12.22	.02
Size decile 1	1.71	13.48	.002
Size decile 10	.91	6.03	.67
Market	.95	6.46	
Avg. corr. of size and BM deciles	-.47		
Sample: 1963-2000			
BM decile 1	.95	5.18	.24
BM decile 10	1.46	5.30	.03
Size decile 1	1.19	6.43	.02
Size decile 10	1.01	4.20	.57
Market	1.03	4.37	
Avg. corr. of size and BM deciles	-.19		

Panel B: State Variables								
	Mean	Std. Dev.	ρ_1	ρ_2	ρ_3	ρ_4	ρ_{12}	ρ_{24}
T-bill	.31	.25	.97	.95	.94	.93	.87	.78
Div. Yield	.042	.015	.98	.96	.93	.91	.73	.56
Growth I.P.	.003	.019	.53	.24	.08	-.02	-.15	-.11
Term spr.	2.11	1.36	.84	.77	.74	.69	.47	.21
Default spr.	1.13	.73	.97	.94	.91	.90	.75	.55

Panel C: Correlation matrix of State Variables					
	T-bill	Term	Def.	DY	Gr. I.P.
T-bill	1				
Term	-0.39	1			
Def.	-0.07	0.41	1		
DY	-0.29	0.09	0.56	1	
Gr. I.P.	-0.07	0.05	-0.11	-0.11	1

Table 1.2: Linear trends in the estimated betas. The table reports OLS estimates from the model $\log \beta_t = \gamma_0 + \gamma_1 trend_t + \varepsilon_t$. The variable β_t is the series of estimates from five-year rolling regressions in which the dependent variable is the portfolio monthly excess return, and the independent variable is the market excess return. The variable $trend_t$ is a time trend. The estimation sample refers to the window over which betas are estimated. The data are at monthly frequency, and the first month is July in both the the 1926-2000 and the 1963-2000 samples. The T-statistic (in paretheses) is computed using the Newey-West estimator with 59 lags of autocorrelation.

Decile	1	2	3	4	5	6	7	8	9	10
Panel A: B/M Portfolios										
				Sample: 1926-2000						
$\gamma_1(\times 100)$	0.029	0.015	0.023	-0.009	-0.006	-0.034	-0.056	-0.068	-0.091	-0.103
	(5.67)	(3.15)	(4.55)	(-1.10)	(-0.76)	(-4.45)	(-6.23)	(-8.71)	(-7.02)	(-7.58)
				Sample: 1963-2000						
$\gamma_1(\times 100)$	0.034	0.004	0.027	-0.004	0.010	0.003	-0.022	-0.067	-0.092	-0.130
	(3.56)	(0.46)	(1.52)	(-0.27)	(0.64)	(0.24)	(-1.16)	(-4.66)	(-3.11)	(-4.11)
Panel B: Size Portfolios										
				Sample: 1926-2000						
$\gamma_1(\times 100)$	-0.086	-0.048	-0.030	-0.020	-0.023	-0.022	-0.018	-0.006	-0.010	0.003
	(-6.10)	(-3.83)	(-3.27)	(-2.13)	(-2.57)	(-3.26)	(-2.72)	(-1.59)	(-1.86)	(1.04)
				Sample: 1963-2000						
$\gamma_1(\times 100)$	-0.166	-0.093	-0.076	-0.067	-0.037	-0.048	-0.046	-0.022	0.007	0.032
	(-8.47)	(-5.93)	(-4.99)	(-4.48)	(-3.17)	(-3.73)	(-7.65)	(-3.06)	(0.83)	(8.64)

Table 1.3: Conditioning information (B/M portfolios). The table reports the coefficients $b_{0,i}$ and $b_{1,i}$ from the estimation of Equation (1.7) in the text: $R_{i,t+1} = \alpha_i + b_{0,i}R_{m,t+1} + (b'_{1,i}z_t)R_{m,t+1} + u_{i,t+1}$. The dependendt variables are the excess returns on B/M deciles portfolios. The z_t variables are the monthly T-bill rate (t_bill), the dividend yield (div_y), the growth rate of industrial production (gr_ip), the term spread ($term$), the default spread (def). The data are at monthly frequency from 7/1926 to 12/2000. T-statistics in parentheses.

Decile	1	4	7	10
Panel A: 1926-2000				
b_0	0.97	0.95	1.12	1.56
	-(47.52)	-(45.88)	-(42.48)	-(30.07)
t_bill	0.22	0.08	-0.41	-0.97
	-(4.36)	-(1.58)	(-6.08)	(-7.27)
div_y	-0.61	0.66	1.52	4.73
	(-0.74)	-(0.79)	-(1.42)	-(2.23)
gr_ip	-1.01	0.39	0	3.36
	(-2.68)	-(1.03)	(0.00)	-(3.49)
$term$	0.01	0.006	0.11	-0.04
	-(1.02)	-(0.57)	(-0.48)	(-1.58)
def	-0.02	0.08	0.09	0.1
	(-1.61)	-(4.74)	-(5.01)	-(2.18)
			continued on next page	

Decile	1	4	7	10
	Panel B: 1926-1962			
b_0	0.89	0.93	1.12	1.94
	(28.89)	(25.87)	(21.98)	(16.90)
t_bill	1.01	0.22	-0.46	-4.38
	(4.24)	(0.81)	(-1.19)	(-4.93)
div_y	0.02	0.49	0.09	-0.51
	(0.03)	(0.53)	(0.08)	(-0.17)
gr_ip	-1.01	0.53	-0.14	3.37
	(-3.12)	(1.40)	(-0.27)	(2.78)
$term$	0.04	0.01	0.08	-0.35
	(1.32)	(0.45)	(1.64)	(-3.11)
def	-0.04	0.08	0.07	0.31
	(-1.89)	(3.15)	(1.96)	(3.54)
	Panel C: 1963-2000			
b_0	1.01	0.85	1.44	2.09
	(8.05)	(7.18)	(11.27)	(11.11)
t_bill	0.15	0.24	-0.83	-1.58
	(0.82)	(1.32)	(-4.20)	(-5.43)
div_y	-0.54	1.63	10.06	22.31
	(-0.22)	(0.70)	(3.95)	(5.96)
gr_ip	2.76	-4.39	-3.17	-7.22
	(1.12)	(-1.89)	(-1.26)	(-1.95)
$term$	0.02	0.02	-0.05	-0.07 ·
	(1.32)	(1.54)	(-2.84)	(-2.42)
def	0.01	-0.03	0.12	0.06
	(0.22)	(-0.47)	(1.59)	(0.57)

Table 1.4: Conditioning information (Size portfolios). The table reports the coefficients $b_{0,i}$ and $b_{1,i}$ from the estimation of Equation (1.7) in the text: $R_{i,t+1} = \alpha_i + b_{0,i}R_{m,t+1} + (b'_{1,i}z_t)R_{m,t+1} + u_{i,t+1}$. The dependendt variables are the excess returns on size deciles portfolios. The z_t variables are the monthly T-bill rate (t_bill), the dividend yield (div_y), the growth rate of industrial production (gr_ip), the term spread ($term$), the default spread (def). The data are at monthly frequency from 7/1926 to 12/2000. T-statistics in parentheses.

Decile	1	4	7	10
Panel A: 1926-2000				
b_0	1.53	1.24	1.17	0.94
	(21.99)	(35.71)	(58.29)	(101.95)
t_bill	-0.80	-0.23	-0.18	0.00
	(-4.47)	(-2.62)	(-3.47)	(-0.04)
div_y	-0.41	-5.07	-3.60	0.43
	(-0.14)	(-3.55)	(-4.35)	(1.12)
gr_ip	8.15	3.46	0.75	-0.21
	(6.32)	(5.40)	(2.03)	(-1.20)
$term$	-0.07	-0.01	-0.01	0.00
	(-1.83)	(-0.34)	(-1.07)	(-0.92)
def	0.26	0.17	0.10	-0.01
	(4.24)	(5.62)	(5.40)	(-1.80)

Decile	1	4	7	10
		Panel B: 1926-1962		
b_0	1.88	1.45	1.33	0.92
	(11.93)	(20.92)	(33.72)	(74.10)
t_bill	-4.31	-2.57	-1.74	0.29
	(-3.54)	(-4.80)	(-5.68)	(3.01)
div_y	-4.24	-7.96	-5.70	0.94
	(-1.04)	(-4.44)	(-5.59)	(2.96)
gr_ip	8.29	3.59	0.79	-0.22
	(5.01)	(4.90)	(1.89)	(-1.72)
term	-0.37	-0.19	-0.13	0.00
	(-2.39)	(-2.78)	(-3.34)	(0.23)
def	0.46	0.30	0.18	-0.02
	(3.84)	(5.73)	(5.81)	(-2.51)
		Panel C: 1963-2000		
b_0	1.77	1.55	1.27	0.89
	(6.62)	(9.14)	(13.06)	(13.77)
t_bill	-1.07	-0.61	-0.28	0.06
	(-2.59)	(-2.33)	(-1.85)	(0.60)
div_y	6.25	5.15	2.20	-1.25
	(1.18)	(1.53)	(1.14)	(-0.97)
gr_ip	-4.05	-6.69	-6.62	2.34
	(-0.77)	(-2.01)	(-3.46)	(1.84)
term	-0.10	-0.03	-0.02	0.00
	(-2.29)	(-1.22)	(-1.36)	(0.40)
def	0.15	0.00	0.05	0.02
	(0.94)	(0.02)	(0.76)	(0.42)

Table 1.5: Beta decomposition (B/M portfolios). Estimates from the following decomposition of beta for B/M decile portfolios: $\beta_{i,m} = \beta_{di,m} - \beta_{r,m} - \beta_{ei,m}$. $\beta_{i,m}$ is the portfolio beta computed using portfolio and market excess return innovations. $\beta_{id,m}$ is the dividend news component of beta. $\beta_{ie,m}$ is the expected return news component of beta. $\beta_{r,m}$ is the real interest rate news component of beta. T-statistics in parentheses.

Decile	1	4	7	10
	Panel A: 1926-2000			
$\beta_{i,m}$	1.04	1.06	1.10	1.29
	(52.06)	(55.72)	(40.61)	(27.93)
$\beta_{id,m}$	0.56	0.17	0.18	-0.10
	(21.21)	(4.36)	(4.01)	(-1.27)
$\beta_{ie,m}$	-0.49	-0.89	-0.93	-1.39
	(-17.17)	(-17.68)	(-19.63)	(-20.69)
$\beta_{r,m}$	0.01			
	(0.39)			
	Panel B: 1926-1962			
$\beta_{i,m}$	0.95	1.14	1.20	1.44
	(46.14)	(51.50)	(33.93)	(22.65)
$\beta_{id,m}$	0.37	0.57	0.43	0.21
	(8.00)	(15.17)	(7.59)	(2.13)
$\beta_{ie,m}$	-0.64	-0.63	-0.83	-1.28
	(-10.61)	(-10.38)	(-12.33)	(-13.89)
$\beta_{r,m}$	0.06			
	(1.74)			
	Panel C: 1963-2000			
$\beta_{i,m}$	1.15	0.98	0.89	0.85
	(27.50)	(27.35)	(20.29)	(14.38)
$\beta_{id,m}$	1.15	-0.80	-0.32	-0.92
	(12.59)	(-6.98)	(-4.10)	(-7.46)
$\beta_{ie,m}$	0.28	-1.50	-0.94	-1.49
	(3.47)	(-14.81)	(-18.15)	(-16.81)
$\beta_{r,m}$	-0.28			
	(-16.68)			

Table 1.6: Beta decomposition (Size portfolios). Estimates from the following decomposition of beta for Size decile portfolios: $\beta_{i,m} = \beta_{di,m} - \beta_{r,m} - \beta_{ei,m}$. $\beta_{i,m}$ is the portfolio beta computed using portfolio and market excess return innovations. $\beta_{id,m}$ is the dividend news component of beta. $\beta_{ie,m}$ is the expected return news component of beta. $\beta_{r,m}$ is the real interest rate news component of beta. T-statistics in parentheses.

Decile	1	4	7	10
Panel A: 1926-2000				
$\beta_{i,m}$	1.50	1.23	1.14	0.94 .
	(29.05)	(44.07)	(64.74)	(105.13)
$\beta_{id,m}$	0.26	0.14	0.23	0.40
	(3.47)	(2.59)	(5.94)	(18.47)
$\beta_{ie,m}$	-1.25	-1.10	-0.92	-0.55
	(-20.25)	(-20.20)	(-20.30)	(-17.61)
$\beta_{r,m}$	0.01			
	(0.39)			
Panel B: 1926-1962				
$\beta_{i,m}$	1.60	1.27	1.14	0.93
	(23.69)	(39.67)	(54.18)	(136.41)
$\beta_{id,m}$	0.48	0.45	0.41	0.45
	(4.57)	(6.96)	(8.12)	(13.69)
$\beta_{ie,m}$	-1.18	-0.88	-0.80	-0.54
	(-12.59)	(-12.39)	(-12.60)	(-10.32)
$\beta_{r,m}$	0.06			
	(1.74)			
Panel C: 1963-2000				
$\beta_{i,m}$	1.36	1.23	1.14	0.93
	(15.09)	(22.58)	(34.31)	(44.62)
$\beta_{id,m}$	0.39	-0.33	0.06	0.08
	(3.48)	(-2.89)	(0.94)	(2.69)
$\beta_{ie,m}$	-0.69	-1.28	-0.80	-0.57
	(-18.44)	(-17.22)	(-26.07)	(-31.46)
$\beta_{r,m}$	-0.28			
	(-16.68)			

Table 1.7: Linear trends in beta and its components. The table reports OLS estimates on the time trend from the model $\beta_t = \gamma_0 + \gamma_1 trend_t + \varepsilon_t$. The series β_t comes from twentyfive-year rolling regressions. β_t is in turn: the overall portfolio beta obtained from excess return innovations ($\beta_{i,m}$), the dividend news component of beta ($\beta_{id,m}$), the expected return news component of beta ($\beta_{ie,m}$), the real interest rate news component of beta ($\beta_{r,m}$). The T-statistic (in paretheses) is computed using the Newey-West estimator with 99 lags of autocorrelation.

$Decile$	1	4	7	10
Panel A: B/M Portfolios				
$\beta_{i,m}$	0.002	0.000	-0.002	-0.006
	(7.40)	(-0.12)	(-8.39)	(-14.26)
$\beta_{id,m}$	0.000	-0.004	-0.003	-0.007
	(-0.22)	(-7.63)	(-4.72)	(-12.56)
$\beta_{ie,m}$	-0.001	-0.003	0.000	0.000
	(-0.85)	(-4.03)	(-0.04)	(0.34)
Panel B: Size Portfolios				
$\beta_{i,m}$	-0.005	-0.002	-0.001	0.0003
	(-14.83)	(-10.20)	(-4.24)	(2.34)
$\beta_{id,m}$	-0.009	-0.007	-0.004	-0.003
	(-6.65)	(-9.39)	(-8.29)	(-3.55)
$\beta_{ie,m}$	-0.002	-0.004	-0.002	-0.002
	(-1.97)	(-4.08)	(-2.82)	(-1.77)
trend in β_r = -.001 (-2.68)				

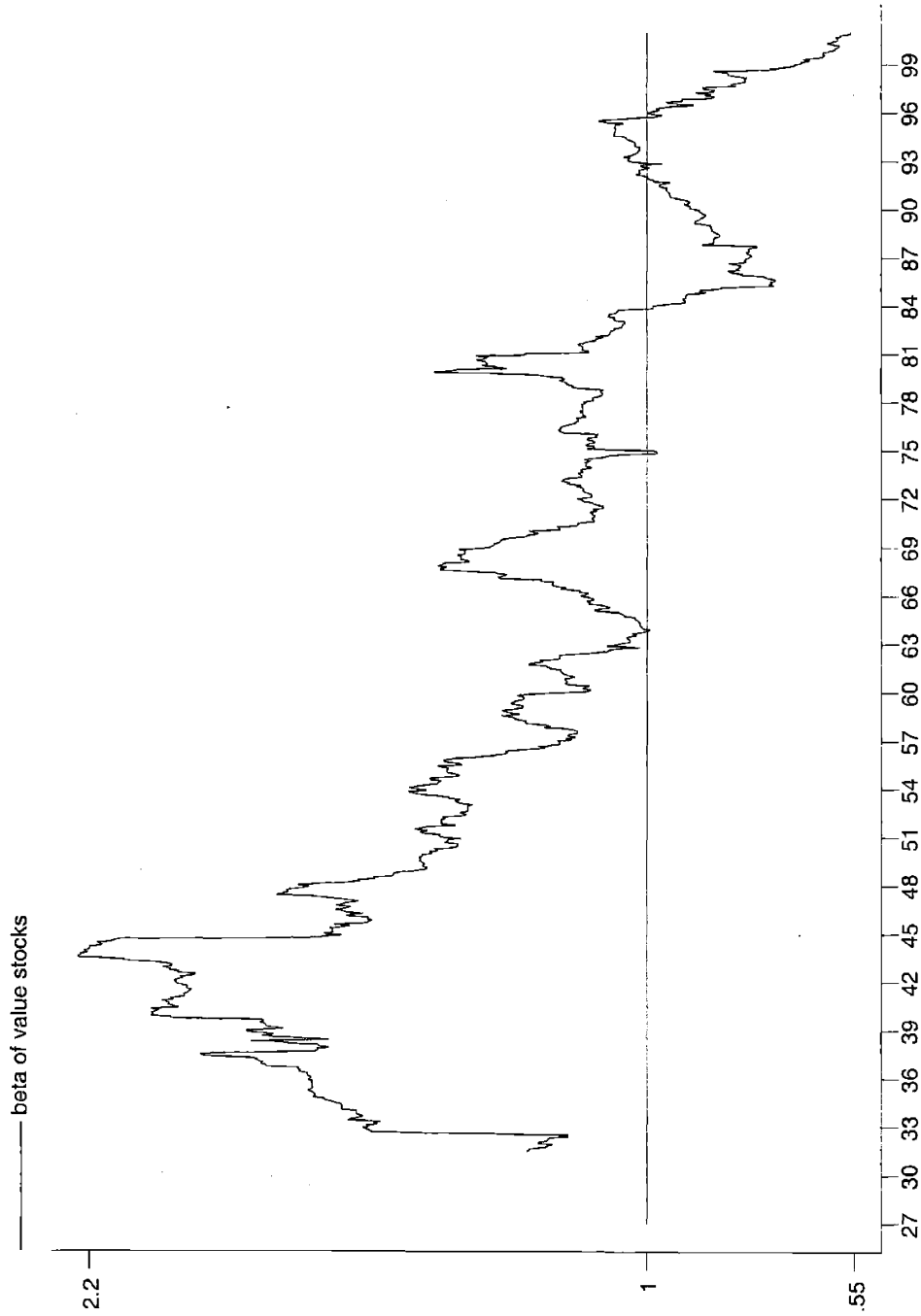**Figure 1-1: Beta of value stocks.** The figure plots the estimates from time-series regressions. The series originates from 5-year rolling window regressions where the dependent variable is the the return on B/M decile 10 portfolio, and the independent variable is the excess return on the market value-weighted portfolio.
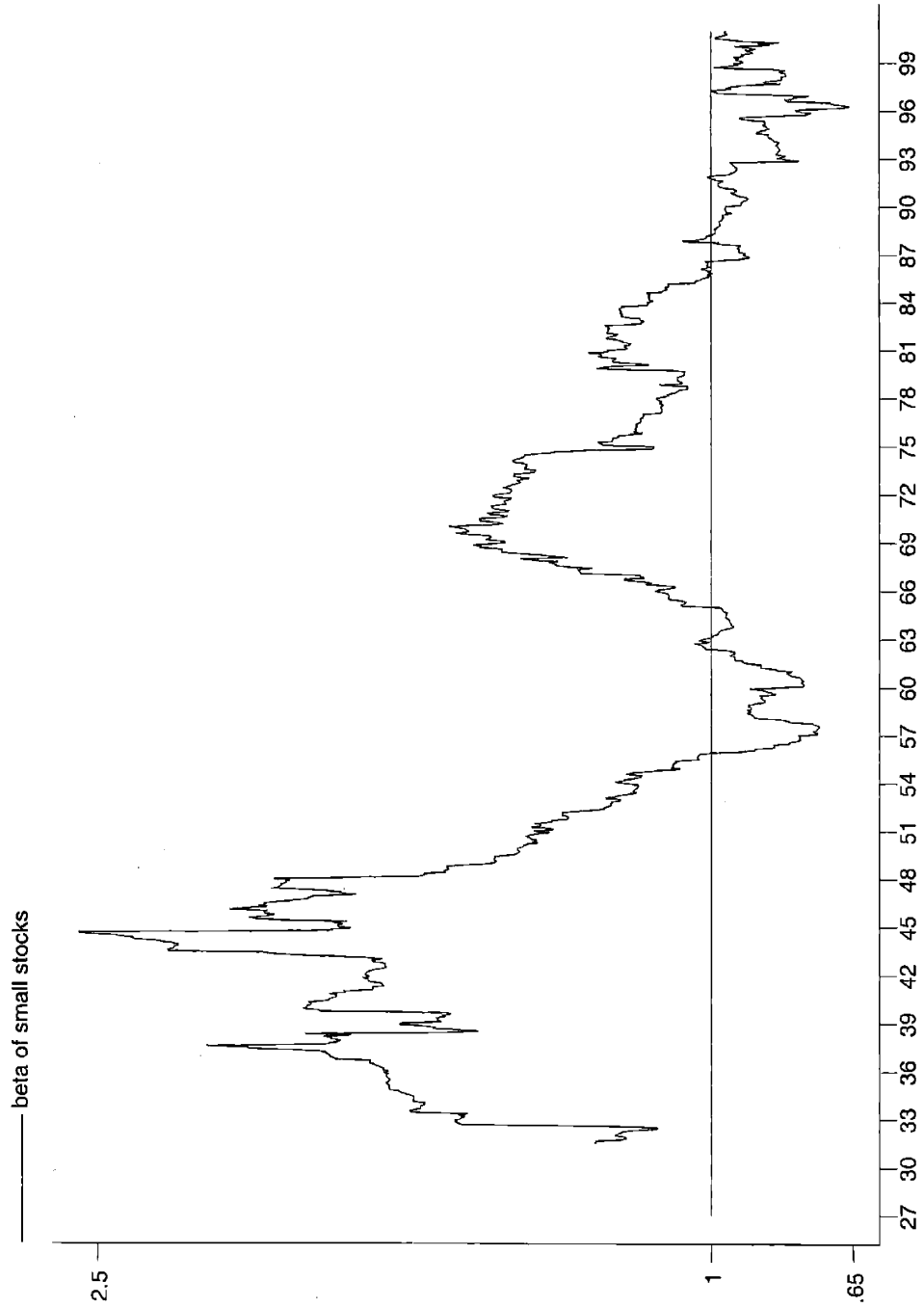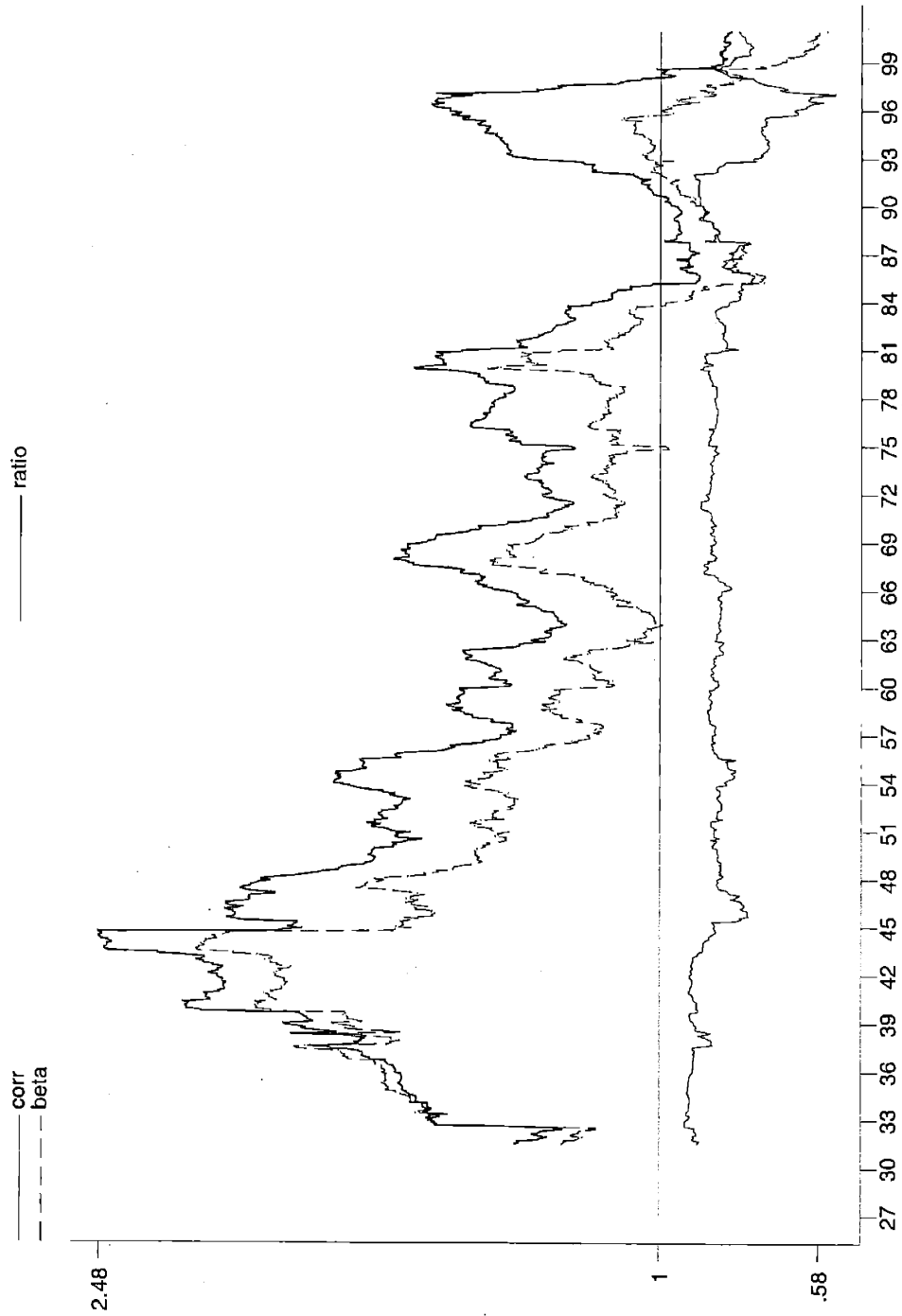
Figure 1-2: Beta of small stocks. The figure plots the estimates from 5-year rolling window regressions. The series originates from regressions where the dependent variable is the return on size decile 1 portfolio, and the independent variable is the excess return on the market value-weighted portfolio.

Figure 1-3: Decomposition of the beta of value stocks. The figure plots series of estimates from 5-year rolling window regressions. The series *beta* (dashed line) originates from CAPM time-series regressions where the dependent variable is the the return on B/M decile 10 portfolio. The series *corr* (thinner solid line) consists of correlation coefficient between the decile 10 portfolio return and the return on the rest of the market in the estimation window. The series *ratio* (thicker solid line) consists of the ratio of the standard deviation of the the decile 10 portfolio return to the standard deviation of the return on the rest of the market in the estimation window.
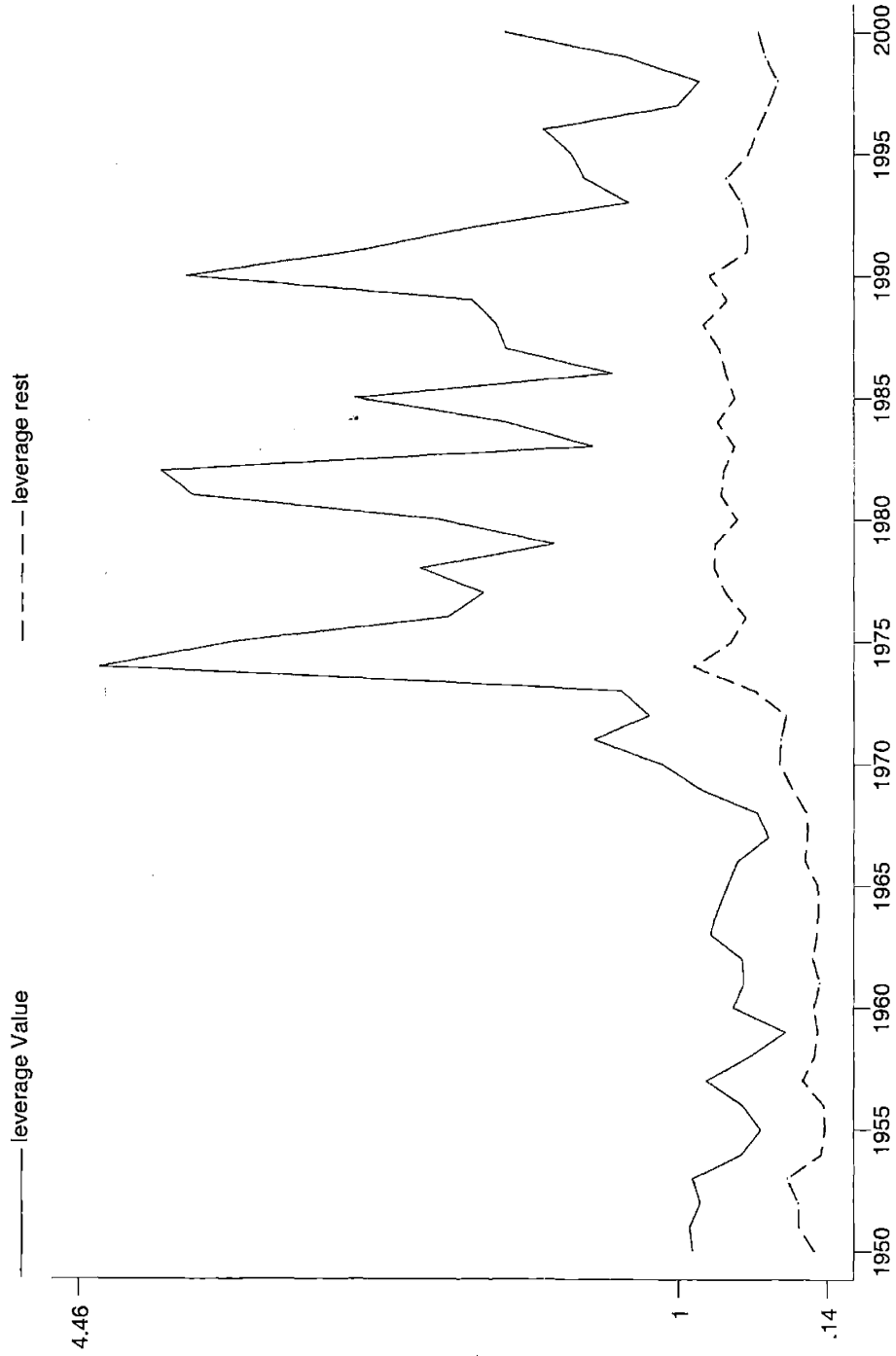
Figure 1-4: Leverage of value companies and the rest of the market. The figure plots the leverage series for the tenth B/M decile (solid line) and for the rest of the market (dashed line). For each company leverage is defined as the ratio of book value of debt to market value of equity. Portfolio average is the value weighted average of company leverage. Debt is defined as the sum of book value of current liabilities, long-term debt, convertible debt and preferred stocks. Accounting data is from the Compustat annual dataset.
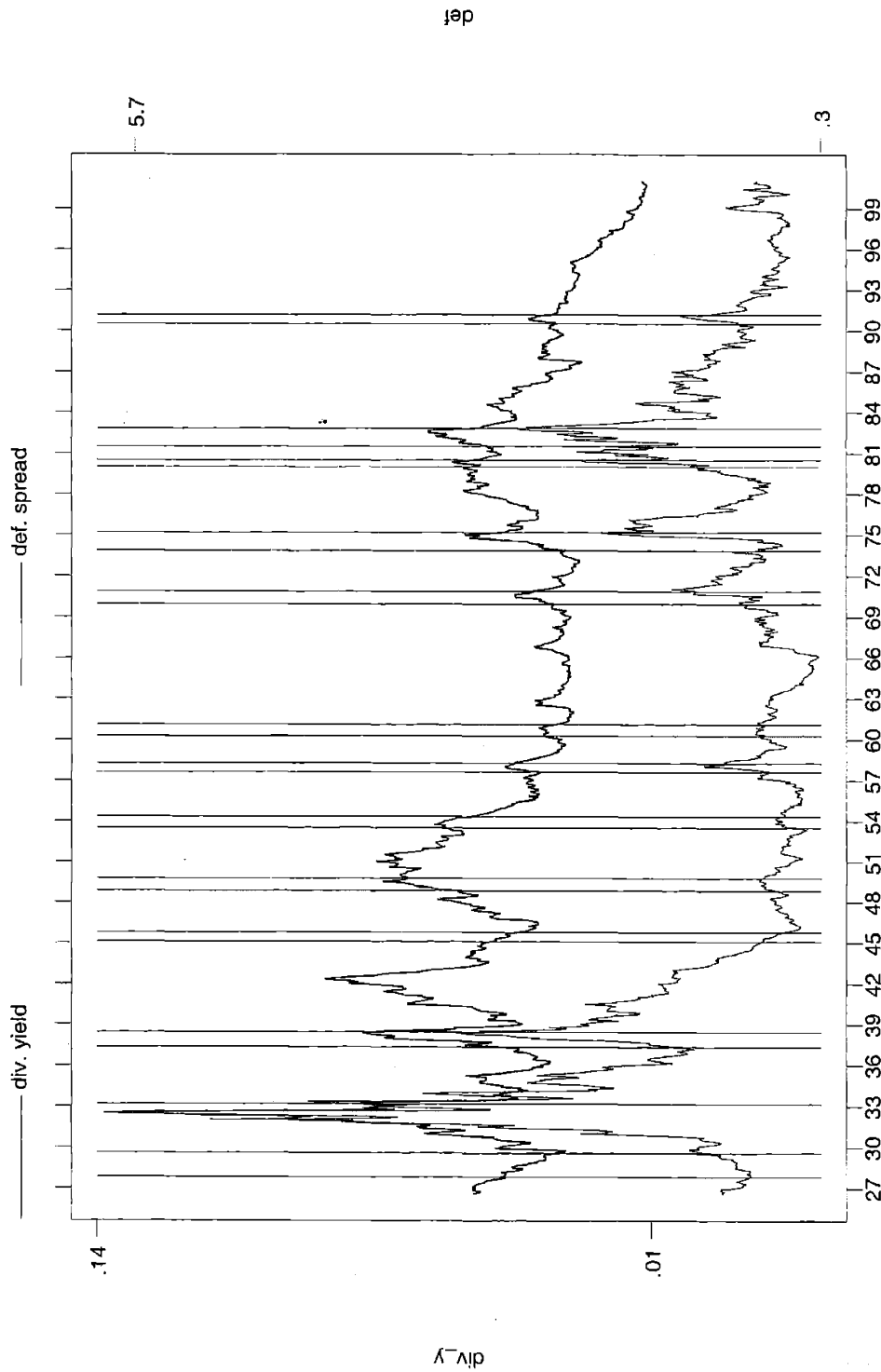
Figure 1-5: Dividend yield and default spread. The figure plots the series of the dividend yield on the S&P Composite Index (thicker line) and the default spread. The default spread is the end-of-month difference between the annualized yields on Aaa and Baa corporate bonds. The scale is different for the two series. The dividend yield is measured on the left vertical axis, and the default spread is measured on the right vertical axis. The vertical grid lines are the NBER business cycle peaks (P) and troughs (T). The dates are: 11/27(T), 8/29(P), 3/33(T), 5/37(P), 6/38(T), 2/45(P), 10/45(T), 11/48(P), 10/49(T), 11/48(P), 7/53(P), 5/54(T), 8/57(P), 4/58(T), 4/60(P), 2/61(T), 12/69(P), 11/70(T), 11/73(P), 3/75(T), 1/80(P), 7/80(T), 7/81(P), 11/82(T), 7/90(P), 3/91(T).
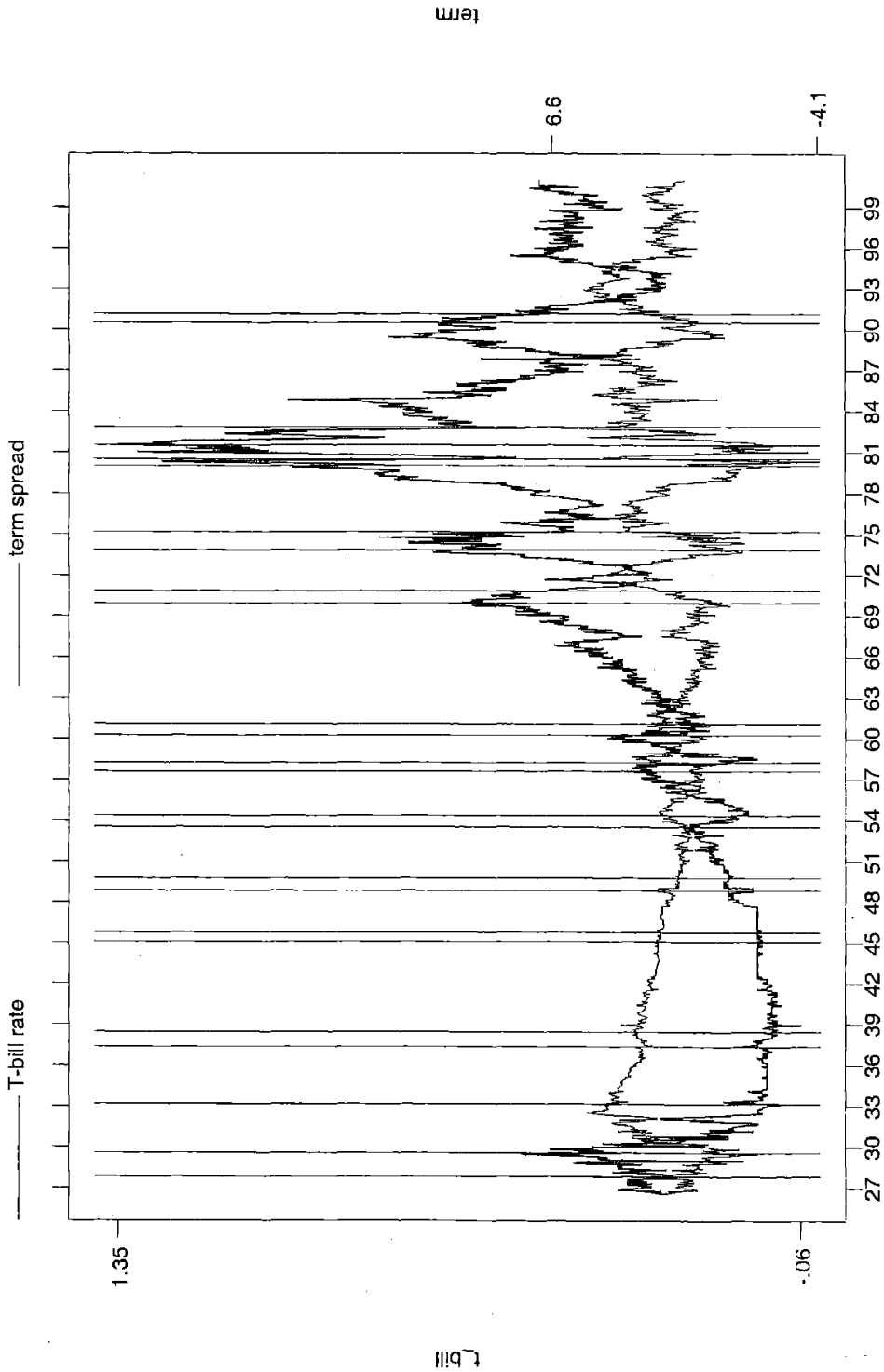
Figure 1-6: T-bill rate and term spread. The figure plots the series of the one-month T-bill rate (thick line) and the term spread (thin line). The term spread is the end-of-month difference between the annualized yields on Aaa corporate bonds and the one-month T-bill rate. The scale is different for the two series. The T-bill rate is measured on the left vertical axis, and the term spread is measured on the right vertical axis. The vertical grid lines are the NBER business cycle peaks (P) and troughs (T). The dates are: 11/27(T), 8/29(P), 3/33(T), 5/37(P), 6/38(T), 2/45(P), 10/45(T), 11/48(P), 10/49(T), 7/53(P), 5/54(T), 8/57(P), 4/58(T), 4/60(P), 2/61(T), 12/69(P), 11/70(T), 11/73(P), 3/75(T), 1/80(P), 7/80(T), 7/81(P), 11/82(T), 7/90(P), 3/91(T).
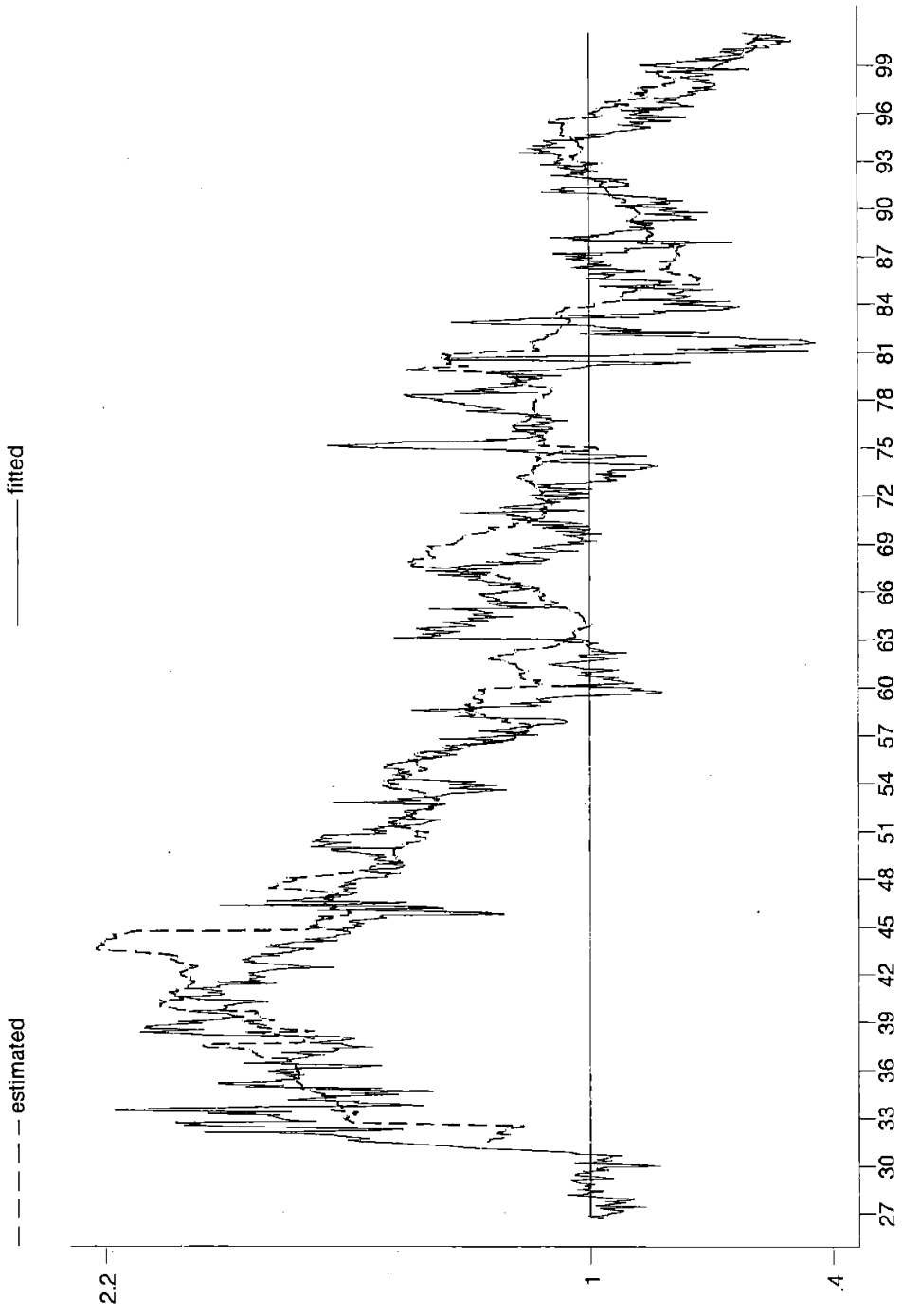
73

Figure 1-7: Tracking the beta of value stocks with state variables. The figure plots the time-series of the fitted beta (solid line) of the tenth B/M decile portfolio. The series has been generated according to Equation 1.8 in the text using monthly data from 7/1926 to 12/2000. The state variables are the t-bill rate, the dividend yield on the S&P Composite Inded, the default spread, the term spread, the growth rate of industrial production. The other series is the estimated beta (dashed line) for the same portfolio using 5-year rolling window regressions.
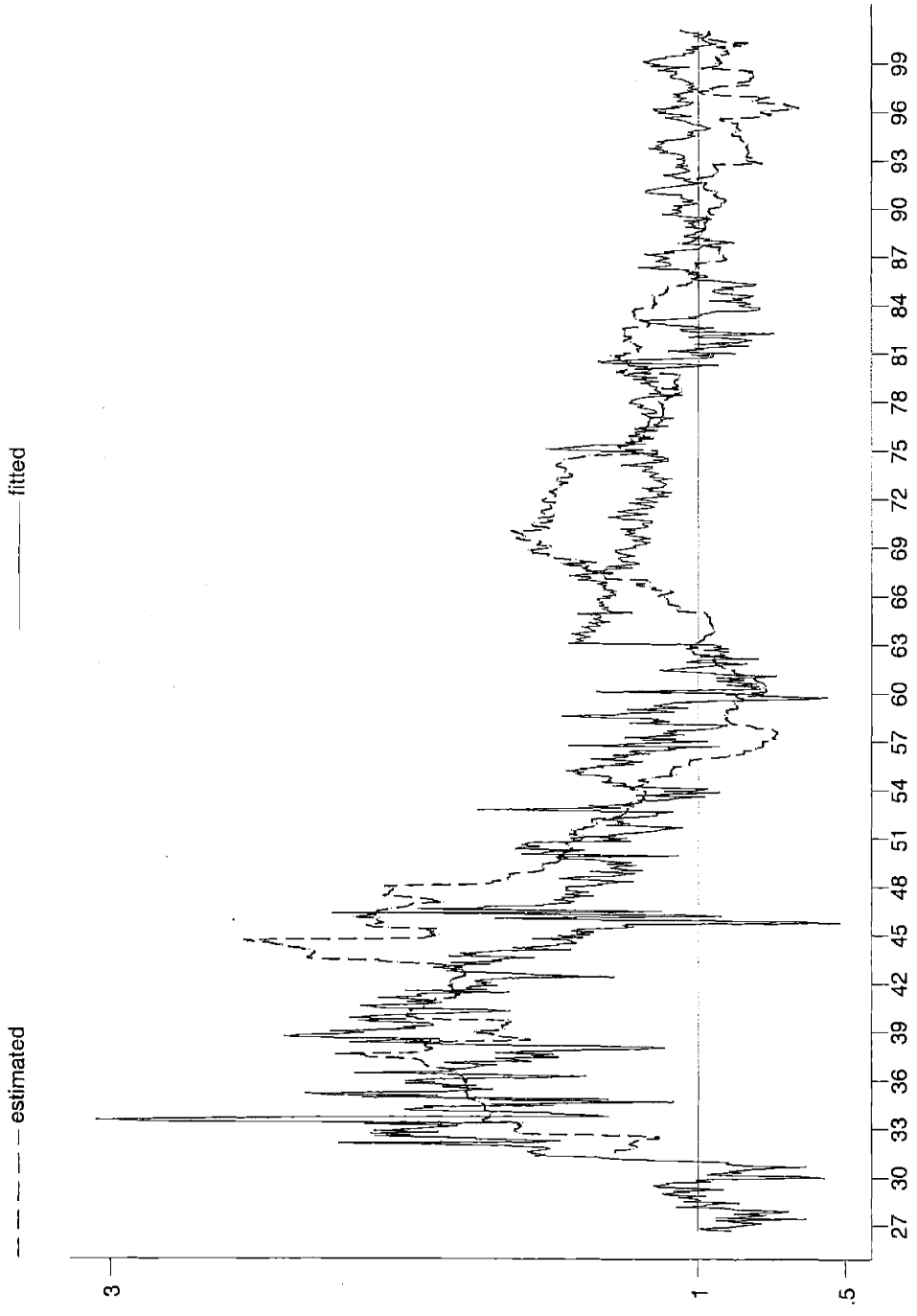
Figure 1-8: Tracking the beta of small stocks with state variables. The figure plots the time-series of the fitted beta (solid line) of the first size decile portfolio. The series has been generated according to Equation 1.8 in the text using monthly data from 7/1926 to 12/2000. The state variables are the t-bill rate, the dividend yield on the S&P Composite Inded, the default spread, the term spread, the growth rate of industrial production. The other series is the estimated beta (dashed line) for the same portfolio using 5-year rolling window regressions.
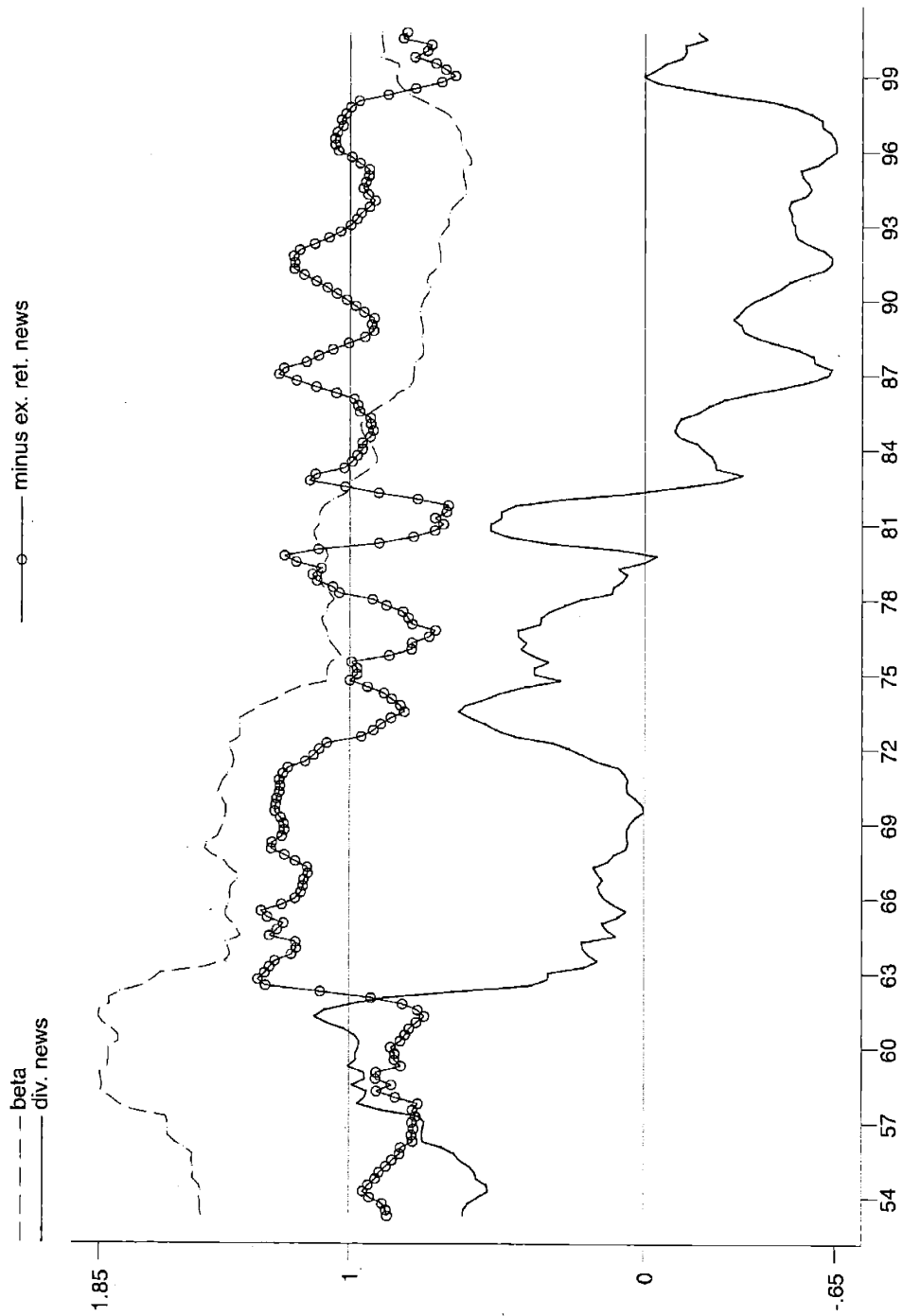
Figure 1-9: Beta of value stocks and its components. The figure plots the estimates from 25-year rolling window regressions where the independent variable is the innovation in the market return. The series *beta* (dashed line) originates from regressions where the dependent variable is the innovation in the return on B/M decile 10 portfolio. The series *div. news* (solid line) originates from regressions where the dependent variable is the dividend news component of the return on B/M decile 10 portfolio. The series *minus ex. ret. news* (circles) is the opposite of a series of estimates originating from regressions where the dependent variable is the excess returns news component of the return on B/M decile 10 portfolio.
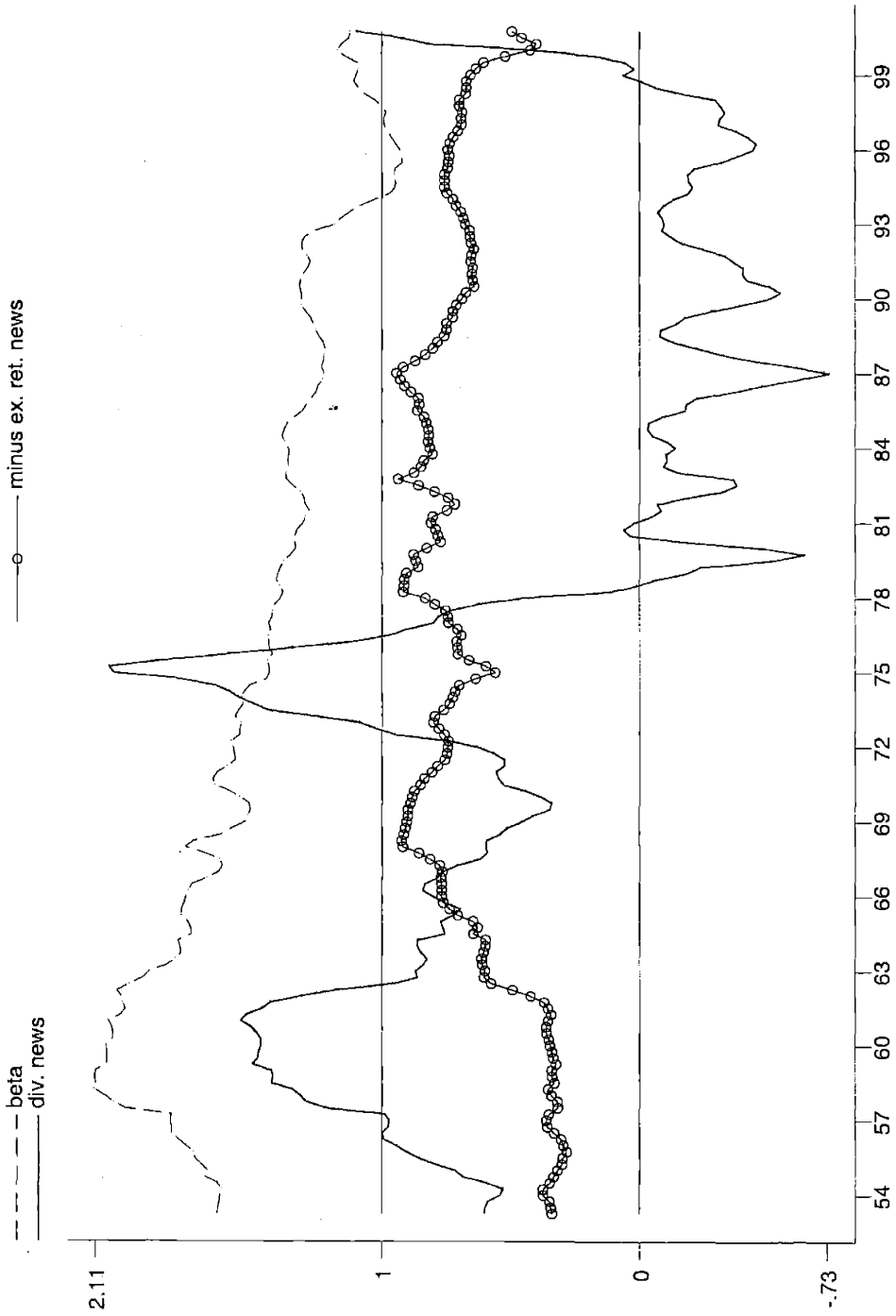
Figure 1-10: Beta of small stocks and its components. The figure plots the estimates from 25-year rolling window regressions where the independent variable is the innovation in the market return. The series *beta* (dashed line) originates from regressions where the dependent variable is the innovation in the return on size decile 1 portfolio. The series *div. news* (solid line) originates from regressions where the dependent variable is the dividend news component of the return on size decile 1 portfolio. The series *minus ex. ret. news* (circles) is the opposite of a series of estimates originating from regressions where the dependent variable is the excess returns news component of the return on size decile 1 portfolio.
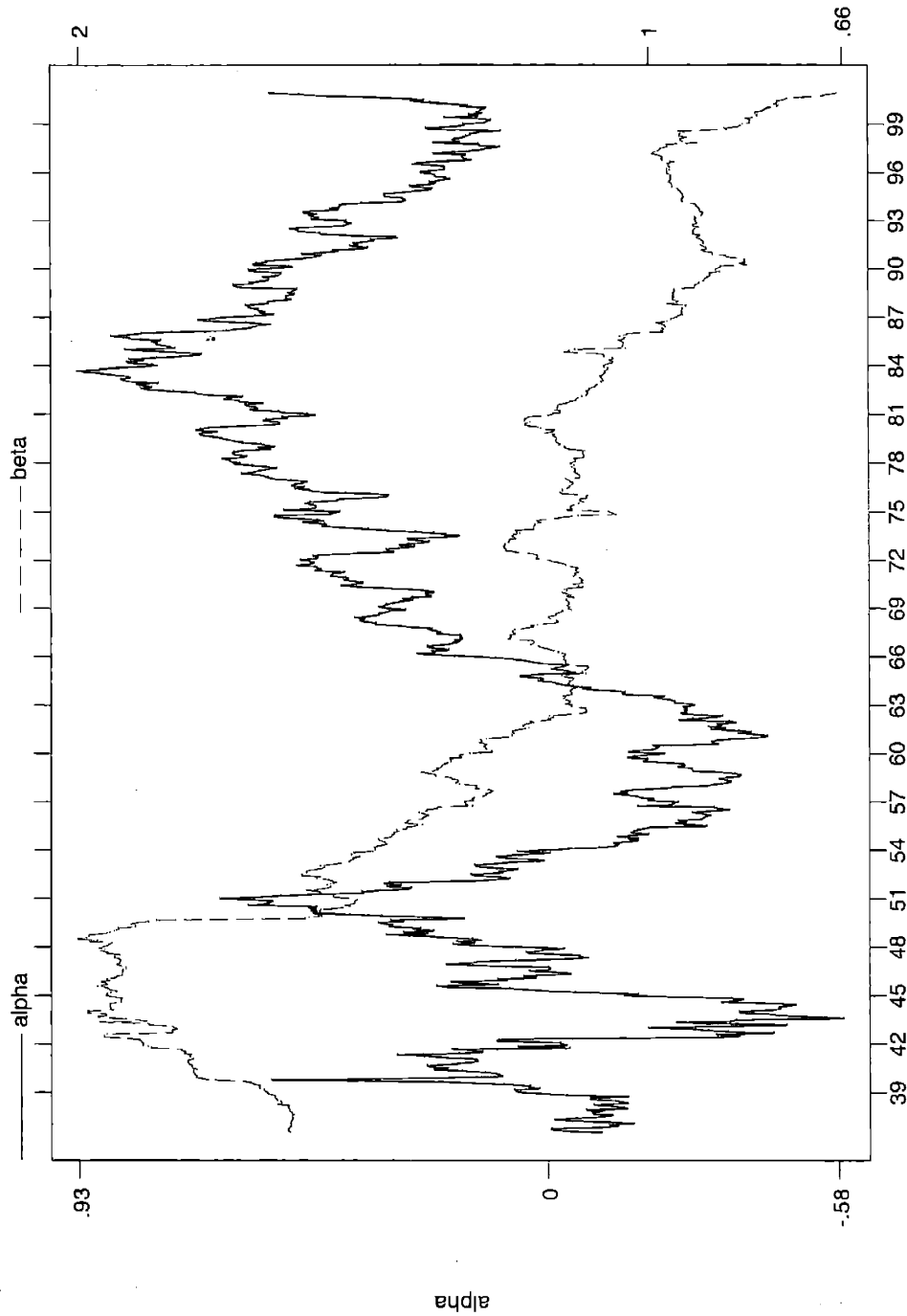
Figure 1-11: Alpha and Beta. The figure plots series of estimates from 10-year rolling window CAPM regressions where the dependent variable is the excess return on B/M decile 10 portfolio. The series *alpha* (solid line, left scale) is the intercept in the CAPM regression. The series *beta* (dashed line, and right scale) is the beta in the CAPM regression.

Bibliography

[1] Adrian, T., and F. Franzoni, 2002, "Learning about Beta: an explanation of the Value Premium", MIT mimeo.

[2] Barberis, N., and A. Shleifer, 2001, "Style investing", Working Paper.

[3] Black, F., M. Jensen, and M. Scholes, 1972, "The capital asset pricing model: Some empirical tests", in M. Jensen ed., *Studies in the Theory of Capital Markets*, (Praeger, New York, NY).

[4] Bollerslev, T., R. Engle, and J. Wooldridge, 1988, "A capital asset pricing model with time varying covariances", *Journal of Political Economy*, 96, pp. 116-131.

[5] Braun, P., D. Nelson, and A. Sunier, 1995, "Good news, bad news, volatility and betas", *Journal of Finance*, 50, pp. 1575-1604.

[6] Campbell, J., 1991, "A Variance Decomposition for Stock Returns", *Economic Journal*, 101, pp. 157-179.

[7] Campbell, J., and J. Mei, 1993, "Where do Betas Come From? Asset Price Dynamics and the Sources of Systematic Risk", *Review of Financial Studies*, 6, pp. 567-592.

[8] Campbell, J., M. Lettau, B. Malkiel, and Y. Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk", *Journal of Finance*, 56, pp. 1-43.

[9] Campbell, J., and R. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies*, 1, pp. 195-227.

[10] Chan, Y., and N. Chen, 1991, "Structural and return characteristics of small and large firms", *Journal of Finance*, 46, pp. 1467-1484.

[11] Chen, N., 1991, "Financial Investment Opportunities and the Macroeconomy", *Journal of Finance*, 46, pp. 529-554.

[12] Daniel, K., and S. Titman, 1997, "Evidence on the characteristics of cross-sectional variation in stock returns", *Journal of Finance*, 52, pp. 1-33.

[13] Davis J., Fama, E., and K. French, 2000, "Characteristics, covariances, and average returns:1929 to 1997", *Journal of Finance*, 55, pp. 389-427.

[14] Fama, E., 1990, "Stock Returns, Expected Returns, and Real Activity", *Journal of Finance*, 45, pp. 1089-1108.

[15] Fama, E., and K. French, 1988, "Dividend Yields and expected Stocks Returns", *Journal of Financial Economics*, 22, pp. 3-25.

[16] Fama, E., and K. French, 1989, "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics*, 25, pp. 23-49.

[17] Fama, E., and K. French, 1992, "The cross-section of expected stock returns", *Journal of Finance*, 47, pp. 427-465.

[18] Fama, E., and K. French, 1993, "Common risk factors in the returns on stocks and bonds", *Journal of Financial Economics*, 33, pp. 3-56.

[19] Fama, E., and K. French, 1995, "Size and book-to-market factors in earnings and returns", *Journal of Finance*, 50, pp. 131-155.

[20] Fama, E., and K. French, 1996, "Multifactor explanations of asset pricing anomalies", *Journal of Finance*, 51, pp. 55-84.

[21] Fama, E., and K. French, 2001, "Disappearing dividends: changing firm characteristics or lower propensity to pay?", *Journal of Financial Economics*, 60, pp. 3-43.

[22] Ferson, W., and C. Harvey, 1991, "The variation of economics risk premiums", *Journal of Political Economy*, 99, pp. 385-415.

[23] Ferson, W., and C. Harvey, 1999, "Conditioning variables and the cross-section of stock returns", *Journal of Finance*, 54, pp. 1325-1360.

[24] Ferson, W., and R. Korajczyk, 1995, "Do arbitrage pricing models explain the predictability of stock returns", *Journal of Business*, 68, pp. 309-349.

[25] Ferson, W., and R. Schadt, 1996, "Measuring fund strategy and performance in changing economic conditions", *Journal of Finance*, 51, pp. 425-462.

[26] Harvey, C., 1989, "Time-varying conditional covariances in tests of asset pricing models", *Journal of Financial Economics*, 24, pp. 289-317.

[27] Jagannathan, R., and Z. Wang, 1996, "The conditional CAPM and the cross-section of expected returns", *Journal of Finance*, 51, pp. 3-53.

[28] Keim, D., and R. Stambaugh, 1986, "Predicting Returns in Stock and Bond Markets", *Journal of Financial Economics*, 17, pp. 357-390.

[29] Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying", *Journal of Political Economy*, 109, pp. 1238-1287.

[30] Lewellen, J., 1999, "The time-series relations among expected return, risk, and book-to-market", *Journal of Financial Economics*, 54, pp. 5-43.

[31] Lintner, J., 1965, "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics*, 47, pp. 13.-37.

[32] Lo, A. W., and A. C. MacKinlay, 1990, "When are contrarian profits due to stock market overreaction?", *Review of Financial Studies*, 3, pp. 175-206.

[33] Scholes, M., and J. Williams, 1977, "Estimating beta from nonsynchronous data", *Journal of Financial Economics*, 5, pp. 309-327.

[34] Schwert, W., 1989, "Why Does Stock Market Volatility Change Over Time?", *Journal of Finance*, 44, pp. 1115-1153.

[35] Shanken, J., 1990, "Intertemporal asset pricing: An empirical investigation", *Journal of Econometrics*, 45, pp. 99-120.

[36] Sharpe, W., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk", *Journal of Finance*, 19, pp. 425-442.

Chapter 2

Learning About Beta: An Explanation of the Value Premium

(joint with Tobias Adrian)

2.1 Introduction

Since its invention by Sharpe (1964) and Lintner (1965), the Capital Asset Pricing Model (CAPM) has experienced varying fortune. While the early tests (Black, Jensen, and Scholes, 1972, Fama and MacBeth, 1973) established its empirical success, the model was still unknown to large part of the financial industry. However, right when the model started to become widely used by practitioners, the academic profession discovered its first empirical failures (Basu, 1977, Banz, 1981) .

Fama and French (1992, 1993) present the most dramatic of these failures. They show that market risk, as measured by beta, does not explain the cross-section of average returns for portfolios sorted on size and book-to-market (B/M). In particular small stocks pay a premium relative to the prediction of the CAPM ('size anomaly'), and so do value stocks ('value puzzle').

The response to this discovery has produced different strands of literature. Fama and French (1993, 1995, 1996) propose a risk based explanation, where the size and B/M factors proxy for some underlying distress risk, to which small and value companies are subject.

This explanation is backed by the theory of multifactor asset pricing models, such as Merton's (1971) Intertemporal CAPM, and Ross's (1976) Asset Pricing Theory.

On the other hand, authors in the 'behavioral' literature suggest that characteristics, rather than risk, are priced in equilibrium. Daniel and Titman (1997), for example, argue that investors have an exaggerated perception of the riskiness of these portfolios, and require a premium to hold them.

More recently, Lettau and Ludvigson (2001) have produced favorable evidence for the version of the CAPM that is implied by the Consumption CAPM of Breeden (1979). In particular, they argue that the CAPM holds under conditional probability distributions. Hence, the correct empirical implementation of the model requires the use of scaled factors, i.e. risk factors multiplied by some appropriate state variables.

So much attention around the CAPM is justified on the grounds that the simplicity of the model makes it theoretically appealing. Moreover, the fact that market participants have widely adopted the CAPM, makes it more puzzling that empirical studies reject it as an equilibrium model. So, why does the CAPM not work well empirically?

We propose an explanation that contains elements of the different strands of literature cited above, but at the same time is innovative, as it draws on new evidence concerning the behavior of value stocks. Specifically, we refer to the recent paper by Franzoni (2002), which shows that the loading on the market factor for value and small stocks has decreased significantly for the past sixty years. More precisely, the beta of these portfolios has dropped by about 75%, from 2.2 in the early forties to below 0.55 in the nineties.

It may be a coincidence that the portfolios that created the biggest problems for the CAPM, are also the ones that experienced a major decline in their riskiness. However, we believe that there is a relation between the two events, and we support our belief with the

evidence in Figure 2-1, which reproduces Figure 11 in Franzoni's (2002) paper. The graph shows that the intercepts (alphas) from rolling window regressions for value stocks become significant in the sixties, right after the estimated beta declines drastically, and they keep increasing as beta goes down. The correlation between the two series is about -50%. Since alpha is a measure of mispricing, the figure suggests that the emergence of mispricing is related to the decline in beta.

The theory that we develop in the paper provides an intuition for this evidence. In an environment where the loadings on risk factors change dramatically, investors may not know the exact riskiness of the portfolios that they are going to hold. Consequently, they need to form beliefs about the betas, and these beliefs are affected by the past levels of the loading. Moreover, as the riskiness of value stocks has been decreasing over time, it is likely the case that investors' expected beta is significantly higher than the actual beta. The implication is that while investors require an expected return that is proportional to the riskiness they perceive, the econometrician observes a premium in excess of the realized riskiness of these stocks, and the CAPM is rejected.

We develop this argument through an equilibrium model of learning, with unobservable and time-varying factor loadings. In the model the CAPM holds under investors' subjective probability distribution. However, an econometrician, who looks at realized returns, observes positive mispricing relative to the CAPM, whenever the expected beta is above the actual beta. A further implication is that idiosyncractic risk contributes to determine the equilibrium expected return through its effect on the expected factor loading. Idiosyincratic risk operates as noise, which affects the speed of learning, and the discrepancy between expected and true factor loadings.

We simulate the model under different assumptions on the process of the loading, and

85

on investors' information set. In the best performing simulations we manage to account for about 80% of the 'value premium'. We obtain this result with a very plausible choice for the underlying parameters.

Our theory has elements of the Lettau and Ludvigson's (2001) explanation, and more generally of the conditional CAPM literature, as it stresses the importance of looking at conditional moments of the return distribution. It is also related to the foundation of the behavioral theories that argue that characteristics are priced in equlibrium (Daniel and Titman, 1997), because investors misperceive the riskiness of value stocks, although in our model this is the outcome of rational learning. Finally, like in Fama and French (1993), value stocks, and more generally all portfolios for which factor loadings change dramatically over time, expose investors to an additional source of risk. In our model, however, the additional risk is not due to a systematic distress factor, but to idiosyncratic estimation risk, and it is not priced ex-ante, but it appears ex-post in the empirical tests.

Learning in a symmetric information setting was first treated in the seminal paper by Gennotte (1986). Gennotte (1986) demonstrates that the optimal portfolio problem under incomplete information can be separated into a learning problem and a maximization problem under the Bayesian posterior distribution, an approach that we adopt as well. More recently, Brennan and Xia (2001) employ a learning model to explain the equity premium puzzle. In their model, the drift of the aggregate dividend process is unobservable. Investors update their belief about the true drift of the dividend process, which leads to a learning premium in aggregate asset prices. The calibration of their model accounts for part of the equity premium. In our model, we focus on cross-sectional asset pricing anomalies, and calibrate parameters, in particular risk aversion, to reproduce the observed equity premium. The aggregate equity premium is not the focus of our paper.

The learning paper that we believe to be most closely related to ours is the one by Lewellen and Shanken (2002). They assume that the mean of the dividend process is unobserved. Their paper accounts for both predictability and excess variance of returns (Shiller, 1981). Ex-ante investors use all the available information, so there are no arbitrage opportunities. Predictability results as the ex-post effect of Bayesian updating of investors' beliefs on the unobservable mean of the dividend, and it is detected only by the econometrician, who looks back in time. Excess volatility is also the consequence of learning, as new realizations of the dividend cause investors to constantly revise their beliefs. Lewellen and Shanken (2002) explain cross-sectional anomalies as the result of learning on fundamentals. Our approach, instead, focuses on learning about riskiness, in an environment with time-varying factor loadings.

Other papers have focused on the portfolio allocation problem for investors with incomplete information. In Barberis (2000), investors are unsure whether returns are predictable or not. This uncertainty leads to an excessive allocation of wealth to stocks, and this allocation is larger, the longer the investment horizon. Barberis (2000) does not address cross-sectional implications of incomplete information, which is the focus of our paper. Finally, Pastor (2002) examines the Bayesian decision problem of investors who are unsure whether the CAPM holds. One of the alternatives to the CAPM that Pastor (2002) considers is the multifactor model proposed by Fama and French (1993). Pastor (2002) finds that even if an investor strongly believes that the market portfolio is mean variance efficient, he should invest a substantial amount of her wealth in value stocks. Unlike our paper, Pastor (2002) does not examine the reasons for the premium generated by the value portfolio, he takes this premium as given.

This paper is organized as follows. In Section 2.2 we develop a model of learning

about time-varying factor loadings. We derive the CAPM as the equilibrium in the model, and compute the population analogs of the moments estimated by the econometrician. In Section 2.3 we describe our strategy for simulating the model, and present the simulation results. In Section 2.4 we modify the original model and let investors learn about the long run level of the factor loading as well. We also present results of the simulations for this different version of the model. Finally, in Section 2.5 we draw the conclusions of this work, and propose directions for future research.

2.2 The model

The conditional asset pricing literature has provided extensive evidence that factor loadings change over time [1]. This finding makes it is reasonable to believe that investors may not have a precise idea of what the true riskiness of an asset is in the moment when they trade it.

In a world with uncertainty about the relevant parameters, investors have to infer the factor loadings from the observation of past returns and dividends. Depending on how fast learning occurs, investors' beliefs can diverge significantly from the truth, especially if factor loadings keep changing.

Since investors' expectations of factor loadings determine the risk premium required to hold assets, it can be the case that the current value of the riskiness of a portfolio, as measured by CAPM tests, does not entirely explain expected returns. This rejection of the CAPM occurs even if the CAPM holds from investors' ex-ante perspective. In particular, if a portfolio beta has been decreasing, as in the case of value stocks (Franzoni, 2002), and investors' expectations are tied to the past high levels of beta, CAPM tests can detect

---

[1]Examples of empirical studies that document time-varying factor loadings are: Shanken (1990), Ferson and Harvey (1991), Ferson and Korajczyk (1995), Braun, Nelson, and Sunier, (1995), Lettau and Ludvigson (2001), and Franzoni (2002).

a premium relative to the measured level of market risk. We will demonstrate that this phenomenon can account for a relevant part of the so-called 'value premium'.

The model that we present below intends to capture the idea that discrepancies between expected and true factor loadings cause the econometrician to observe systematic deviations from the CAPM, for assets whose riskiness is changing over time.

2.2.1 The set-up

We consider an overlapping generations economy. There is only one physical good, which can be allocated to investment or consumption. We assume that there are N risky assets, and a riskless asset. The dividend of stock i evolves according to a factor structure

$$d_t^i = D^i + x_t b_t^i + \varepsilon_t^i \tag{2.1}$$

where the factor x_t has zero mean and variance σ^2. The idiosyncractic component ε_t^i is uncorrelated across stocks. The realized processes of d_t and x_t are part of investors' information set, but investors can not observe the factor loading b_t.

The factor loading varies according to an auto-regressive process

$$b_{t+1}^i - B^i = F\left(b_t^i - B^i\right) + u_{t+1}^i \tag{2.2}$$

where u_{t+1}^i is a white noise, uncorrelated across stocks, and uncorrelated with ε_{t+1}^i. We assume for now that the long run mean of the loading B is known to investors. Later on we will remove this assumption, and let the investors learn about B as well.

The investors' inference problem can be solved by applying the Kalman filter to the system of equations (2.1) and (2.2) (see, e.g., Hamilton, 1994). Following Gennotte (1986), we separate the learning problem from the portfolio optimization. In order to apply the Kalman filter to this system with time-varying coefficients, we need to assume normality of

the disturbances conditional on information up to $t-1$ and on x_t, that is

$$\varepsilon_t | x_t, I_{t-1} \quad \sim \quad N(0, \sigma_\varepsilon^2)$$

$$u_{t+1} | x_t, I_{t-1} \quad \sim \quad N(0, \sigma_u^2). \tag{2.3}$$

The representative agent lives for two periods and maximizes consumption in the second period. The utility function is quadratic

$$u(c_{t+1}) = c_{t+1} - \frac{\tilde{A}}{2} c_{t+1}^2. \tag{2.4}$$

The assumption of quadratic utility will allow us to obtain a CAPM even if the unconditional distribution of dividends is not normal.

2.2.2 The pricing function

The problem of the representative consumer is therefore

$$\text{Max} \ E_t c_{t+1} - \frac{\tilde{A}}{2} E_{t+1} c_{t+1}^2$$

$$\text{s.t.} \quad c_{t+1} = (1+r)W_t + \alpha(p_{t+1} + d_{t+1} - (1+r)p_t) \tag{2.5}$$

where α is the $N \times 1$ vector of quantities of stocks invested in each risky asset, p_t and d_t are $N \times 1$ vectors of prices and dividends.

The usual first order conditions apply

$$E_t(u'(c_{t+1})(p_{t+1} + d_{t+1})) = E_t(u'(c_{t+1}))(1+r)p_t. \tag{2.6}$$

Replacing for the utility function gives

$$E_t(p_{t+1} + d_{t+1}) - (1 + r)p_t = \tilde{A}E_t\left(c_{t+1}(p_{t+1} + d_{t+1} - (1 + r)p_t)\right). \tag{2.7}$$

The equilibrium condition is that total consumption is equal to total dividends

$$c_t = \sum_{i=1}^{N} d_t^i = ND + x_t \sum_{i=1}^{N} b_t^i + \sum_{i=1}^{N} \varepsilon_t^i \tag{2.8}$$

where D is assumed to be the same across stocks, for simplicity.

We normalize the sum of the loadings to one, and we assume that the Law of Large Numbers applies to the idiosyncratic components of dividends

$$\frac{\sum_{i=1}^{N} b_t^i}{N} = 1$$
$$\frac{\sum_{i=1}^{N} \varepsilon_t^i}{N} \rightarrow 0 \quad \text{as} \quad N \rightarrow \infty. \tag{2.9}$$

The assumptions in (2.9) imply the following equilibrium condition

$$c_t = N(x_t + D) \tag{2.10}$$

which provides the interpretation for x_t as the random component of the average dividend at time t.

From now on we drop the superscript i for notational convenience, and all variables are referred to a single asset i, unless otherwise specified. Replacing (2.10) into the pricing equation (2.7) we obtain

$$E_t p_{t+1} + E_t\left(D + b_{t+1}x_{t+1} + \varepsilon_{t+1}\right) - (1 + r)p_t$$

$$= \tilde{A}NE_t\left((x_{t+1} + D)\left(p_{t+1} + D + b_{t+1}x_{t+1} + \varepsilon_{t+1} - (1 + r)p_t\right)\right). \tag{2.11}$$

Let $A = \tilde{A}N$, and let b^e_t be the time t expectation of b_{t+1} under investors' information set. Knowing that x_t is i.i.d. with distribution $f(0, \sigma^2)$, the pricing equation becomes

$$E_t p_{t+1} + D - (1+r)p_t = ADE_t(p_{t+1}) + AD^2 - AD(1+r)p_t$$

$$+ AE_t\left(x_{t+1}p_{t+1}\right) + Ab^e_t\sigma^2 \tag{2.12}$$

where we have used the assumption that x_{t+1} does not contain any information about b_{t+1}, so that

$$E_t\left(b_{t+1}x_{t+1}\right) = b^e_t E_t x_{t+1} = 0.$$

We prove in the appendix that under these assumptions the pricing function is (equation (A-2))

$$p_t = \frac{D}{r} - \frac{A\sigma^2}{r\left(1 - AD\right)}b^e_t. \tag{2.13}$$

Notice that $(1 - AD)$ has to be positive for marginal utility to be positive at the expected value of dividends. The asset price in equation (2.13) has the usual interpretation. The first component is the expected discounted cash flow. The second component is the risk premium, which depends on investors' expectation of the risk factor loading, risk aversion, and the factor's volatility.

2.2.3 The CAPM

In this model the CAPM holds under the subjective probability distribution, i.e. the distribution of dividends as perceived by investors. In order to prove this result, we need to compute returns. We look at absolute excess returns, defined as

$$R_{t+1} = p_{t+1} + d_{t+1} - (1 + r)p_{t+1}. \tag{2.14}$$

This definition implies that the return on stock i is

$$R_{t+1} = x_{t+1}b + \varepsilon_{t+1} + \frac{A\sigma^2\left((1+r)b_t^e - b_{t+1}^e\right)}{r(1 - AD)}. \tag{2.15}$$

So that the expected return under investors' probability distribution is

$$E_t R_{t+1} = \frac{A\sigma^2}{1 - AD}b_t^e. \tag{2.16}$$

We define the absolute market return as the average of individual stock returns

$$R_{t+1}^m = \frac{\sum_{i=1}^N R_{t+1}^i}{N}. \tag{2.17}$$

Hence, the absolute market return is

$$R_{t+1}^m = x_{t+1} + \frac{A\sigma^2}{1 - AD} \tag{2.18}$$

due to the assumptions in (2.9). Moreover, the expected absolute market return is

$$E_t R_{t+1}^m = \frac{A\sigma^2}{1 - AD}. \tag{2.19}$$

Now, in order to prove that the CAPM holds under the subjective probability distribution, we just need to show that

$$\beta_t \equiv \frac{Cov_t\left(R_{t+1}^i, R_{t+1}^m\right)}{Var_t\left(R_{t+1}^m\right)} = b_t^e. \tag{2.20}$$

93

In the appendix, we prove that the covariance term is

$$Cov_t \left(R_{t+1}^i, R_{t+1}^m \right) = \sigma^2 b_t^e \tag{2.21}$$

while the variance of the market return is simply σ^2 (see equations (A-3) and (A-5)).

Hence, the CAPM holds under the subjective probability distribution, and the market beta is equal to b_t^e, i.e. the expectation of the factor loading, conditional on investors' information set. While we have derived the CAPM for absolute returns, it is straightforward to show that it holds for relative returns as well.

2.2.4 Observed Mispricing

So far we have looked at investors' beliefs, as they are the relevant probability distribution for pricing. However, the properties of empirical tests depend on the objective probability distribution. In an efficient market the two distributions should be the same, but this is not the case when there is parameter uncertainty. As in Lewellen and Shanken (2002), also in our set-up parameter uncertainty causes the econometrician to observe mispricing, while from investors' perspective markets are efficient.

To represent the econometrician's point of view, we look at the objective probability distribution, in which the process for b_t is treated as observable, unlike under the subjective distribution, where investors have to form beliefs about b_t.

In the time-series tests of the CAPM, such as the ones in Fama and French (1993), the excess return on stock i is regressed on the market excess return over some interval of time. If the intercept in the regression is statistically and economically significant, the CAPM is rejected. In the case of value stocks (fifth B/M quintile) Fama and French find a premium of about 0.5% monthly over the 1963-1993 sample. This evidence, along with the finding of a significant B/M factor in cross-sectional regressions, is referred to as 'value premium'.

94

In terms of population moments, the analog of the regression intercept is

$$
\begin{aligned}
\hat{\alpha}_t &= E_t^o R_{t+1} - \hat{\beta}_t E_t^o R_{t+1}^m \\
&= E_t^o R_{t+1} - E_t R_{t+1} + (\beta_t - \hat{\beta}_t) E_t R_{t+1}^m
\end{aligned} \tag{2.22}
$$

where $\hat{\beta}$ is the population equivalent of the beta estimated by the econometrician, whereas β is the subjective beta. The subscript o denotes moments of the objective distribution. Note that this population moment is the limit in a repeated sampling thought experiment, where the average of the estimated premia is taken. To obtain this expression, we used the fact that the market return has the same expectation under the subjective and objective distributions. We refer to $\hat{\alpha}$ as to the estimated mispricing.

In order to compute the mispricing, we need to provide an expression for the estimated beta

$$
\hat{\beta}_t = \frac{Cov_t^o \left(R_{t+1}^i, R_{t+1}^m \right)}{Var_t^o \left(R_{t+1}^m \right)}. \tag{2.23}
$$

In the appendix (see equation (A-6)) we prove that if the distribution of x_t is symmetric, then the covariance is simply

$$
Cov_t^o \left(R_{t+1}^i, R_{t+1}^m \right) = b_{t+1} \sigma^2. \tag{2.24}
$$

The variance of the market return is σ^2, the same as under the subjective probability distribution. So, the theoretical value of the estimated beta coincides with the true factor loading

$$
\hat{\beta}_t = b_{t+1}. \tag{2.25}
$$

The other component of the estimated mispricing is the difference between objective

and subjective expected returns

$$E_t^o R_{t+1} - E R_{t+1} = \frac{\left(b_t^e - b_{t+1}\right) A \sigma^2}{r\left(1 - AD\right)} E_t^o \left(\frac{F P_{t+1|t} x_{t+1}^2}{x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2} \right) \tag{2.26}$$

where $P_{t+1|t}$ is the mean squared forecast error from the prediction of b_{t+1} made at time t.

Since $P_{t+1|t}$ is positive the expectation in equation (2.26) is clearly larger than zero.

To obtain equation (2.26) we have used the fact that the updating formula for b_{t+1}^e is

$$b_{t+1}^e = B + F\left(b_t^e - B\right) + K(x_{t+1})\left(\varepsilon_{t+1} + b_{t+1} x_{t+1} - b_t^e x_{t+1}\right) \tag{2.27}$$

where

$$K\left(x_{t+1}\right) = \frac{F P_{t+1|t} x_{t+1}}{x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2}.$$

By replacing equations (2.20), (2.23), and (2.26) into (2.22) we obtain the final expression

for the estimated mispricing

$$\hat{\alpha}_t = \left(b_t^e - b_{t+1}\right)\left(1 + \frac{G_t}{r}\right) E_t R_{t+1}^m \tag{2.28}$$

where

$$G_t = E_t^o \left(\frac{F x_{t+1}^2 P_{t+1|t}}{x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2} \right)$$

and G_t is positive.

From equation (2.28), it is evident that whenever investors' belief of the factor loading

is above the true level, the population equivalent of the observed mispricing is positive.

In conclusion, the bottom line of this model is that even if the CAPM holds ex-ante, the

empirical tests can still reject CAPM, as a result of the discrepancy between the subjective

and objective probability distributions. In particular, the tests can find a premium, if

investors' belief of the factor loading is higher than the true value. Moreover, it is the case

that idiosyncratic risk contributes to determine equilibrium expected returns, as it affects the learning process, and contributes to determine the difference between b_t^e and b_{t+1}.

The next section presents the results of simulations aimed at assessing whether the theoretical result derived in this model can account for a relevant fraction of the 'value premium'.

2.3 Simulations

The first issue that one faces when simulating the model is the choice of the risk factor. If we took the model literally, the natural choice for the factor would be aggregate consumption. On the other hand, the 'equity premium puzzle' literature started by Mehra and Prescot (1985), has pointed out that the correlation between consumption and market returns is not enough to explain the expected return on the market. In other words, for plausible levels of risk aversion consumption risk is just too low to explain the realized returns on risky assets.

In terms of our simulations, using consumption as risk factor would produce returns that are not volatile enough compared to actual market returns. It would then be difficult to establish a connection between the simulated variables and the real ones.

A related issue is what interpretation to give to the dividend process in our model. Dividend news are not volatile enough as to justify the volatility of realized returns (Shiller, 1981, Campbell and Shiller, 1988). Hence, dividends in our model must be interpreted in a broader sense, as all sort of news that are relevant for investors' wealth.

Although recent literature (Lettau and Ludvigson, 2001) has provided evidence that the consumption CAPM can be a valid asset pricing model once scaled factors are taken into account, we believe there is a more direct way to implement our model. In particular,

we will assume that x_t is the projection on the space of asset returns of some unspecified meta-factors that are relevant for asset pricing. Consequently, equation (2.18) allows us to set x_t equal to the de-meaned excess market return.

Our choice of the risk factor does not deal with the equity premium and volatility puzzles. On the other hand, this does not interfere with our exercise, as we intend to explain cross-sectional anomalies.

We use monthly return data from July 1926 to December 2000[2]. The market return is computed as the value-weighted portfolio of the universe of stocks in CRSP. The data set also contains returns on B/M decile portfolios, formed as in Fama and French (1993).

The model allows us to normalize the price of the market portfolio to one. This normalization implies that the relative market return is equal to the absolute market return, while the relative portfolio returns are equal to the absolute returns times the weight of the portfolio in the market.

It follows that σ can be set equal to the volatility of the excess market return, which in our data is 5.5% monthly. The parameter A is determined by setting the theoretical equity premium in equation (2.19) equal to the realized equity premium (0.68% monthly). The risk free rate r is set to its average realized value of 0.31%. Finally, from the normalization of the market price we deduce that D has to be equal to the mean excess market return plus r. Its value is therefore 0.99%.

In our simulations we will focus on the returns of value stocks, defined as the tenth B/M decile portfolio. This choice drives us in the specification of the variance of ε_t that represents the noise in the observational equation in the Kalman system (equation (2.2)). Its variance is not a free parameter, because ε_t is also part of the idiosyncratic risk of stock

---

[2]The data can be obtained from Ken French's website.

returns (equation (2.15)). Hence, we set σ_ε^2 to the level such that the R^2 in the CAPM regressions is about 70%, which is the approximate level for value stocks.

In order to generate a path for the underlying factor loading, one could generate series of disturbances u_t. However, it is very unlikely that we would obtain a series of factor loadings that is decreasing in a similar fashion to the observed estimates of beta for value stocks. Hence, given that our goal is to explain the effect on the value premium of a particular decreasing realization of the factor loading process, we need to adopt a different strategy.

equation (2.25) in the model tells us that the theoretical value of the estimated beta is equal to the underlying factor loading. This result suggests the simulation strategy of setting the realized process for b_t equal to the estimates of beta. These estimates are obtained from five-year rolling window regressions with one-month increments (for details, see Franzoni, 2002). In this way, we let the data tell us what the realizations of u_t are. Investors learn about a realization of factor loading that is clearly decreasing. Accordingly, a sensible choice for σ_u^2 comes from looking at the residuals of a regression, in which we fit an AR1 process to the estimated betas.

A separate issue is the choice of F, which is a crucial parameter in the learning process, and as such deserves more care. The estimate of F from an AR1 process for the betas is 0.99, and a Dickey-Fuller test does not reject the hypothesis of a unit root. However, the estimated F could be far from the true underlying parameter, as it could be strongly affected by the overlapping windows that we use in the estimation of beta. Also, it can be different from what investors expect F to be, which is what matters in the learning process. Therefore, we carry out our simulations for different values of F. In particular, we will let F vary between 0.9 and 1. Values of F below 0.9, do not affect the results.

The main updating equation is given in (2.27). Also, the mean squared forecast error

evolves according to

$$P_{t+1|t} = (1 - K(x_t) x_t)^2 P_{t|t-1} + K(x_t)^2 \sigma_\varepsilon^2 + \sigma_u^2. \qquad (2.29)$$

We take as initial condition for $P_{t+1|t}$ the squared standard error of the estimate of beta in the first window of the sample. The initial condition for b_t^e is equal to the estimate of beta in the same window. The choice of the initial conditions turns out not to be crucial.

Hence, equations (2.27) and (2.29) produce investors' expectation of the factor loading, which is then used in equation (2.15), along with randomly generated ε's, to create a series of monthly portfolio returns. The simulated returns are then regressed on the market excess return to obtain a series of alphas that will be compared with the alphas coming from the CAPM rolling window regressions on real data.

The first case we consider is when $F = 1$, i.e. the factor loading follows a random walk. Figure 2-2 plots the average series of expected loadings b_t^e from 250 repetitions. The graph also plots the estimates of beta from real data, which coincide with the true underlying factor b_{t+1}. As a result of learning, the expected loading is above the true loading for most of the sample, and in particular when b_{t+1} is decreasing[3].

According to equation (2.28) in our model, this situation should cause the econometrician to observe a premium relative to the measured riskiness of the portfolio. However, when $F = 1$ investors update their beliefs very quickly, so that the expected loading is not far enough from the true loading as to generate a premium. Consequently, while the series of alpha from real data is significantly different from zero starting from the late sixties, over the same interval of time the alpha from simulated data is significant only at the end of the

---

[3]For this parameter choice, as well as for all the others, it is the case that the beta estimated on simulated data is decreasing in a similar fashion to the underlying factor loading, and to the beta estimated on real data.

nineties. Figure $2-3$ graphs this situation.

Table 2.1 provides summary statistics on the ability of the simulated data to explain the observed mispricing out of 250 repetitions. In particular, we report the percentage of average simulated alphas that are significant in the windows when also the real alphas are significant (Column 1), the average fraction of mispricing explained in those occurrences (Column 2), and the explained share of the intercept in the regression on the 1963-2000 sample (Column 3) [4]. Consistent with the evidence in Figure $2-3$, the first row of the table shows that when F equals 1 the average alpha estimated on simulated data is significant in only 5% of the windows in which the alpha from real data is significant (Column 1).

To generate significant premia we need the expectation of the loading to be sufficiently distant from the truth.

This is more likely to be the case if F is smaller than 1. When the process for the loading is mean reverting, investors' expectations give weight also to the long run mean B, which instead does not matter if F equals 1. Hence, depending on the level of B, the expected factor loading can be well above the true one over most of the sample.

For now we are assuming that investors know the value of B. This assumption can be justified by saying that investors have clear in their minds what the long run value should be, towards which the factor loading has to converge. Later on, we will let them learn about B as well. It is worth presenting the results with known B, in order to understand the simple effect of mean reversion on the speed of learning. The choice of B turns out to be crucial. Therefore, we carry out the simulations for different levels of B.

As one might expect, when F is below 1, the higher is B, the higher are the alphas that we manage to generate. The intuition is straightforward. When investors believe the factor

---

[4]We look at this sample, because also Fama and French's (1993) sample starts in 1963, and we have data until the year 2000.

loading is going to mean revert, their expectation of b_{t+1} tends to adhere to the long run mean. If B is high, then b_t^e tends to be on average higher than the realized factor loading. Hence, the econometrician observes higher premia.

On the other hand, a decrease in F moves the b_t^e series away from the b_{t+1} series, and closer to B. Therefore, if B lies mostly above the realized factor loading, the econometrician tends to observe positive and significant alphas. Instead, when B is low, a decrease in F has the opposite effect.

These results are displayed in Figures $2-4$ through 2-11, and in Table 2.1.

In Figures $2-8$ and $2-9$, for example, we constructed the series assuming $F = 0.97$ and setting $B = 1.27$, which corresponds to the average estimated beta for value stocks. The series of expected factor loadings appears to track less closely the true loading than in the case of random walk. Also, since the long run mean is high enough, the alphas tend to be significant starting from the sixties, although still far from the real ones. In fact, from Table 2.1 we infer that although the series of simulated alphas is significant 81% of the times in which the real alphas are significant (Column 1), the average fraction of the observed mispricing that we can explain is only 21% (Column 2).

In order to obtain levels of mispricing that are closer the ones that are observed, we need to assume that investors believe the long run mean to be higher. For example, Figure 2-11 shows that with $B = 1.5$ and $F = 0.97$ the alphas estimated on simulated data are much closer to real ones, over the relevant period that starts in the late sixties. Moreover, Table 2.1 tells us that for this choice of parameters the simulated alphas are significant 100% of the times when the real alphas are significant, and in those occurrences the ratio of simulated to real alphas is on average 39%. Also, the estimated premium in the 1963-2000 subsample is 0.30% monthly for the simulated data, which represents 68% of the 0.44%

monthly premium estimated over the same period on real data (Column 3).

A value of B equal to 1.5, although significantly higher than the average estimated beta, is not implausible considering that the historical peak of the estimated beta for value stocks is about 2.2, which was reached in the early forties. Moreover, a level of F below 1 is totally justifiable, as it is hard to believe that factor loadings will drift away without ever reverting to some long run mean.

In the next section, we will let investors learn about the long run mean, and we will be able to explain sizeable amounts of the observed mispricing, almost irrespective of the assumed value for B.

2.4 Learning about the long run mean

2.4.1 The model

To work under more realistic assumptions about investors' information set, we modify the original model and let them learn about the long run mean of the factor loading as well. Most of the set-up of the model in Section 2.2.1 still applies here. In particular, the factor loading still evolves according to equation (2.2), but now B is unknown. So, the state equation in the Kalman Filter, which describes the evolution of the unobservable variables, is actually a system of two equations

$$\xi_{t+1} = \tilde{F}\xi_t + \nu_{t+1} \tag{2.30}$$

where

$$\xi_t = \begin{bmatrix} B \\ b_t \end{bmatrix}$$

$$\tilde{F} = \begin{bmatrix} 1 & 0 \\ 1-F & F \end{bmatrix}$$

$$\nu_t = \begin{bmatrix} 0 \\ u_t \end{bmatrix}.$$

and the usual distributional assumption on u_t applies[5].

By taking expectations of the system in (2.30) under the subjective probability distribution, we can state that

$$E_t b^e_{t+1} = (1 - F) B^e_t + F b^e_t$$

$$E_t B^e_{t+1} = B^e_t \tag{2.31}$$

Using similar arguments to Section 2.2.1, in the appendix we prove that the new pricing function for asset i is

$$p_t = \frac{D}{r} - \frac{A\sigma^2}{(1+r-F)(1-AD)} \left(b^e_t + \frac{(1-F) B^e_t}{r} \right) \tag{2.32}$$

which is the same as in the previous model except that B is replaced with investors' expectation.

---

[5]Note that if F equals 1, the factor loading follows a random walk and learning about the long run mean B becomes irrelevant. In that case the model would coincide with the model in the previous section for the case where F equals 1.

Similarly, absolute excess returns are

$$R_{t+1} = x_{t+1}b_{t+1} + \varepsilon_{t+1} + \frac{A\sigma^2(1+r)}{(1+r-F)(1-AD)}\left(b_t^e + \frac{B_t^e(1-F)}{r}\right)$$
$$- \frac{A\sigma^2}{(1+r-F)(1-AD)}\left(b_{t+1}^e + \frac{B_{t+1}^e(1-F)}{r}\right) \qquad (2.33)$$

So that expected returns are the same as in the previous model

$$E_t R_{t+1} = \frac{A\sigma^2}{(1-AD)}b_t^e \qquad (2.34)$$

Along the lines of Section 2.2.1 we can prove that the CAPM holds in this model as well, and the market beta computed under the subjective probability distribution β_t is equal to the expected factor loading b_t^e, as in equation (2.20).

We now turn to considering the econometrician's point of view, i.e. we look at the moments of the objective probability distribution, where b_t and B are treated as known. Again, we want to provide an expression for the population analog of the estimated mispricing (equation (2.22)). As in the previous section, it can be shown that the population equivalent of the estimated beta $\hat{\beta}_t$ coincides with the true factor loading b_{t+1}.

In the appendix we use this result to derive the estimated mispricing, which turns out to be

$$\hat{\alpha}_t = (b_t^e - b_{t+1})\left(1 + \frac{rF\pi_{2,2} + (1-F)(1+r)\pi_{1,2}}{r(1+r-F)}H_t\right)E_t R_{t+1}^m \qquad (2.35)$$

where

$$H_t = E_t^o\left(\frac{x_{t+1}^2}{\pi_{2,2}x_{t+1}^2 + \sigma_\varepsilon^2}\right)$$

and $\pi_{i,j}$ is the element on the i-th row and j-th column of $P_{t+1|t}$. In particular

105

$$\pi_{1,2} = E\left((B - B_t^e)(b_{t+1} - b_t^e)\right)$$

$$\pi_{2,2} = E\left(b_{t+1} - b_t^e\right)^2.$$

So, it is still the case that the mispricing is proportional to the difference between the expected and the true factor loading. The sign of the factor of proportionality depends on $\pi_{2,2}$, which is positive, and $\pi_{1,2}$, which can take either sign. The simulations that we present below, will show that, as in the model with known B, the mispricing is clearly an increasing function of the difference between b_t^e and b_{t+1}. This evidence suggests that when the factor loading is trending down, it is the case that the error in forecasting the long run mean covaries positively with the error in forecasting the level of the factor loading.

2.4.2 Simulations

The simulation strategy for this model is analogous to the one we adopted in Section 2.3. The only difference now is that we need generate two forecasts, one for b_{t+1} and one for B, using the following updating formula

$$\xi_{t+1}^e = F\xi_t^e + K(\tilde{x}_{t+1})\left(\varepsilon_{t+1} + \tilde{x}_{t+1}'\left(\xi_{t+1} - \xi_t^e\right)\right) \tag{2.36}$$

where

$$\tilde{x}_{t+1} = \begin{bmatrix} 0 \\ \\ x_{t+1} \end{bmatrix}$$

$$K\left(\tilde{x}_{t+1}\right) = \frac{\tilde{F}P_{t+1|t}\tilde{x}_{t+1}}{\tilde{x}_{t+1}'P_{t+1|t}\tilde{x}_{t+1} + \sigma_\varepsilon^2}.$$

Also, $P_{t+1|t}$ is a 2×2 matrix defined as

$$P_{t+1|t} = E\left((\xi_{t+1} - \xi_t^e)(\xi_{t+1} - \xi_t^e)'\right) \tag{2.37}$$

and its updating formula is given in Hamilton (1994).

We need to set initial conditions for both b_t^e and B_t^e. We choose as initial condition for b_t^e the beta estimated in the first five-year window of the sample, but this choice is not crucial for our results The initial condition for B_t^e is not influential as long as it is high enough for the path of expected B to be mostly decreasing. Therefore, we set it to 1.5, but 1.27 would not affect the results. The choice of 1 for B_0^e, instead, would strongly weaken the significance of our results. The initial conditions for $P_{t+1|t}$ turns out not to matter. Hence, we set every element of $P_{t+1|t}$ equal to the value of σ_u^2 in the simulations of the previous model[6].

The general result from the simulation of the model with unknown B is that the series of expected factor loadings decreases more slowly relative to the model with known B. Compare, for example, Figures 2-8 and 2-12, where the choice of parameters is $F = .97$ and $B = 1.27$. In the latter graph, where investors learn about B, the decline in the b_t^e is much slower, as this series is attracted towards the B_t^e series, which in turn declines smoothly. This behavior is fairly insensitive to the choice of B and F, as one can see from the comparison of Figure 2-12 ($F = .97$, $B = 1.27$) and Figure 2-14 ($F = .97$, $B = 1$).

The intuition for this result is again simple. In both models, when F is below 1, investors' expectation of the factor loading tends to adhere to the long run mean. However, if they do not know B, also the expectation of B is likely to be higher than the true value, and this contributes to keep the b_t^e series well above b_{t+1}. Moreover, the fact that investors do not observe the long run mean B, makes their expectation of b_{t+1} fairly insensitive to the true level of B.

As one might expect, the consequence of this behavior is that the econometrician ob-

---

[6]Even the choice of negative values for the off-diagonal elements of $P_{t+1|t}$ would not affect the results.

serves premia that are overall higher than the ones we presented in Section 2.3. This fact can be inferred, for example, from the comparison of Figures 2-9 and 2-13.

Another way to assess the relative of performance of the model with unknown B, is by looking at Table 2.2, which replicates Table 2.1 in the new set-up. The superiority of the model with unknown B is overwhelming. For almost any choice of parameters, learning about the long run mean produces more significant alphas, and explains a higher fraction of the observed value premium.

Not only does this model perform well in relative terms, but also it has a striking ability of explaining a large share of the observed mispricing. Look, for example, at the entries of Table 2.2 for which F equals 0.90. The average simulated alpha is significantly positive 100% of the times when the real alpha is significant, irrespective of the underlying B. Also, the average ratio of these alphas that the simulated series can explain is always well above 40%. Finally, the fraction of the value premium in the 1963-2000 sample that we can account for is between 70% and 80%.

The performance of the model is good also for higher levels of F. In fact the fraction of the value premium in the 1963-2000 sample explained by the simulated data is never below 36% (Table 2.2, $F = .99$, $B = 1.27$).

In summary, the key to the success of our simulations is the combination of two elements: a mean-reverting process of the factor loading (F smaller than 1), and an unknown long run mean B. The first element seems to be justified on theoretical grounds by the belief that factor loadings do not diverge in the long run. The assumption of an unknown long run mean is totally consistent with an environment, where factor loadings undergo dramatic changes over time, and investors have to learn about the underlying process[7].

---

[7]It has been remarked that a behavior of the series b_t^e, such as the one described in this section, could be generated by a model with adaptive expectations. To this purpose we have developed a variation of our

2.5 Concluding remarks

The assumption that investors know the variance-covariance structure of asset returns is especially implausible if this structure varies over time. Dramatic changes in factor loadings, like the ones documented by Franzoni (2002) for value and small stocks, can cause investors to continuously revise their expectations of the riskiness of an asset. Depending on the speed of learning, and on the amount of noise in returns, these expectations can diverge significantly from the true level. When the risk premia required by investors do not reflect the underlying riskiness, the econometrician can observe mispricing. Hence, parameter uncertainty may play a role in the explanation of some the CAPM failures.

We develop an equilibrium model with learning about time-varying factor loadings, which captures this intuition. In our model the CAPM holds from investors' point of view, but the econometrician observes mispricing, when the expected factor loading diverges from the actual beta. In particular, after a decreasing realization of the loading process, such as the one that characterized value stocks, the expected loading tends to be above the true one. As a consequence, investors require a risk premium that appears excessive through the econometrician's glasses, and which can cause the rejection of the CAPM.

---

set-up, in which the updating formula is simply

$$b_t^e = \lambda b_{t-1}^e + (1 - \lambda) b_t \tag{2.38}$$

where λ is the parameter that determines the speed of learning. The details of this model can be provided upon request. We simulated the model for several values of λ. To obtain a series for investors' beliefs, which generates value premia comparable to the ones in the above simulations, we need a value of λ between 0.99 and 1. The reason why we prefer our models of learning, is that adaptive expectations are difficult to give a rational foundation in this environment. In fact, adaptive expectations are rationally justified if the underlying process for the factor loading can be represented as the sum of a random walk and a MA1 process (see, e.g., Hamilton, 1994). Under these assumptions, for λ to be positive, it has to be the case that the first difference of the factor loading is negatively autocorrelated (intuitively, we give positive weight to our past beliefs, if the changes in the variable that we learn about tend to revert themselves). However, this is not the case for the changes in estimated betas, for which the measured autocorrelation tends to be positive. Hence, the optimal level of λ should be smaller than zero, therefore very far from 0.99. For this reason the adaptive expectation model is deprived of a rational foundation, and our fully rational model is preferable.

The simulation of our model points out that in order to produce levels of mispricing that are close to the 'value premium', investors need not to update their beliefs too quickly. We can achieve this result if the factor loading follows a mean reverting process, and if investors have to learn about the long run mean of this process as well. The first condition assures that the expectations of the loading is fairly stable in the face of new information, as it tends to adhere to the long run mean. The second condition causes the belief of the loading to decrease slowly, while investors learn about the long run mean as well. Neither condition seems implausible: factors loading are not likely to wander around without a hinge, but investors may not know what value they will revert to, especially after dramatic changes.

So, under these assumptions, our model can generate statistically and economically significant premia for value stocks. In particular, for some parameter values, we can explain between 70% and 80% of the value premium in the 1963-2000 sample.

The immediate implication of our results is in line with the conclusions from the conditional CAPM literature (see, e.g., Lettau and Ludvigson, 2001). The tests of the CAPM that do not account for time-varying factor loadings are misspecified.

However, scaling the factors with state variables is not enough, because it does not take into account investors' inference problem. Since the CAPM holds only under the subjective probability distribution, to correctly test the model the econometrician would have to replicate the Bayesian-updating process of rational investors, as argued by Lewellen and Shanken (2002).

This consideration suggests an ambitious direction for future research. Parameter uncertainty is another source of risk to which investors are subject. In particular, estimation risk is more serious for those portfolios, whose factor loadings tend to vary more dramatically. Hence, the challenge is to build a multifactor pricing model that takes into account

110

estimation risk, possibly as a function of how volatile factor loadings have been in the past.

As a preliminary step we will assess whether parameter uncertainty is related to the good performance of the Fama and French three-factor model. More concretely, we intend to look for a correlation between the B/M factor and the volatility of the factor loading for value stocks. A significant correlation will provide additional evidence in favor of learning as an explanation of the 'value premium'.

Appendix

Model with known B

In this appendix we derive the pricing function for the model with known B. A guess for the pricing function is

$$p_t = B_0 + B_1 b_t^e. \tag{A-1}$$

Replacing this pricing function into the pricing equation (2.12), we obtain

$$B_0 + B_1 b_t^e + D - (1+r) B_0 - (1+r) B_1 b_t^e = ADB_0 + ADB_1 b_t^e + AD^2$$

$$-AD(1+r) B_0 - AD(1+r) B_1 b_t^e + Ab_t^e \sigma^2.$$

This equation has to hold for all values of b_t^e, so the coefficients are determined by setting the constant and the coefficient on b_t^e equal to zero

$$B_0 + D - (1+r) B_0 = ADB_0 + AD^2 - AD(1+r) B_0.$$

The solution to this is

$$B_0 = \frac{D}{r}.$$

The coefficients in front of b_t^e give the following equation

$$B_1 - (1+r) B_1 = ADB_1 - AD(1+r) B_1 + A\sigma^2$$

with the solution

$$B_1 = -\frac{A\sigma^2}{r(1 - AD)}.$$

So, replacing the coefficients into the pricing function $(A-1)$

$$p_t = \frac{D}{r} - \frac{A\sigma^2}{r(1 - AD)} b_t^e. \qquad \text{(A-2)}$$

In order to check that CAPM holds under the investors information set (the "subjective" probability distribution), the covariance of individual returns with the market return must be computed

$$
\begin{aligned}
&Cov_t\left(R_{t+1}^i, R_{t+1}^m\right) \\
&= Cov_t\left(x_{t+1}b_{t+1} + \varepsilon_{t+1} + \frac{A\sigma^2\left((1+r)b_t^e - b_{t+1}^e\right)}{r(1 - AD)}, x_{t+1} + \frac{A\sigma^2}{1 - AD}\right) \\
&= \sigma^2 b_t^e - \frac{A\sigma^2}{r(1 - AD)} Cov_t\left(b_{t+1}^e, x_{t+1}\right) \\
&= \sigma^2 b_t^e - \frac{A\sigma^2}{r(1 - AD)} Cov_t\left(b_t^e + K(x_{t+1})x_{t+1}\left(b_{t+1} - b_t^e\right), x_{t+1}\right) \\
&= \sigma^2 b_t^e - \frac{A\sigma^2}{r(1 - AD)} Cov_t\left(K(x_{t+1})x_{t+1}, x_{t+1}\right) E_t\left(b_{t+1} - b_t^e\right) \\
&= \sigma^2 b_t^e.
\end{aligned}
$$

Therefore, the covariance is simply

$$Cov_t\left(R_{t+1}^i, R_{t+1}^m\right) = \sigma^2 b_t^e \qquad \text{(A-3)}$$

where the third equality follows from replacing the updating rule for the conditional expectation of b_{t+1}, and the fourth one from the assumption that x_{t+1} does not contain any information on b. We introduced the following notation

$$K(x_{t+1}) = \frac{FP_{t+1|t}x_{t+1}}{x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2} \qquad \text{(A-4)}$$

and $P_{t+1|t}$ is the mean square forecast error for b_{t+1}.

The variance of the market return is simply:

$$Var_t\left(R_{t+1}^m\right) = Var_t\left(x_{t+1} - \frac{\theta - A\lambda^2}{1 - AD}\right) = \sigma^2 \tag{A-5}$$

In order to compute the average mispricing, we need to solve for the CAPM relationship under the econometricians information set. The econometrician can only observe realized data, which is, asymptotically, like conditioning on the true evolution of b. The covariance of individual returns with the market under the econometrician's "objective" probability distribution is therefore

$$
\begin{aligned}
& Cov_t^o\left(R_{t+1}^i, R_{t+1}^m\right) \\
= \; & Cov_t^o\left(x_{t+1}b_{t+1} + \varepsilon_{t+1} + \frac{A\sigma^2\left((1+r)b_t^e - b_{t+1}^e\right)}{r\left(1 - AD\right)}, x_{t+1} + \frac{A\sigma^2}{1 - AD}\right) \\
= \; & b_{t+1}\sigma^2 - \frac{A\sigma^2}{r\left(1 - AD\right)}Cov_t^o\left(b_{t+1}^e, x_{t+1}\right) \\
= \; & b_{t+1}\sigma^2 - \frac{A\sigma^2}{r\left(1 - AD\right)}Cov_t^o\left(b_t^e + K(x_{t+1})x_{t+1}\left(b_{t+1} - b_t^e\right), x_{t+1}\right) \\
= \; & b_{t+1}\sigma^2 - \frac{A\sigma^2\left(b_{t+1} - b_t^e\right)}{r\left(1 - AD\right)}Cov_t^o\left(K(x_{t+1})x_{t+1}, x_{t+1}\right) \\
= \; & b_{t+1}\sigma^2 - \frac{A\sigma^2\left(b_{t+1} - b_t^e\right)}{r\left(1 - AD\right)}Cov_t^o\left(\frac{P_{t+1|t}x_{t+1}^2}{x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2}, x_{t+1}\right).
\end{aligned}
$$

Assuming that x_{t+1} is normally distributed, we can apply Stein's lemma to the above covariance. If $\tilde{x}$ and $\tilde{y}$ are normally distributed, then according to Stein's lemma

$$Cov(f(\tilde{x}), \tilde{y}) = E(f'(\tilde{x}))Cov(\tilde{x}, \tilde{y}).$$

In particular, we consider the first term in the covariance as a function of x_{t+1}, so that the the covariance boils down to the variance of x_{t+1} times the expectation of the derivative:

$$
\begin{aligned}
& Cov_t^o\left(R_{t+1}^i, R_{t+1}^m\right) \\
= \; & b\sigma^2 - \frac{A\sigma^4\left(b - b_t^e\right)}{r\left(1 - AD\right)}C_t
\end{aligned}
$$

where

$$C_t = E_t^o \left(\frac{2x_{t+1}}{\left(x_{t+1}^2 P_{t+1|t} + \sigma_\varepsilon^2\right)^2} \right) = 0$$

because of the symmetry of the distribution of x_{t+1}. Indeed, the assumption of Normality

is not even required. For the covariance term in the expression for $\hat{\beta}_t$ to be zero, we just

need the symmetry of the distribution of x_{t+1}.

The objective covariance term therefore boils down to

$$Cov_t^o \left(R_{t+1}^i, R_{t+1}^m \right) = b_{t+1}\sigma^2. \tag{A-6}$$

Model with unknown B

We now derive the pricing function for the model with learning about B. A guess for the

pricing function is

$$p_t = B_0 + B_1 b_t^e + B_2 B_t^e. \tag{A-7}$$

Replacing this pricing function into the pricing equation (2.12), we obtain

$$-rB_0 + B_1 \left(1 - AD\right) \left(B_t^e \left(1 - F\right) + F b_t^e\right) + B_2 B_t^e + D - \left(1 + r\right) B_1 b_t^e - \left(1 + r\right) B_2 B_t^e$$

$$= ADB_0 + ADB_2 B_t^e + AD^2 - AD \left(1 + r\right) B_0 - AD \left(1 + r\right) B_1 b_t^e - AD \left(1 + r\right) B_2 B_t^e + Ab_t^e \sigma^2$$

Using the method of undetermined coefficients we now get three equations in three

unknowns. The first equation is

$$B_0 + D - \left(1 + r\right) B_0 = ADB_0 + AD^2 - AD \left(1 + r\right) B_0$$

which has solution

$$B_0 = \frac{D}{r}.$$

The second equation is

$$B_1 (1 - AD) F - (1 + r) B_1 = -AD (1 + r) B_1 + A\sigma^2$$

and the solution is

$$B_1 = -\frac{A\sigma^2}{(1 + r - F)(1 - AD)}.$$

The third equation is

$$B_1 (1 - AD)(1 - F) + B_2 - (1 + r) B_2 = ADB_2 - AD (1 + r) B_2$$

with solution

$$B_2 = (1 - F)\frac{B_1}{r}.$$

So, replacing the coefficients into the pricing function $(A-7)$

$$p_t = \frac{D}{r} - \frac{A\sigma^2}{(1 + r - F)(1 - AD)} \left(b_t^e + \frac{(1 - F) B_t^e}{r} \right). \tag{A-8}$$

In order to check that CAPM holds under the investors information set (the "subjective" probability distribution), the covariance of individual returns with the market return must be computed

We now want to prove the expression for the mispricing given in equation (2.35). Since the population equivalent of the estimated beta $\hat{\beta}_t$ coincides with the true factor loading b_{t+1}, we just need to compute the difference between objective and subjective expected returns. To this purpose we need to use the expression for returns in equation (2.33), and the updating formula for the Kalman system given in (2.36). Then, the difference in

expected returns becomes

$$E_t^o R_{t+1} - ER_{t+1}$$

$$= \frac{A\sigma^2}{(1+r-F)(1-AD)} \left((1+r)\left(b_t^e + \frac{B_t^e(1-F)}{r} \right) - \left(E_t^o b_{t+1}^e + \frac{E_t^o B_{t+1}^e (1-F)}{r} \right) \right)$$

$$- \frac{A\sigma^2}{1-AD} b_t^e$$

$$= \frac{A\sigma^2}{(1+r-F)(1-AD)} \left((1+r)\left(b_t^e + \frac{B_t^e(1-F)}{r} \right) - \left((1-F)B_t^e + Fb_t^e + \frac{(1-F)B_t^e}{r} \right) \right)$$

$$- \frac{A\sigma^2}{(1+r-F)(1-AD)} \left(\left(\iota_2' + \frac{(1-F)}{r} \iota_1' \right) E_t^o \left(K(\tilde{x}_{t+1}) \tilde{x}_{t+1}' \right) (\xi_{t+1} - \xi_t^e) \right) - \frac{A\sigma^2}{1-AD} b_t^e$$

$$= \frac{A\sigma^2 \left((1+r-F) b_t^e - \left(\iota_2' + \frac{(1-F)}{r} \iota_1' \right) E_t^o \left(K(\tilde{x}_{t+1}) \tilde{x}_{t+1}' \right) (\xi_{t+1} - \xi_t^e) \right)}{(1+r-F)(1-AD)} - \frac{A\sigma^2}{1-AD} b_t^e$$

$$= \frac{A\sigma^2}{(1+r-F)(1-AD)} \left(\iota_2' + \frac{(1-F)}{r} \iota_1' \right) E_t^o \left(K(\tilde{x}_{t+1}) \tilde{x}_{t+1}' \right) (\xi_t^e - \xi_{t+1})$$

where we have introduced the notation

$$\iota_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \iota_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

while $\tilde{x}_{t+1}$ and $K(\tilde{x}_{t+1})$ are defined in the text.

Let $\pi_{i,j}$ be the element on the i-th row and j-th column of $P_{t+1|t}$. Then, given the above definitions, we can write

$$E_t^o R_{t+1} - ER_{t+1} = \frac{A\sigma^2 (b_t^e - b_{t+1})}{(1+r-F)(1-AD)} \left(F\pi_{2,2} + \frac{(1-F)(1+r)}{r} \pi_{1,2} \right) E_t^o \left(\frac{x_{t+1}^2}{\pi_{2,2} x_{t+1}^2 + \sigma_\varepsilon^2} \right)$$

where the expectation is positive because $\pi_{2,2}$ is the mean squared error for the forecast of b_{t+1}.

Hence, the expression for mispricing becomes

117

$$
\begin{aligned}
\hat{\alpha}_t &= E_t^o R_{t+1} - E_t R_{t+1} + (\beta_t - \hat{\beta}_t) E_t R_{t+1}^m \\[2mm]
&= \frac{(b_t^e - b_{t+1}) A \sigma^2}{(1 + r - F)(1 - AD)} \left(F \pi_{2,2} + \frac{(1 - F)(1 + r)}{r} \pi_{1,2} \right) E_t^o \left(\frac{x_{t+1}^2}{\pi_{2,2} x_{t+1}^2 + \sigma_\varepsilon^2} \right) \\[2mm]
&\quad + (b_t^e - b_{t+1}) E_t R_{t+1}^m \\[2mm]
&= (b_t^e - b_{t+1}) E_t R_{t+1}^m \left(1 + \frac{r F \pi_{2,2} + (1 - F)(1 + r) \pi_{1,2}}{r(1 + r - F)} E_t^o \left(\frac{x_{t+1}^2}{\pi_{2,2} x_{t+1}^2 + \sigma_\varepsilon^2} \right) \right)
\end{aligned}
$$

which coincides with equation (2.35) in the text.

Table 2.1: Summary statistics (known B). The table reports summary statistics on the average alhpa's estimated on the simulated data when B is known out of 250 repetitions. Column 1 reports the percentage of significant simulated alpha's in the times when the real alpha is also significantly different from zero. Both series are estimated with 10-year rolling window regressions of the excess returns of value stocks on the market excess return in the 1926-2000 sample. The significance of the simulated alpha is computed using the standard deviation across repetitions. Column 2 reports the average ratio between the simulated and real alpha in the windows when both are significant. Column 3 reports the ratio of the simulated alpha to the real alpha obtained from a single regression in the 1963-2000 sample. In that sample the real alpha is 0.44% monthly.

		(1) % signif.	(2) avg. ratio of α	(3) ratio of α ('63-'00)
$F = 1$		5	.23	.09
$F = .99$	$B = 1$	3	.24	.05
$F = .99$	$B = 1.27$	64	.15	.27
$F = .99$	$B = 1.5$	92	.23	.45
$F = .97$	$B = 1$	1	.51	0
$F = .97$	$B = 1.27$	81	.21	.37
$F = .97$	$B = 1.5$	100	.39	.68
$F = .95$	$B = 1$	1	.56	0
$F = .95$	$B = 1.27$	84	.23	.39
$F = .95$	$B = 1.5$	100	.41	.68
$F = .90$	$B = 1$	1	.66	0
$F = .90$	$B = 1.27$	92	.23	.44
$F = .90$	$B = 1.5$	100	.46	.80

Table 2.2: Summary statistics (unknown B). The table reports summary statistics on the average alhpa's estimated on the simulated data when B is unknown out of 250 repetitions. Column (1) reports the percentage of significant simulated alpha's in the times when the real alpha is also significantly different from zero. Both series are estimated with 10-year rolling window regressions of the excess returns of value stocks on the market excess return in the 1926-2000 sample. The significance of the simulated alpha is computed using the standard deviation across repetitions. Column 2 reports the average ratio between the simulated and real alpha in the windows when both are significant. Column 3 reports the ratio of the simulated alpha to the real alpha obtained from a single regression in the 1963-2000 sample. In that sample the real alpha is 0.44% monthly.

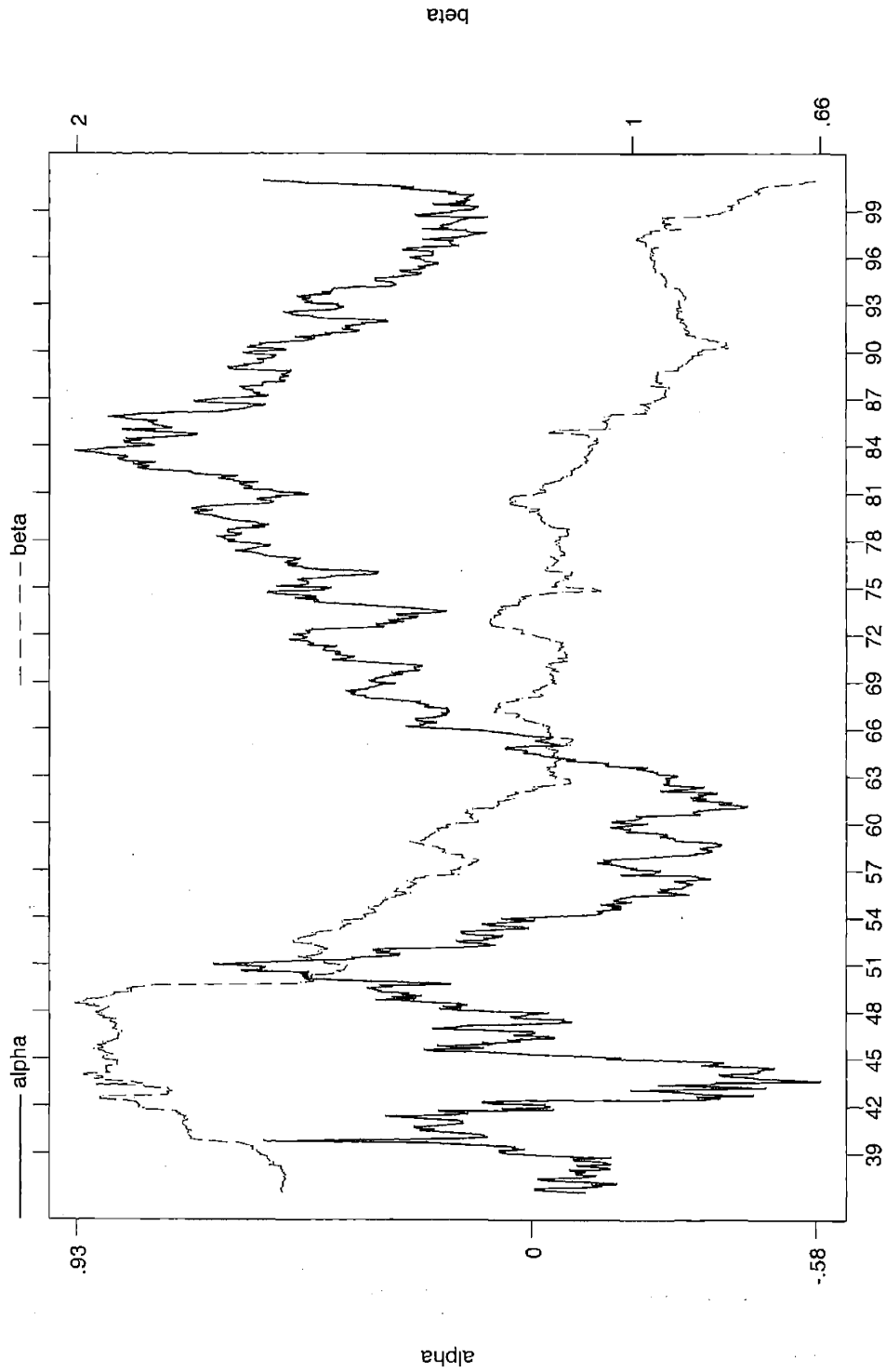|		(1)	(2)	(3)
		% signif.	avg. ratio of α	ratio of α ('63-'00)
$F = .99$	$B = 1$	67	.24	.39
$F = .99$	$B = 1.27$	74	.19	.36
$F = .99$	$B = 1.5$	75	.23	.41
$F = .97$	$B = 1$	100	.38	.60
$F = .97$	$B = 1.27$	100	.33	.61
$F = .97$	$B = 1.5$	100	.35	.65
$F = .95$	$B = 1$	100	.36	.70
$F = .95$	$B = 1.27$	100	.33	.62
$F = .95$	$B = 1.5$	100	.44	.73
$F = .90$	$B = 1$	100	.45	.78
$F = .90$	$B = 1.27$	100	.46	.73
$F = .90$	$B = 1.5$	100	.45	.79

Figure 2-1: Alpha and Beta. The figure plots series of estimates from 10-year rolling window CAPM regressions where the dependent variable is the excess return on B/M decile 10 portfolio. The series *alpha* (solid line, left scale) is the intercept in the CAPM regression. The series *beta* (dashed line, and right scale) is the beta in the CAPM regression.
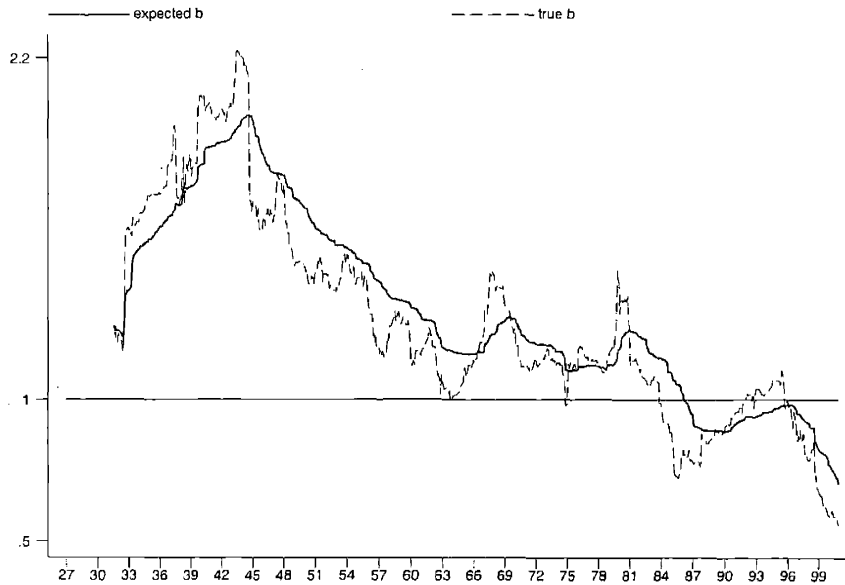
Figure 2-2: True and expected factor loadings, $F = 1$(**known B**). The figure plots the series of average expected betas (solid line) over 250 repetitions of the model with known B. The series of the true underlying factor loadings is also reported (dashed line).
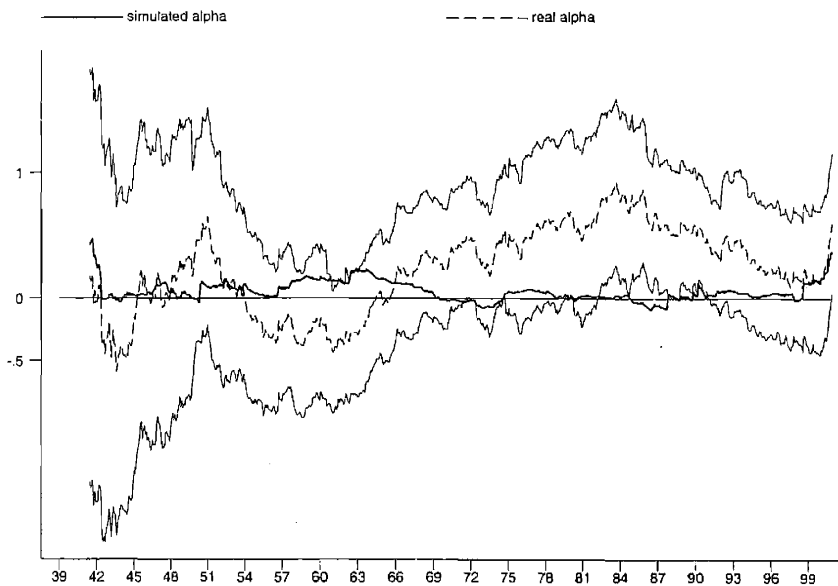

Figure 2-3: Simulated and real alpha, $F = 1$ (**known B**). The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with known B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.
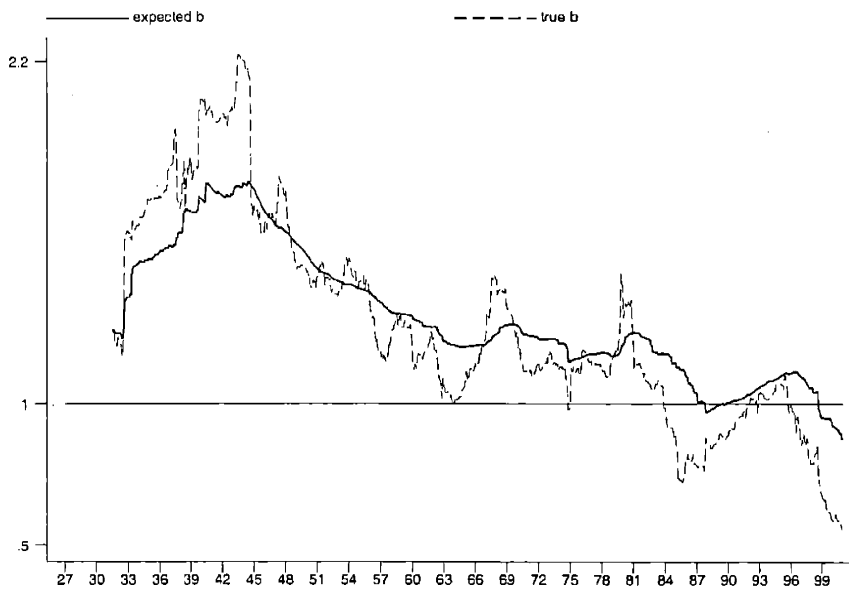
Figure 2-4: True and expected factor loadings, $F = .99$ $B = 1.27$ **(known B).** The figure plots the series of average expected betas (solid line) over 250 repetitions of the model with known B. The series of the true underlying factor loadings is also reported (dashed line).
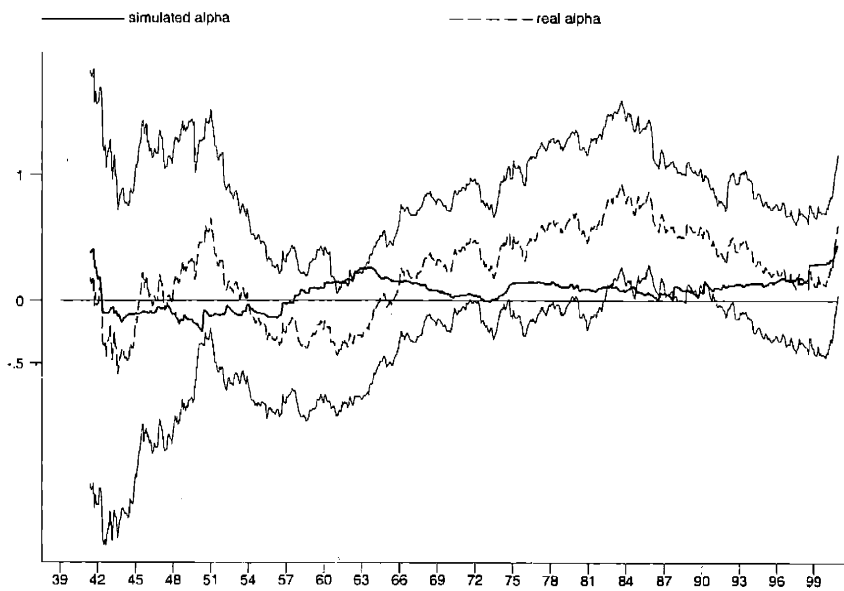

Figure 2-5: Simulated and real alpha, $F = .99$ $B = 1.27$ **(known B).** The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with known B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.
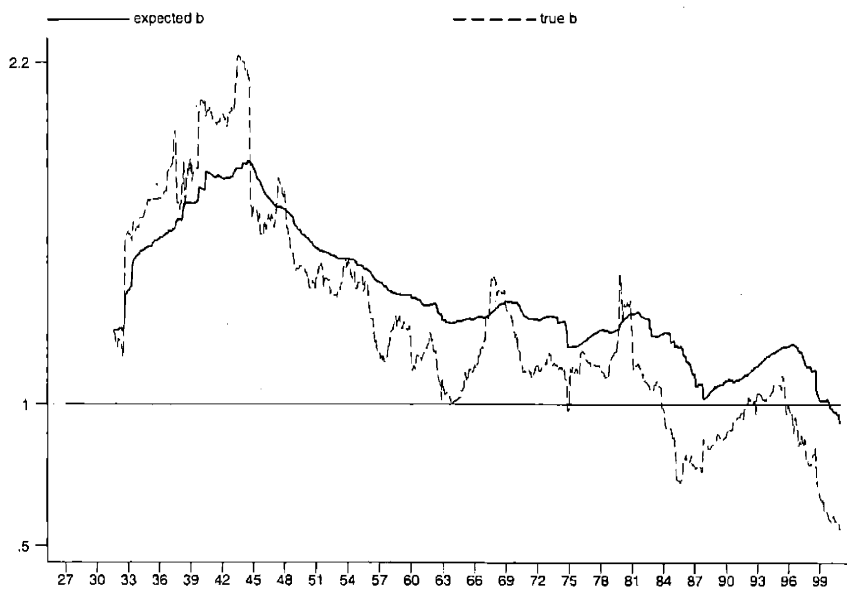
Figure 2-6: True and expected factor loadings, $F = .99$ $B = 1.5$ (known B). The figure plots the series of average expected betas (solid line) over 250 repetitions of the model with known B. The series of the true underlying factor loadings is also reported (dashed line).
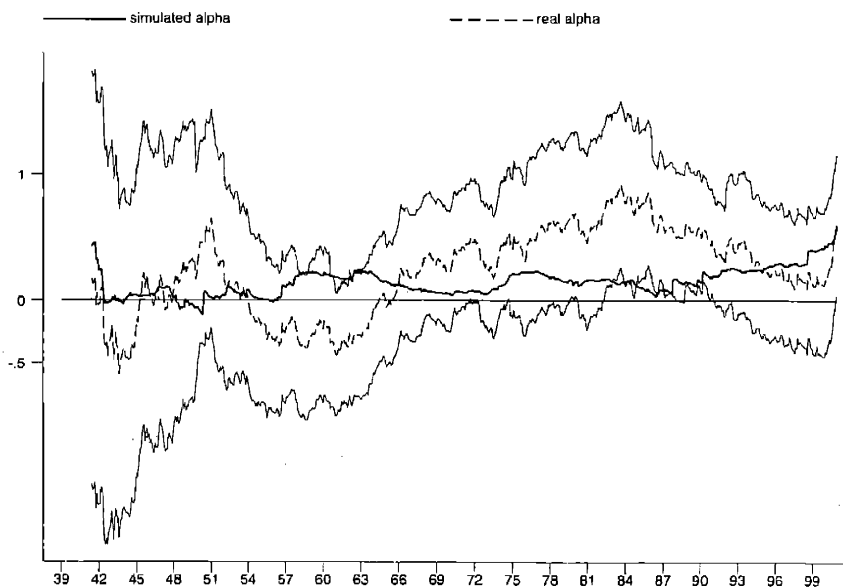

Figure 2-7: Simulated and real alpha, $F = .99$ $B = 1.5$ (known B). The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with known B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.

Figure 2-8: True and expected factor loadings, $F = .97$ $B = 1.27$ **(known B).** The figure plots the series of average expected betas (solid line) over 250 repetitions of the model with known B. The series of the true underlying factor loadings is also reported (dashed line).

Figure 2-9: Simulated and real alpha, $F = .97$ $B = 1.27$ **(known B).** The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with known B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.
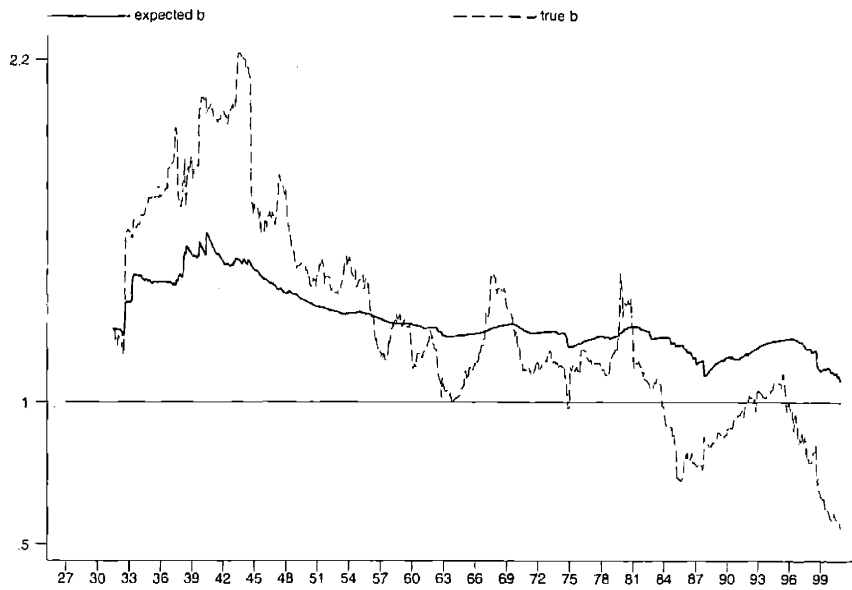
Figure 2-10: True and expected factor loadings, $F = .97$ $B = 1.5$ (known B). The figure plots the series of average expected betas (solid line) over 250 repetitions of the model with known B. The series of the true underlying factor loadings is also reported (dashed line).
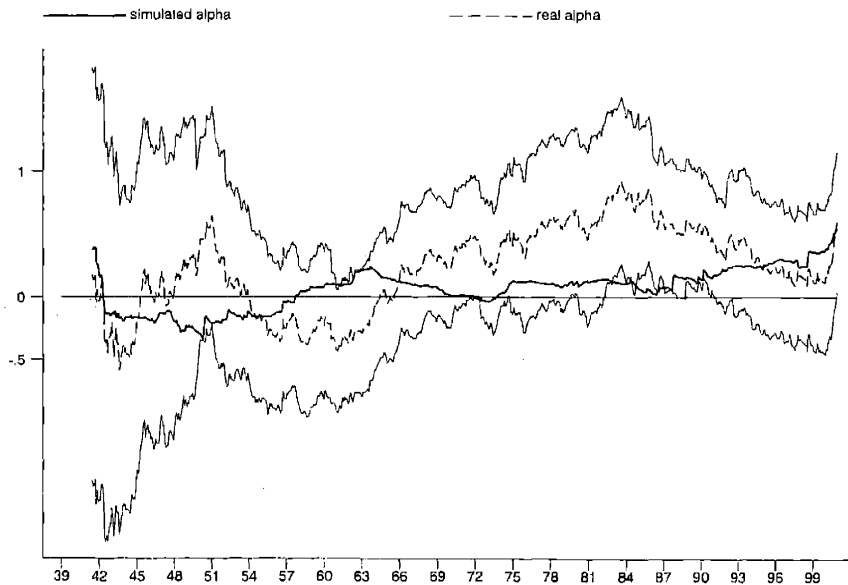

Figure 2-11: Simulated and real alpha, $F = .97$ $B = 1.5$ (known B). The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with known B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.
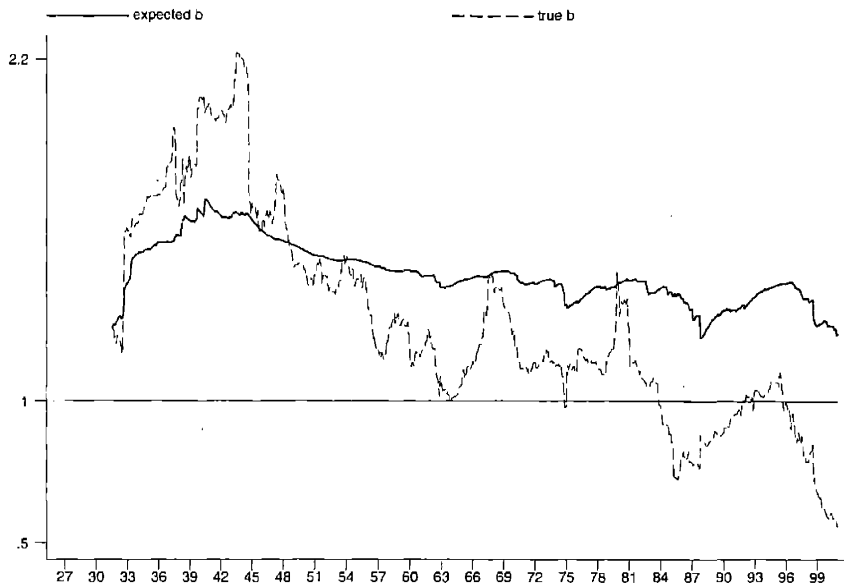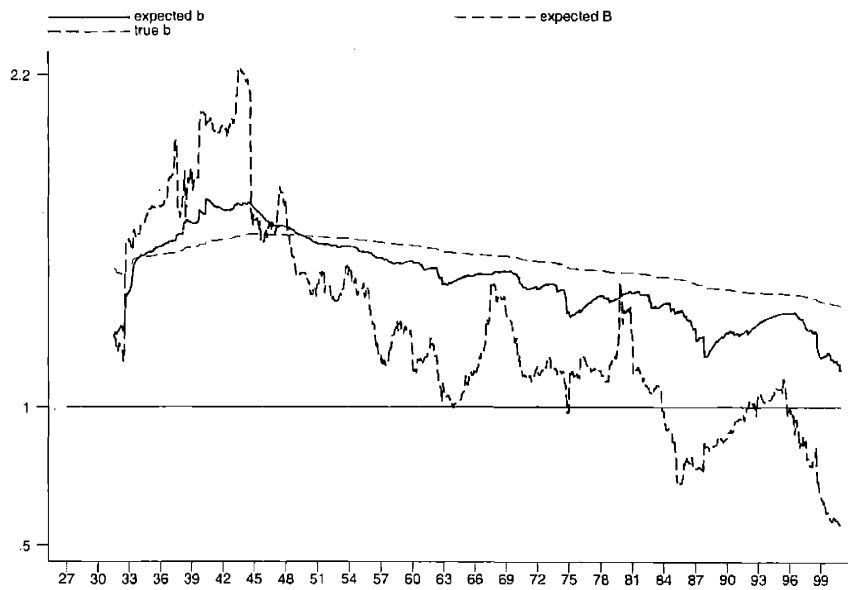
Figure 2-12: True and expected factor loadings, $F = .97$ $B = 1.27$ (unknown B). The figure plots the series of average expected betas (solid line) and B (thin dashed line) over 250 repetitions of the model with unknown B. The true factor loadings are also reported (thick dashed line).
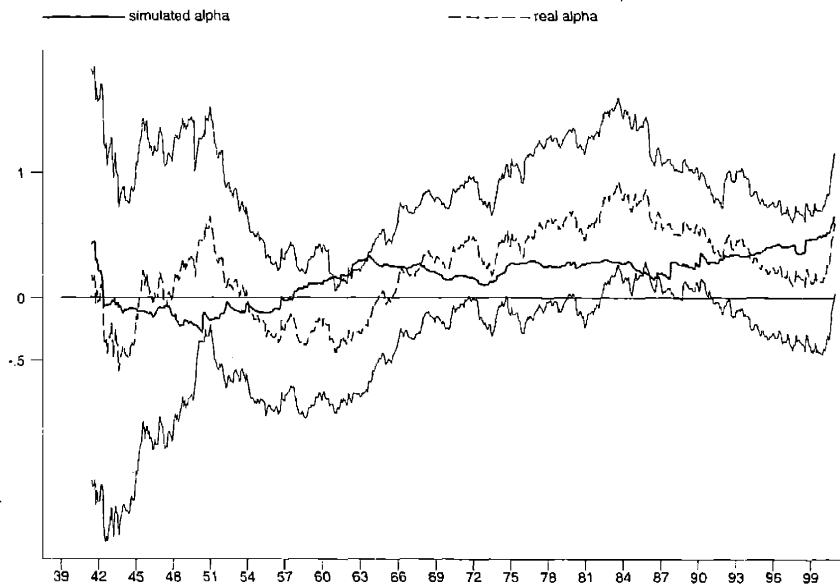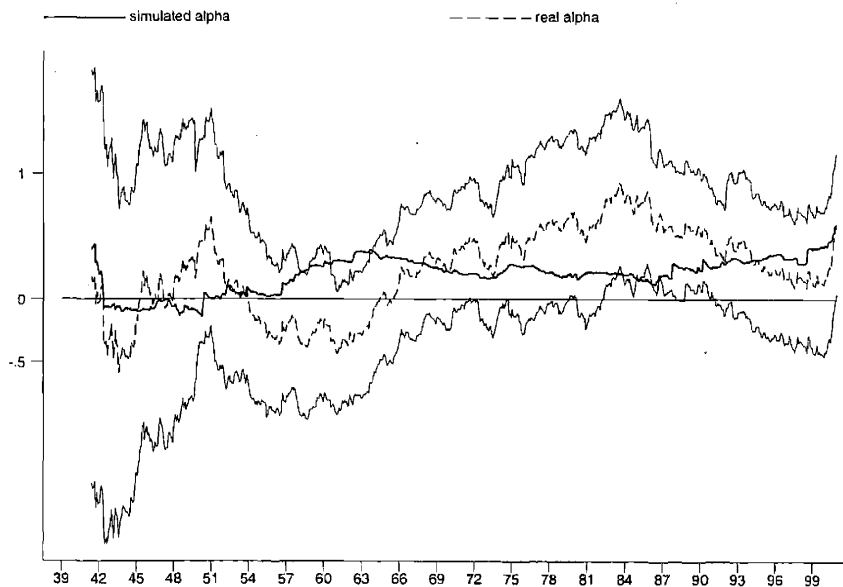

Figure 2-13: Simulated and real alpha, $F = .97$ $B = 1.27$ (unknown B). The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with unknown B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.
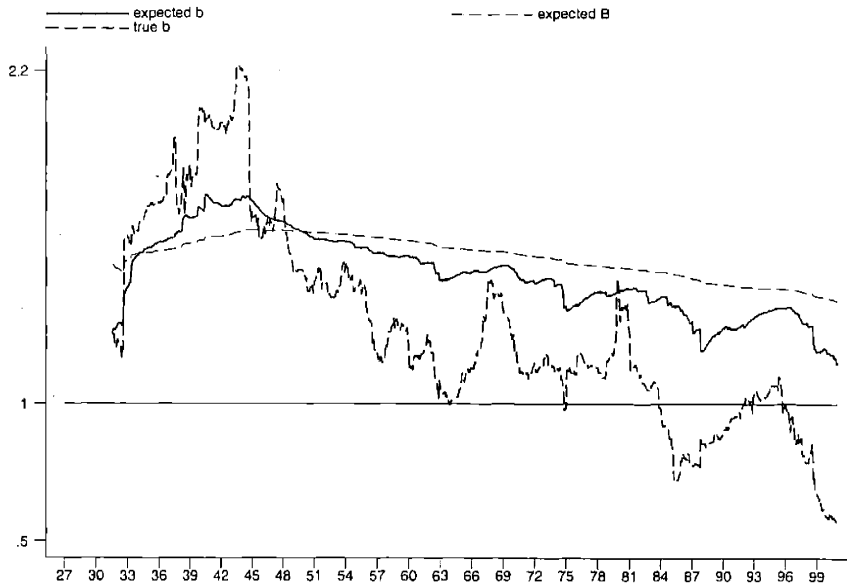
Figure 2-14: True and expected factor loadings, $F = .97$ $B = 1$ (unknown B). The figure plots the series of average expected betas (solid line) and B (thin dashed line) over 250 repetitions of the model with unknown B. The true factor loadings are also reported (thick dashed line).
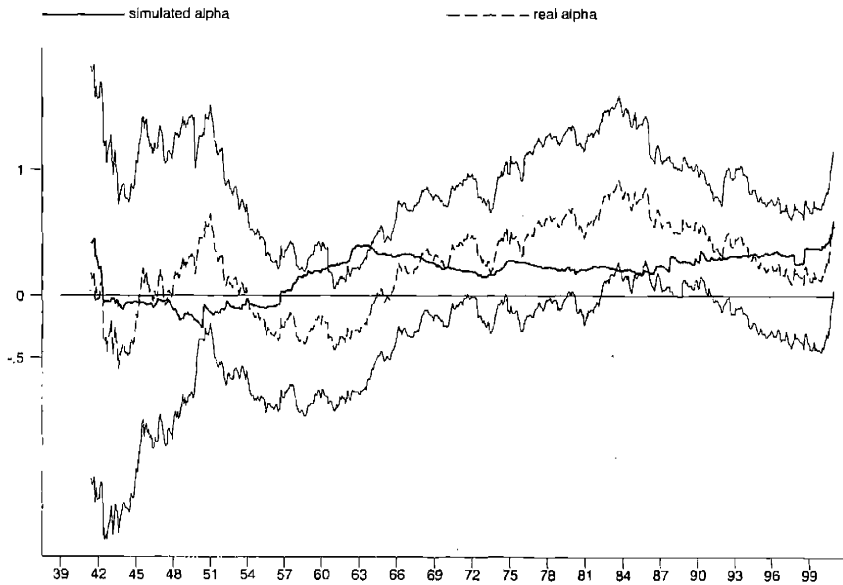
Figure 2-15: Simulated and real alpha, $F = .97$ $B = 1$ (unknown B). The figure plots the series of the average alpha (solid line) over 250 repetitions of the model with unknown B. This series is the average of the estimated alpha series from 10-year rolling window regressions on simulated data. The alpha series estimated on real data for value stocks is also reported (dashed line). The 95% confidence interval is plotted around the alpha from real data.

Bibliography

[1] Banz, R., 1981, "The Relationship between Return and Market Value of Common Stocks", *Journal of Financial Economics, 9*, pp. 3-18.

[2] Barberis, N., 2000, "Investing for the Long Run when Returns are Predictable", *Journal of Finance 55*, pp. 225-264.

[3] Basu, S., 1977, "The Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis", *Journal of Finance, 32*, pp. 663-682.

[4] Black, F., M. Jensen, and M. Scholes, 1972, "The capital asset pricing model: Some empirical tests", in M. Jensen ed., *Studies in the Theory of Capital Markets*, (Praeger, New York, NY).

[5] Braun, P., D. Nelson, and A. Sunier, 1995, "Good news, bad news, volatility and betas", *Journal of Finance, 50*, pp. 1575-1604.

[6] Breeden, D., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics, 7*, pp. 265-296.

[7] Brennan, M. J. and Y. Xia, 2001, "Stock Return Volatility and Equity Premium", *Journal of Monetary Economics 47*, pp. 249-283.

[8] Campbell, J., and R. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies*, 1, pp. 195-227.

[9] Daniel, K., and S. Titman, 1997, "Evidence on the characteristics of cross-sectional variation in stock returns", *Journal of Finance*, 52, pp. 1-33.

[10] Fama, E., and K. French, 1995, "Size and book-to-market factors in earnings and returns", *Journal of Finance*, 50, pp. 131-155.

[11] Fama, E., and K. French, 1996, "Multifactor explanations of asset pricing anomalies", *Journal of Finance*, 51, pp. 55-84.

[12] Fama, E., and K. French, 1992, "The cross-section of expected stock returns", *Journal of Finance*, 47, pp. 427-465.

[13] Fama, E., and K. French, 1993, "Common risk factors in the returns on stocks and bonds", *Journal of Financial Economics*, 33, pp. 3-56.

[14] Fama, E. and J. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests", *Journal of Political Economy, 81*, pp. 607-636.

[15] Ferson, W., and C. Harvey, 1991, "The variation of economics risk premiums", *Journal of Political Economy*, 99, pp. 385-415.

[16] Ferson, W., and R. Korajczyk, 1995, "Do arbitrage pricing models explain the predictability of stock returns", *Journal of Business*, 68, pp. 309-349.

[17] Franzoni, F., 2002, "Where is beta going? The Riskiness of Value and Small Stocks", Ph.d. Dissertation, MIT.

[18] Gennotte, G., 1986, "Optimal portfolio choice under incomplete information", *Journal of Finance 61*, pp. 733-749.

[19] Hamilton, J. D., 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.

[20] Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying", *Journal of Political Economy*, 109, pp. 1238-1287.

[21] Lewellen, J. and J. Shanken, 2002, "Learning, asset-pricing tests, and market efficiency", *Journal of Finance 57*, pp. 1113-1145.

[22] Lintner, J., 1965, "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics*, 47, pp. 13.-37.

[23] Mehra, R. and E. Prescott, 1985, "The Equity Premium: A Puzzle", *Journal of Monetary Economics 15*, pp. 145-161.

[24] Merton R. C., 1973 "An intertemporal Asset Pricing Model", *Econometrica 41*, pp. 867-887.

[25] Pastor, L., 2002, "Portfolio selection and asset pricing models", *Journal of Finance 55*, pp. 179–223.

[26] Ross, S., 1976, "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory 13*, pp. 341-360.

[27] Shanken, J., 1990, "Intertemporal asset pricing: An empirical investigation", *Journal of Econometrics*, 45, pp. 99-120.

[28] Sharpe, W., 1964, "Capital asset prices: A theory of market equilibrium under conditions of risk", *Journal of Finance*, 19, pp. 425-442.

[29] Shiller, R. J., 1981 "Do Stock Prices Move too much to be Explained by Subsequent Changes in Dividends?", *American Economic Review 71*, pp. 421-436.

Chapter 3

Do Stock Prices Neglect the Implications of Current Earnings for Future Earnings?

3.1 Introduction

The modern view of an efficient market can be summarized by Fama's [1970] definition that a market is efficient if prices 'fully reflect' all the available information. As few other pieces of information are more informative about a stock's ability to generate payoffs than accounting earnings, it is not surprisig that a great deal of research in the field of market efficiency has been dedicated to the response of prices to earnings news. Since the early study by Ball and Brown [1968], this literature has provided evidence that cumulative abnormal stock returns continue to drift up after good earnings news, and down after bad news, for a period of time that can last up to three quarters. If one believes that prices in efficient markets should react immediately to news, this fact is a challenge to the efficient markets hypothesis, and for this reason it has been labelled the "post earnings announcement drift anomaly".

The early studies of post-earnings announcement drift suffered from a variety of method-ological limitations that could have biased the results (see Ball [1978], Foster, Olsen and Shevlin [1984], and Bernard and Thomas [1989], for a discussion). In particular, the sup-porters of market efficiency argued that these studies did not properly account for the

change in the underlying risk factors, following an earnings announcement. These criticism has been rebutted in Bernard and Thomas's [1989] paper, where they show that the drift is present, even when the change in risk is accounted for. Lately, even strenuous supporters of market efficiency like Fama [1997], admit that the post-earnings announcement drift is the most resistent of the financial anomalies.

Given that good (bad) earnings news appear to predict future positive (negative) abnormal returns over short horizons (up to three quarters), the post-earnings announcement drift has been considered as evidence that prices underreact to news. At the same time, the evidence that good (bad) news firms earn negative (positive) abnormal returns after four quarters, has led researchers to talk about long-run overreaction of prices. A great deal of theoretical work in behavioral finance has been intended to provide explanations of why prices do not incorporate immediately all the available information (see, e.g., Barberis, Shleifer and Vishny [1997]), and to develop settings where the drift is not eliminated by arbitrage (see, e.g., Hong and Stein [1997]).

This paper has been inspired by Bernard and Thomas's [1990] conjecture that market expectations are formed naively. That is, investors expect that earnings in a quarter will be the same as earnings in the same quarter of the prior year, plus a drift. From this hypothesis, I derive the implication that abnormal returns depend only on the seasonal difference of earnings. The paper tests this implication.

Subsequently, I consider the problem from the point of view of efficient markets. If prices fully impounded all the available information, then the abnormal return should not depend on any predictable component of earnings. This implication is also tested in the paper.

The results of the tests produce a picture where stock prices of large firms appear to

behave efficiently. Whereas, for small firms, neither the efficient markets hypothesis, nor Bernard and Thomas's conjecture seem appropriate. This finding suggests the consideration that some intermediate view could explain the behavior of small firms' stocks. The paper provides evidence that supports this conjecture.

The paper is organized as follows. In Section 3.2 a formal model for abnormal returns is described in some detail. This model can encompass different hypothesis on market expectations of earnings. The two tests mentioned above are then developed. Section 3.3 describes the sample used for the analysis. The time-series properties of the earnings process are then described, and variance ratio tests are performed to devide the sample between firms whose earnings process appears to be a random walk and firms for which it does not. The two tests are implemented in Sections 3.4 and 3.5. A possible explanation of the behavior of small firms' stock prices is considered in Section 3.6, and evidence is provided in favor of this hypothesis. Finally, Section 3.7 draws the conclusions of the analysis.

3.2 Theoretical framework

3.2.1 Development of the testable implications

In order to give an explanation of the post-earnings announcement drift anomaly, some researchers have moved away from the assumption of informational efficiency of the stock market.

A seminal paper in this line of research is Bernard and Thomas's 1990 article. On the basis of previous evidence (e.g., Bernard and Thomas [1989]), they provide an interpretation of the post-earnings announcement drift centered on the failure of investor expectations to fully incorporate the implications of current earnings for future earnings. Their hypothesis is motivated by the finding that the 3-day stock price reaction around the time of quarterly earnings reports is predictable, conditional upon knowledge of the past four quarterly

earnings.

Since their explanation of this evidence is the basis of the tests in this paper, it will be examined in some detail.

In Box-Jenkins [1976] ARIMA notation, a discrete stochastic process with a seasonal pattern (over S periods) is represented as

$$\phi(L)\Phi(L)X_t = \theta(L)\Theta(L)u_t + \delta \tag{A-1}$$

where X_t is a stationary process; L is the lag operator; $\phi(L)$ and $\theta(L)$ are polynomials in L, namely $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$; $\Phi(L)$ and $\Theta(L)$, the seasonal parts of the model have the form $\Phi(L^S) = 1 - \Phi_{1S}L - \cdots - \Phi_{PS}L^{PS}$ and $\Theta(L^S) = 1 - \Theta_{1S}L - \cdots - \Theta_{QS}L^{QS}$; u_t is a white noise process; and δ is a constant. The model is of order $(p, d, q) \times (P, D, Q)$, where the capital letters refer to the seasonal part. Autoregressive parameters are denoted by p and P; moving average parameters by q and Q; and d and D are the degrees of ordinary and seasonal differencing required to achieve stationarity.

In the accounting literature there is consensus that a random firm's earnings can be represented by the Brown and Rozeff [1979] ARIMA $(1, 0, 0) \times (0, 1, 1)$ model

$$Q_t = \delta + Q_{t-4} + \phi(Q_{t-1} - Q_{t-5}) + \epsilon_t - \theta\epsilon_{t-4} \tag{A-2}$$

where Q_t are quarterly earnings in quarter t, ϵ_t is a white noise shock, $0 < \phi < 1$, and θ is sufficiently positive to ensure that the fourth order autocorrelation in seasonally differenced earnings is negative.

Bernard and Thomas [1990] consider the Brown-Rozeff model as a good description of their sample of firms, where seasonally differenced earnings tend to exhibit an autocorrela-

tion of about 0.34 at a lag of one quarter, of 0.19 at two quarters, of 0.06 at three quarters, and -0.24 after four quarters. That is, the change of earnings with respect to the same quarter of the previous year, displays momentum at one, two and three quarter horizons, and reversal after a year.

Previous research (e.g. Bernard and Thomas [1989]) had shown that a disproportionately large part of the post-announcement drift was concentrated in the three day window around the next quarterly announcement. In other words, given that a firm announces positive (negative) unexpected earnings for quarter t, the market tends to be positively (negatively) surprised in the days sorrounding the announcement for quarter $t+1$. Bernard and Thomas [1990] investigate the possibility that stock prices fail to reflect fully the implications of current earnings for future earnings. Specifically, they entertain the hypothesis that the market forms expectations in a naive way: i.e. investors believe that earnings follow a seasonal random walk with drift (ARIMA $(0,0,0) \times (0,1,0)$)

$$Q_t = \delta + Q_{t-4} + \epsilon_t \tag{A-3}$$

Accordingly, the market's expectation of next quarter's earnings is simply equal to the earnings of the corresponding quarter of the previous year plus a drift

$$E^M_{t-1}(Q_t) = \delta + Q_{t-4} \tag{A-4}$$

where the superscript M indicates market expectations.

On the other hand, if the true process is the one in Equation A-2, as it is assumed, the conditional expected value of earnings is

$$E^M_{t-1}(Q_t) = \delta + Q_{t-4} + \phi(Q_{t-1} - Q_{t-5}) - \theta\epsilon_{t-4} \tag{A-5}$$

At this point, it is necessary to specify how the abnormal returns are generated. The

135

assumption is that the reduced form of abnormal returns (AR) can be expressed as

$$AR_t = \lambda(Q_t - E^M_{t-1}(Q_t))$$ (A-6)

The abnormal return is the excess return over what the security would earn in the absence of the event (in this case the earnings announcement), which in turn depends on the model that one believes is underlying the stock market (e.g. CAPM, APT, etc.). Equation A-6 assumes that the residual return is entirely generated by company specific information about earnings. This assumption, common in the event study research, is convenient because it relates the abnormal return to the extent of earnings surprise, with no need of fully specifying the generating model. Its cost is that it postulates a linear relation, disregarding possible interactions of earnings with other variables that belong to the investor's information set.

Whenever the market expectations of earnings are unbiased, Equation A-6 implies that the abnormal return associated with the earnings announcement is equal to the innovation in time t earnings

$$AR_t = \lambda \epsilon_t$$ (A-7)

Instead, if expectations are biased, the abnormal return depends on the unpredictable component, $Q_t - E_{t-1}(Q_t)$, and on its complement, $E_{t-1}(Q_t) - E^M_{t-1}(Q_t)$, that is predictable on the basis of information available at time $t - 1$

$$AR_t = \lambda(Q_t - E_{t-1}(Q_t)) + \lambda(E_{t-1}(Q_t) - E^M_{t-1}(Q_t))$$ (A-8)

Relying on the Brown-Rozeff model, and on the naive market expectations of Equation A-4, Bernard and Thomas replace $Q_t - E_{t-1}(Q_t)$ with the difference between Equations A-2 and A-5, and $E_{t-1}(Q_t) - E^M_{t-1}(Q_t)$ with the difference between Equations A-5 and A-4

$$AR_t = \lambda \epsilon_t + \lambda \phi D_{t-1} - \lambda \theta \epsilon_{t-4}$$ (A-9)

where $D_{t-1} = Q_{t-1} - Q_{t-5}$. Since D_{t-1} can be expressed in terms of previous shocks, Equation A-9 can be rewritten as

$$AR_t = \lambda\epsilon_t + \lambda\phi\epsilon_{t-1} + \lambda\phi^2\epsilon_{t-2} + \lambda\phi^3\epsilon_{t-3} + \lambda(\phi^4 - \theta)\epsilon_{t-4} + \lambda v \qquad \text{(A-10)}$$

where v is a linear combination of shocks from periods prior to $t - 4$, and it is therefore uncorrelated with the other shocks in the equation. The coefficients on ϵ_t to ϵ_{t-3} are positive and declining in magnitude, while $(\phi^4 - \theta)$ is negative by the assumption of negative fourth order autocorrelation.

Bernard and Thomas [1990] estimate a Brown-Rozeff model for each company's earnings, and use the fitted residuals as a proxy for the earnings innovations in a regression based on Equation A-10. The main finding of their paper is that the estimates, based on a sample of 2,626 firms over the 1974-1986 period, have the sign and magnitude predicted by Equation A-10. Besides, they show that the knowledge of past earnings allows to construct a portfolio that earns an abnormal return of about 2.6% over the three-day window around the announcement, which can get to 3.35% when the portfolio is restricted to small firms. They take these results as evidence of the fact that the market expectation is formed on the basis of a seasonal random walk model, causing the failure of prices to incorporate the implications of current earnings for future earnings.

In a later paper, Bernard [1992] considers some explanations for why the market participants' expectations would be so heavily anchored to the comparable earnings figure of the prior year. One category of explanations resorts to the psychological literature documenting a tendency for individuals in certain prediction contexts to anchor on some value and place little weight on recent changes in a series, unless these changes are considered salient and are attributed to a stable underlying cause (Andreassen [1987], Andreassen and Kraus [1990]).

Alternatively (or in addition), the reasons for a sluggish updating of expectations can be searched in the incentive structure at work for a key group of players, financial analysts. Bernard argues that there may be little to gain and much to lose by adjusting a forecast to a level far from current consensus, if the analyst believes that his or her forecast is already likely to be the most accurate.

Much theoretical work in behavioral finance has been devoted to develop models that incorporated the findings of the empirical literature concerning the earnings anomaly. In particular, Barberis, Shleifer and Vishny [1997] share Bernard and Thomas's view that "...investors typycally (not always) believe that earnings are more stationary than they really are." In other words, they believe that market expectations fail to recognize the autocorrelation in the seasonal differences of quarterly earnings. On this assumption, they build a model where stock prices can underreact to news, and future returns can be predicted using past information.

In this paper I develop two empirical tests that address the question of whether market expectations of earnings are based on a seasonal random walk or they are based on the correct autoregressive process. Moving from the results of these tests, further analysis is conducted to evaluate the extent of informational efficiency for stocks in different size groups.

3.2.2 The testable implications

When Bernard and Thomas's hypothesis is taken literarily, namely in the way it is expressed in Equation A-4, it implies that what counts in the generation of abnormal returns, is only the difference between current earnings and the earnings in the corresponding quarter of the prior year.

Replacing the market expectation of earnings, as given by Equation A-4, into the model

for abnormal returns, as expressed by Equation A-6, yields

$$AR_t = \lambda(D_t - \delta) \tag{A-11}$$

where $D_t = Q_t - Q_{t-4}$ is the seasonal difference in earnings. Equation A-11 indicates that one implication of Bernard and Thomas's hypothesis is that no other piece of information about earnings, besides D_t, is a significant determinant of the abnormal return. I label the empirical test of this implication Test 1. Note that Equation A-11 has been obtained without any assumption on the true process of earnings. Therefore, Test 1 controls the validity of Bernard and Thomas's assumption in a way that is more direct than Bernard and Thomas's test based on Equation A-10, since the latter involves the joint hypothesis that earnings follow the Brown-Rozeff model.

The second test takes a different perspective [1]. From Equation A-8 it is evident that, if expectations are formed rationally, the abnormal return depends only on the innovation in quarterly earnings, as in Equation A-7, reported below

$$AR_t = \lambda \epsilon_t \tag{A-12}$$

On the other hand, if expectations are formed according to a seasonal random walk, the correct specification is the one in Equation A-9, redisplayed below

$$AR_t = \lambda \epsilon_t + \lambda \phi D_{t-1} - \lambda \theta \epsilon_{t-4} \tag{A-13}$$

Therefore, once that the response of abnormal returns to the innovation in earnings is appropriately controlled for, the failure of D_{t-1} and ϵ_{t-4} to be significant determinants of the abnormal return would depose against Bernard and Thomas's hypothesis that market

---

[1] Under many respects this test resembles the tests for excess sensitivity of consumption to income, present in the consumption literature (see, e.g., Flavin [1981]).

expectations are biased. In other words, if there is underreaction of prices to earnings news, then the abnormal return should display sensitivity to the predictable component of earnings. I label the test of this implication Test 2. Note that the assumption that earnings follow a Brown-Rozeff model constrains the signs of D_{t-1} and ϵ_{t-4} to be positive and negative, respectively. Besides, this assumption rules out the significance of other prior earnings terms as determinants of the abnormal return. More generally, if one assumes the efficiency of expectations as the null hypothesis (Equation A-7), the absence of significance of earnings information, controlling for the earnings surprise, would cause the rejection of any alternative, including Bernard and Thomas's hypothesis, no matter what the process of earnings is. Further, note that if the process of earnings is actually a seasonal random walk, the distinction between Bernard and Thomas's hypothesis and market efficiency would collapse, the two being observationally equivalent. Allowing for this possibility, the sample will be partitioned into firms whose earnings process appears to be a seasonal random walk, and firms for which it does not.

Finally, Test 1 allows one to draw some more inference on the issue of market efficiency. At the other extreme with respect to Bernard and Thomas's hypothesis that the market disregards the autocorrelation of earnings, lies the assumption that the market treats efficiently the available information about earnings. In this case, the determinants of the abnormal return are obtained replacing the true expectation of earnings from Equation A-5 into Equation A-6

$$AR_t = \lambda(D_t - \delta) - \lambda\phi D_{t-1} + \lambda\theta\epsilon_{t-4} \qquad (A-14)$$

Hence, the efficient market hypothesis constrains the coefficient on $(D_t - \delta)$ to be equal in absolute value to the coefficient on ϕD_{t-1}. However, the test involves the joint hypothesis that earnings follow a Brown-Rozeff model, therefore the failure to find such equality would

140

not establish the inefficiency of expectations.

3.3 The sample and the time-series of earnings

3.3.1 Sample selection and variable construction

Earnings data have been extracted from Compustat Industrial and Full Coverage 1999 quarterly files, and include companies traded in the NYSE/AMEX and NASDAQ, over the period from 1988 to 1999. To be included in the sample, a firm must have at least thirtytwo consecutive quarterly reports of earnings before extraordinary items (DATA8 in Compustat file). Information about returns and capitalization has been extracted from the CRSP daily dataset, and merged with the earnings sample. The available CRSP data covers the period up to 1998, therefore the analysis that does not concern returns information, such as the estimation of the time series process of earnings, has been conducted using the longer Compustat sample. The merged sample includes 61,632 firm-quarters of data, for 1,665 companies, from 1988 to 1999. This sample is more recent than the 1974-1986 sample used by Bernard and Thomas [1990].

A variable that will be used in the analysis is the abnormal stock return (AR) associated with the quarterly earnings announcement. After finding that a disproportionate part of the post-earnings announcement drift is concentrated in the period including the two days before and the day of the announcement , the literature has focused on this short time span to evaluate the impact of information on prices (see, e.g., Bernard and Thomas [1990], Brown and Han [forthcoming]) [2]. This approach has the advantage of limiting the importance of the choice of the model for normal returns. Given that daily normal returns are close to

---

[2]The assumption underlying the inclusion of the days before the announcement in the time window is that the news about the quarterly earnings begin to materialize before the announcement itsself. This is confirmed by the fact that stock returns of good news firms display a drift before the news is released (see, e.g., Bernard and Thomas [1989]).

zero, the model for normal returns does not have a big effect on inferences about abnormal returns. Hence, I use the CRSP equally-weighted index as a measure of the normal return. The AR is obtained subtracting the index from the daily returns and summing the residuals over the 3-day window. If the return is not available on the day of the announcement, it is replaced with the first available return in the seven days following the announcement (in the absence of that, the quarter-firm observation is dropped from the sample). Similarly, the returns of the two days before the announcement can be replaced with data from the previous seven days.

3.3.2 The time-series properties of earnings

As mentioned before, the cumulative evidence of the studies concerned with the time-series behavior of earnings (e.g., Foster [1977], Brown and Rozeff [1979], Brown, Griffin, Hagerman, and Zmijewski [1987]) indicates a precise pattern in the autocorrelations of seasonal differences: positive up to lag three, and negative for higher lags.

Although the sample studied here does not temporally overlap with the ones used in the other studies, the time series properties of earnings appear to respect this pattern, as testified by Table 3.1. Panel A presents the firm-specific sample autocorrelations of seasonally differenced earnings, up to lag eight. The sample means of the autocorrelations estimated separately for each firm are 0.23, 0.12, 0.03 for lags 1, 2 and 3 respectively. In line with prior research, the autocorrelation turns negative at lag 4 (mean = -0.25). This pattern is consistent across the 26 industries (two-digit SIC codes) for which the sample includes at least twenty companies.

To make it comparable across firms in the regression analysis, the seasonal difference variable is scaled down by its firm-specific standard deviation. This new variable is often referred to as standardized unexpected earnings (SUE) in the literature on the post-earnings

142

announcement drift. I will continue to name it seasonal difference, to avoid the a priori implication that market expectations are formed on the basis of a seasonal random walk. In Panel B the means of the sample autocorrelations for this standardized variable are reported. They are quite consistent with the ones in Panel A, indicating that the process of scaling does not have any influence on the degree of autocorrelation.

3.3.3 Variance ratio tests

The distinction between the hyptothesis that market expectations are unbiased and the hypothesis that they are formed on the basis of a seasonal random walk, collapses when the actual earnings generating process is a seasonal random walk, as in Equation A-3. Therefore, the two tests developed in Section 3.2.2 would gain power if they could be conducted on a sample of firms, whose earnings process is not a seasonal random walk. To this purpose, the sample is partitioned on the basis of the results of variance ratio (VR) tests, performed on each firm's earnings series separately.

In this case, the VR test is preferred over the other tests for two reasons. First, given the seasonal nature of the non-stationarity, the unit root tests in their standard version could not be applied. Instead, the VR test are applied to the differenced data: it does not matter whether the non stationarity is at lag 1, or whether it is seasonal, because the test checks whether the differenced [3] data follow a white noise process. Second, one particular version of VR test can accomodate for general heteroskedasticity in the innovations of the earnings process. In fact, it is very well possible that the shocks affecting earnings in the different seasons are not identically distributed, even if these shocks are uncorrelated across periods.

---

[3]In this case, the data is seasonally differenced and it is subtracted an estimate of the drift term. The estimate of the drift is obtained as the mean of the seasonal differences over time.

For the construction of the heteroskedasticity-consistent VR test, as well as for the derivation of its asymptotic distribution, I refer to Lo and MacKinlay [1988], or to the book by Campbell, Lo and MacKinlay [1997] [4].

The VR test requires the specification of the number of periods to include in the computation of the cumulative variance. The choice has fallen on four periods, since the focus of the paper is on the failure of agents to recognize the existence of autocorrelations of earnings within the year [5].

A firm is included in the non-random-walk (NRW) group if the VR test at a size of 5% rejects the null hypothesis. Otherwise it belongs to the random-walk (RW) group. The results of the VR tests are reported in Table 3.2. Almost 55% of the firms in the sample belong to the NRW group (Panel A).

Another dimension along which the sample can be partitioned is firm size. The size of a firm is generally considered to be a proxy for information availability. In fact, larger companies have more analyst coverage and, in general, are followed with more attention. Therefore, one would expect mistakes in expectations to be more likely for small stocks. This conjecture will be further investigated in the next sections. One quarter-firm observation is assigned to size deciles on the basis of the average market value of equity over the year of interest. Small, medium and large firms are in size deciles 1 to 4, 5 to 7, 8 to 10, respectively. Around 40% of the observations pertain to small firms, while the remaining 60% is equally split between medium and large firms.

The results of the VR tests by size group are reported in Panel B of Table 3.2. Small quarter-firm observations are more concentrated in the NRW group, whereas large compa-

---

[4]To be precise, the statistic that I use in these tests is the one denoted as Equation (2.4.44), on page 55 of that book.

[5]Other VR tests have been performed, using different numbers of periods, and the results were not affected substantially.

nies are evenly split between the RW and the NRW groups.

Finally, Panel C of Table 3.1 gives the pattern of autocorrelations of seasonal differences of earnings by VR test result groups. As expected, the pattern is more pronounced for NRW firms than the whole sample, and the numbers are almost identical to the ones in Table 1 of Bernard and Thomas [1990], where they have the autocorrelations of earnings with no distinction between RW and NRW firms. For RW firms, the only autocorrelation that is still far from zero is the one at lag 4. This is the consequence of restricting the focus of the VR tests to the autocorrelations up to lag 3.

3.4 Implementation of Test 1

3.4.1 Econometric issues

Test 1 has been developed in Section 3.2.2 on the basis of the null hypothesis that market expectations of earnings are formed assuming a seasonal random walk process. In that case, the seasonal difference of earnings at time t should be the only significant determinant of time t abnormal return, as in Equation A-11. A possible alternative hypothesis is that expectations correctly reflect the underlying process of earnings. If this process is described by a Brown-Rozeff model (Equation A-2), then the specification of the abnormal return is given by Equation A-14. However, more general alternative hypotheses are possible, where the agents understand that earnings are correlated from period to period, but do not react in the correct measure to this correlation. In this case, lagged seasonal differences of earnings are also significant determinants of the abnormal return.

Since a choice has to be made in terms of what seasonal differences to include in the model, the most natural candidate, which is also suggested by the alternative hypothesis of efficient markets, is the seasonal difference at lag 1.

145

The problem of a test based on a regression like

$$AR_{jt} = b_0 + b_1 D_{jt} + b_2 D_{j,t-1} + u_{jt} \tag{A-15}$$

(where j stands for company , and t for the quarter), is that both D_t and D_{t-1} are endogenous variables in the regression. For example, if expectations are rational and the earnings process is a Brown-Rozeff model, then the correct specification for AR is Equation A-14. In that case, error u_t in the above regression contains the shock ϵ_{t-4}, which is correlated with both D_t and D_{t-1}. This suggests that the endogeneity problem has to be taken into account in the implementation of Test 1.

One possibility is trying to understand in what direction the bias affects the coefficient on D_{t-1}, which is the variable of interest. If the correct model for earnings is the Brown-Rozeff, then D_{t-1} is positively correlated with ϵ_{t-4}. Hence, if ϵ_{t-4} is the only omitted variable in the regression, as it is the case when expectations are unbiased (Equation A-14), then the negative coefficient on D_{t-1} would be biased toward zero, in a way that is favorable to the null hypothesis of Test 1. Therefore, to cope with the endogeneity problem one could estimate a regression like Equation A-15, and reject safely the null hypothesis whenever D_{t-1} turns out significant. This is one solution that will be pursued in the analysis.

Alternatively, one could estimate a Brown-Rozeff model for each company's earnings, obtaining the fitted residuals. These residuals can then be included in the regression of interest to control for the omitted variable ϵ_{t-4}. The estimated model becomes

$$AR_{jt} = b_0 + b_1(D_{jt} - \hat{\delta}_j) + b_2 D_{j,t-1} + b_3 e_{j,t-4} + v_{jt} \tag{A-16}$$

where $e_{j,t-4}$ are the fitted residuals from the estimation of the Brown-Rozeff earnings model, and $\hat{\delta}_j$ is the drift term in the same model (and it is included for coherence with Equation A-

14). As in the case of the previous solution, the validity of this approach relies on the assumptions that earnings are described by a Brown-Rozeff model and that the omission of ϵ_{t-4} is the only cause of endogeneity. The results in Panel C of Table 3.1, indicate that the Brown-Rozeff model is not a strong assumptions for NRW companies. A further issue concerns the estimate of the Brown-Rozeff model. In a time-series as short as that used here, the iterative techniques necessary to estimate the Brown-Rozeff model are often unreliable. To deal with this problem, the error e_{t-4} is obtained from the estimation of a Foster [1970] model, which differs from the Brown-Rozeff model for the omission of the moving-average term. This should not create further problems in the implementation of Test 1, in so far as the correlation of the innovation at lag 8 with the other regressors is negligible.

Both of the above described solutions to the endogeneity problem introduce a joint hypotheses problem, which detracts from the appeal of Test 1. A more convenient alternative would be using instrumental variables estimation. The regressors D_t and D_{t-1} in Equation A-15 can be instrumented by other seasonal differences of earnings at lags prior to lag 4. The validity of such instruments is testified by the results in Table 3.1. Their exogeneity relies on the assumption that the omitted variables, included in the error term u_t, do not involve lags higher than 4. The exploratory analysis conducted using this approach did not yield significant estimates for either D_t or D_{t-1}. In the light of the results obtained using the other two approaches, I am inclined to impute this lack of significance entirely to the weakness of the instruments. Therefore the estimates obtained using IV have not been reported. The appeal of this approach, however, suggests that future research should look for better instruments.

3.4.2 Regression results

The regression results based on Equations A-15 and A-16 are reported in Table 3.3. The construction of the variables used in the analysis has been described in the previous sections. Note that, in order to make them homogeneous cross-sectionally, the variables used for the seasonal differences have been standardized by their firm-specific standard error. The estimations have been performed partioning the sample along two dimensions. The distinction by firm size is intended to separate the effect of large firms, characterized by better information, from that of small firms, for which mistakes in expectations are a priori more likely. The separation between random-walk (RW) and non-random-walk (NRW) firms serves to create a control sample (the RW firms), with which to confront the results obtained on the NRW firms, given that the theory imposes different restrictions on the two groups of companies.

The estimates based on the whole sample are in Panel A of Table 3.3. The significant coefficients for D_{t-1} in both Equation A-15 and Equation A-16 indicate that the first response of Test 1 is a rejection of the null hypothesis. The contemporaneous seasonal difference D_t is not the only determinant of the abnormal return. This means that the earnings surprise is not entirely captured by D_t and a correction for D_{t-1} is required. In fact, the sign of the coefficient on D_{t-1} is negative, as it should be if agents incorporate the positive autocorrelation of earnings into their expectations (see, e.g., Equation A-14).

Note that the coefficient on D_{t-1} in Equation A-15 is smaller in absolute value than the corresponding estimate in Equation A-16, suggesting that the omitted variable bias operates in favor of the null hypothesis, as envisaged. Further, the coefficient on e_{t-4} is positive, consistently with Equation A-14, namely with rational expectations and a Brown-Rozeff model for earnings. The fact that in Panel A the coefficient on D_t is significant also for RW

firms, however, is as much in contrast with the efficient market hypothesis as it is with the Bernard and Thomas's hypothesis. When the earnings process is a random walk, rational agents should realize that the parameter ϕ in Equation A-14 is zero, and this would yield an insignificant estimate of β_2. A possible explanation of why β_2 is significant for RW firms is discussed in Section 3.6.

The analysis by firm size helps to define more clearly the picture. The large firms' sample in Panel B confirms the rejection of the null hypothesis, according to both regression models. Moreover, the fact that D_{t-1} is not a significant determinant of the abnormal return of RW firms, indicates that the unbiased expectations hypothesis is a plausible alternative in the case of large firms. Further investigations conducted in Section 3.4.3 confirm this impression.

The results for small firms (Panel C) are in line with the previous rejections of the null hypothesis. However, the significance of D_{t-1} for RW firms excludes the possibility of explaining the behavior of small stocks on the basis of unbiased expectations. This is not surprising since small stocks are characterized by coarser information.

Other regressions have been run in which higher lags (lag 2 and lag 3) of the seasonal difference were included. The estimates (not reported) are significant for small firms, and insignificant for large firms and for RW firms in any size group. These results strengthen the rejection of the null, and sustain the idea that the behavior of large stocks, unlike that of small stocks, can be described by a model of unbiased expectations.

Note, finally, that the R^2 coefficients of the regressions in Table 3.3 are in general low, not exceeding 3%. This suggests that earnings information is far from being the main determinant of the abnormal stock return.

To summarize, the results of Test 1 reject the null hypothesis of earnings expectations formed on the basis of a seasonal random walk. In the case of small firms, however, this

rejection does not imply that expectations are unbiased. For large firms rational expectations cannot be ruled out so far, and further evidence sustaining this hypothesis is provided in the next section.

3.4.3 Large firms and efficiency

The evidence presented until now (Panel B of Table 3.3), suggests that the hypothesis of unbiasedness of expectations in the case of large firms deserves some credit, for two reasons. First, the null hypothesis that expectations are biased in the direction predicted by Bernard and Thomas is rejected. Second, for RW firms the abnormal return depends only on the contemporaneous seasonal difference, that corresponds exactly to the earnings innovation when earnings follow a seasonal random walk.

If expectations correctly incorporate the fact that earnings follow a Brown-Rozeff model, then the reduced form of the abnormal return is given in Equation A-14. In that equation, both D_t and ϕD_{t-1} are multiplied by λ. Hence, this hypothesis constrains the coefficients b_1 and b_2, in the following regression

$$AR_{jt} = b_0 + b_1(D_{jt} - \delta_j) + b_2(\phi D_{j,t-1}) + b_3 e_{j,t-4} + v_{jt} \qquad (A\text{-}17)$$

to be equal in absolute value. Of course, δ and ϕ are not known. So, for estimation purposes, they have to be replaced by their estimates obtained from a Foster [1970] model. As said before, this model is more reliably estimated than the Brown-Rozeff model, and it has been shown to perform similarly.

The regression

$$AR_{jt} = b_0 + b_1(D_{jt} - \hat{\delta}_j) + b_2(\hat{\phi} D_{j,t-1}) + b_3 e_{j,t-4} + w_{jt} \qquad (A\text{-}18)$$

has been estimated for NRW firms in all size groups. The estimates are reported in Panel

A of Table 3.4, and the results of the Wald test of the hypothesis that b_1 and b_2 add to zero are in Panel B.

The estimates of b_1 and b_2 are very close in absolute value in all the samples that have been considered. However, it is only for large firms that the Wald test does not reject the null hypothesis. Hence, the conjecture that earnings expectations are unbiased in the case of large firms, finds a further confirmation in the results of this test. Again, this conclusion is consistent with the consideration that large stock are characterized by better information.

3.5 Implementation of Test 2

3.5.1 Econometric Issues

As discussed in Section 3.2.2, Test 2 can be interpreted as a test of the null hypothesis that expectations are unbiased, jointly with the assumption of a Brown-Rozeff model for earnings. The alternative hypothesis is that expectations fail to understand fully the implications of current earnings for future earnings, so that the abnormal return responds to predictable components of earnings. Included in this alternative is the possibility that expectations are incorrectly based on a seasonal random walk, as conjectured by Bernard and Thomas. In this event, the abnormal return is generated according to Equation A-9.

Bernard and Thomas [1990] proposed the regression

$$AR_{jt} = \beta_0 + \beta_1 D_{j,t-1} + +\beta_2 e_{j,t-4} + v_{jt} \tag{A-19}$$

as a test of their hypothesis, where $e_{j,t-4}$ are the fitted residuals from a firm-specific Foster model. They interpreted the positive coefficient on D_{t-1} and the negative coefficient on e_{t-4} as evidence in favor of their hypothesis (Equation A-9).

I will also adopt this model, but I will restrict my interpretion of any significant estimate to the rejection of the joint null hypotheses of unbiased expectations and Brown-Rozeff

earnings process.

Alternatively, one could recognize that the omission of ϵ_{t-4} in a regression of AR_t on D_{t-1} causes an endogeneity problem, since ϵ_{t-4} and D_{t-1} are positively correlated. Hence, one could apply instrumental variable estimation to the following model

$$AR_{jt} = \beta_0 + \beta_1 D_{j,t-1} + w_{jt} \qquad \text{(A-20)}$$

using D_{t-5} as an instrument for D_{t-1}. This yields a consistent estimate, and there is no need to use the fitted residuals e_{t-4}. Since only four lags separate D_{t-1} from D_{t-5}, the variable D_{t-5} is a strong enough instrument (the first-stage R^2 is around 2.5%), and the IV estimates will be reported.

3.5.2 Regression results

The results from the OLS and IV estimations of models A-19 and A-20, respectively, are reported in Table 3.5. When both models are estimated on the whole sample (Panel A), the response is a rejection of the null hypothesis. In fact, the coefficient on D_{t-1} is significant and positive indicating that prices at time $t-1$ do not fully incorporate the implications of current earnings for future earnings. In other words, it appears that agents fail to understand the degree to which a change in earnings at time t is likely to be followed by a change in earnings of the same sign at time $t+1$. Consequently, the abnormal return, that depends on the extent to which the market is surprised by the earnings announcement, can be predicted using past earnings information.

Although the negative coefficient on e_{t-4} in the OLS estimation is also consistent with the implication of Bernard and Thomas's hypothesis in Equation A-9, the results do not allow to establish more than a general failure of expectations. In fact, as it will be discussed later, it can be that the market anticipates that seasonal differences are autocorrelated,

152

though it does not appreciate the whole extent of this autocorrelation. The partion by NRW and RW firms of the sample, confirms that when autocorrelation in seasonal differences is present (NRW firms), the abnormal return is sensitive to the predictable component of earnings.

The situation of large firms (Panel B) is different, as neither estimation method gives significant estimates of β_1. In addition, the impact of e_{t-4} is barely significant at the 5% level. The results in Panel B are consistent with the response of Test 1, indicating that, in the case of large firms, market expectations of earnings are close to be unbiased.

The evidence for small firms in Panel C suggests that this group of firms is responsible of the rejection of the null hypothesis in the whole sample. Both estimation procedures indicate that the lagged seasonal difference is a significant predictor of the abnormal return. However, this rejection of the null hypothesis does not imply the acceptance of Bernard and Thomas's conjecture. In fact, the possibility that expectations are based on a seasonal random walk, has been ruled out by Test 1, which is a direct test of this hypothesis.

The R^2 in all the regressions in Table 3.5 is by far smaller than the R^2's in Table 3.3, where the estimate for the earnings innovation is included in the regression. This suggests that, although the expectations are biased, still the extent to which past earnings information can help to predict the abnormal return is limited.

In summary, the compound null hypothesis of unbiased expectations formulated on the basis of a Brown-Rozeff model for earnings has been rejected in this sample of stocks. The responsibility of the rejection lies on small firms. Instead, for large firms the unbiasedness of market expectations seems plausible, as it was found by Test 1.

153

3.6 The case of small firms

Under several respects, the evidence presented in the previous sections indicates that neither the unbiased expectations model, nor the naive expectation hypothesis is appropriate for small firms. First, the Bernard and Thomas's view of naive expetations based on a seasonal random walk is rejected by Test 1. Second, the joint hypoteses of unbiased expectations and Brown-Rozeff earnings process has been rejected by Test 2. Third, for small firms the coefficient on the regressor D_{t-1} in Table 3.3 is significant also in the case of random-walk firms.

This combined evidence suggests a possible explanation of the behavior of small firms' prices. I conjecture that investors form their expectations assuming an intermediate position between the two extreme hypotheses mentioned at the beginning of this section. Specifically, according to this new hypothesis the market assumes that the seasonal differences of earnings are positively autocorrelated for all firms in the small class (both RW and NRW), but it understates the level of this autocorrelation. This may be due to the fact that, in the absence of further information, investors assume that earnings follow a process intermediate between the seasonal random walk and the true Brown-Rozeff model, disregarding what the actual process is. Therefore, no matter whether the earnings process is random walk or Brown-Rozeff, the market expectations of earnings for small firms are given by

$$E_{t-1}^M(Q_t) = \delta + Q_{t-4} + \phi_m(Q_{t-1} - Q_{t-5}) + \epsilon_t - \theta\epsilon_{t-4} \qquad \text{(A-21)}$$

where ϕ_m is smaller than the ϕ that characterizes the true process of earnings. From the comparison of Equation A-21 with Equations A-5 and A-4, it is evident that the parameter of autoregression ϕ_m is intermediate between the true autoregression parameter ϕ, and the zero value of autoregression assumed in the naive expectations case, and typical of RW

154

firms.

Replacing Equation A-21 into Equation A-6, one gets the abnormal returns under this new hypothesis

$$AR_t = \lambda(Q_t - Q_{t-4} - \delta - \phi_m D_{t-1} + \theta\epsilon_{t-4}) \tag{A-22}$$

For the purposes of Test 1, this equation can be rewritten as follows

$$AR_t = \lambda(D_t - \delta) - \lambda\phi_m D_{t-1} + \lambda\theta\epsilon_{t-4} \tag{A-23}$$

or, evidentiating the parameter ϕ of the correct process,

$$AR_t = \lambda(D_t - \delta) - \lambda(\frac{\phi_m}{\phi})\phi D_{t-1} + \lambda\theta\epsilon_{t-4} \tag{A-24}$$

Given the assumption $\phi_m < \phi$, it follows that $\phi_m/\phi < 1$. Consequently, the estimation of the model that was given in Equation A-18 (reported here)

$$AR_{jt} = b_0 + b_1(D_{jt} - \hat{\delta}_j) + b_2(\hat{\phi}D_{j,t-1}) + b_3 e_{j,t-4} + w_{jt} \tag{A-25}$$

should yield an estimate of b_1 smaller than the estimate of b_2 in absolute value, since these two coefficients correspond to the coefficients multiplying $(D_t - \delta)$ and ϕD_{t-1}, respectively, in Equation A-24. This prediction is confirmed by the data. In Table 3.4, the point estimate of b_1 for small firms is smaller than the estimate of b_2 in absolute value. Moreover, the test of the null hypothesis that this two coefficients are equal, rejects in the case of small firms.

Further, since Equation A-21 gives the market expectations no matter whether the actual process is a random walk or not, Equation A-23 should hold also for small firms in the RW class. Consequently, the estimate of the parameter premultiplying D_{t-1} should turn out significant also for RW firms. This is exactly what was found in the review of Table 3.3, when it was noted that neither the naive expectations model, nor the unbiased

155

expectations assumption were appropriate for small firms.

In terms of the regressions for Test 2, using the assumption of a Brown-Rozeff model for earnings (Equation A-2), Equation A-23 can be rewritten as

$$AR_t = \lambda\epsilon_t + \lambda(\phi - \phi_m)D_{t-1} - \lambda\theta\epsilon_{t-4} \tag{A-26}$$

where ϵ_t is again the innovation in the earnings process. This equation implies that when the earnings process is not a random walk ($\phi > 0$) the coefficient on D_{t-1} is positive. This implication is confirmed by the evidence in Panel C of Table 3.5, in the case of NRW firms. Alternatively, if the actual process is a seasonal random walk, then the correct expression for the abnormal return is

$$AR_t = \lambda\epsilon_t + \lambda - \phi_m D_{t-1} - \lambda\theta\epsilon_{t-4} \tag{A-27}$$

which implies that the estimate of the coefficient on D_{t-1} should turn out negative, as if prices overreacted to earnings information. Though not significant, this coefficient results negative , for RW firms using both OLS and IV estimation, as testified by Panel C of Table 3.5.

Therefore, several pieces of evidence concerning small firms depose in favor of the hypthesis developed in this section. From what presented in the paper, it appears that prices do underreact to earnings news, but in a form less extreme than the one envisaged by Bernard and Thomas. Possible explanations of why this is the case, can appeal to the scarcity of information characterizing small firms. In fact, investors may have not enough information to decide whether a firms' earnings process follow a random walk or not. Hence, they may decide to take an intermediate position between the earnings process of NRW and RW firms. This causes underreaction for NRW firms and ovverreaction for RW firms. Similar

arguments have been proposed by Barberis, Shleifer and Vishny [1997], with the difference that they base their explanations on psychological factors driving investors' expectations, rather than informational arguments. Whether my explanation can be sustained in a model where informed and uninformed investors coexist (e.g., Grossman and Stiglitz [1980]) is to be established on theoretical grounds. Instead, whether the informational argument is to be preferred to the psychological explanation, is the scope of further empirical work. Though, the fact that this finding characterizes only small firms, deposes in favor of the informational explanation.

3.7 Coclusions

This paper develops and implements two tests concerned with the reaction of stock prices to earnings information.

The first test assumes as null hypothesis Bernard and Thomas' [1990] conjecture that expectations are formed assuming a seasonal random walk process for earnings, while the actual process is autoregressive in seasonal differences. This hypothesis predicts that the market completely neglects the fact that a change in earnings with respect to the previous year, is likely to be followed by a change in earnings of the same sign in the next quarter. If this is the case, the abnormal return is a function of the contemporaneous seasonal difference only. This implication is the scope of the first test.

The results depose against this naive expectations hypothesis. In fact, the evidence indicates that prices incorporate (at least some of) the positive correlation of seasonal differences. Moreover, in the case of large firms, the coefficients appear to satisfy the restrictions imposed by the hypothesis of unbiased expectations. This is not the case for small firms, suggesting that different structures of expectations underlie the two groups of

157

firms.

The second test assumes unbiased expectations as the null hypothesis. The testable prediction is that the abnormal return should depend only on the innovation in the earnings process. The results of the test are articulated by firm size. For large firms, past earnings information does not seem to be a relevant predictor of the abnormal return, reestablishing the conclusion of efficiency. This finding is confirmed by instrumental variable estimation that corrects for a possible attenuation bias. In the case of small firms, instead, the hypothesis of unbiased expectations is rejected by the fact that abnormal returns can be predicted using past earnings information.

The conbined evidence of the first and the second test has suggested the hypothesis that investors form their expectations about small firms' earnings, by taking an intermediate view between the naive seasonal random walk model and the correct autoregressive model. This behavior may cause underreaction for firms whose earnings process is positively autocorrelated in seasonal differences, and ovverreaction for firms, whose earnings process is actually a random walk. Some of the evidence in the paper supports this conjecture. Whether findings depend on the scarcity of information surrounding small stocks, or on some psychological bias, is an open question on both theoretical and empirical grounds. However, the fact that the behavior of small firms is not mirrored in the large firms' sample, deposes in favor of the first explanation.

Table 3.1: Time-series behavior of quarterly earnings

	Panel A: Autocorrelations in seasonally differenced earnings							
Lag	1	2	3	4	5	6	7	8
	Firm-specific autocorrelations in seasonally differenced earnings							
Mean	0.23	0.12	0.03	-0.25	-0.03	-0.03	-0.02	-0.02
	Number of positive mean autocorrelations for 26 industries[a]							
	26	25	25	0	3	1	4	2
Panel B: Autocorrelations in in standardized seasonally differenced earnings								
Lag	1	2	3	4	5	6	7	8
Mean	0.23	0.12	0.03	-0.25	-0.03	-0.03	-0.02	-0.02
Panel C: Autocorrelations in seasonally differenced earnings by VR outcome[b]								
Lag	1	2	3	4	5	6	7	8
NRW	0.36	0.20	0.06	-0.22	-0.07	-0.06	-0.05	-0.04
RW	0.08	0.03	0.00	-0.29	0.00	0.00	0.00	0.00

[a] Only the 26 industries with at least 20 members are included. Industries are defined on the basis of the 2-digit SIC.

[b] Firms are grouped in random-walk (RW) companies and non-random-walk (NRW) companies on the basis of the results of variance ratio (VR) tests.

Table 3.2: Distribution of firms according to the outcome of variance ratio tests

Panel A: Percentage of companies in each group[a]			
	NRW		RW
	54.95%		45.05%
Panel B: Distribution of VR test outcome by size group[b]			
VR test outcome	Small	Medium	Large	Total
RW	10321	7944	9323	27588
	41.81%	43.01%	50.46%	44.76%
NRW	14360	10524	9154	34044
	58.19%	56.89%	49.54%	55.24%
Total	24687	18408	18477	61632
	100.00%	100.00%	100.00%	100.00%

[a] Firms are grouped in random-walk (RW) companies and non-random-walk (NRW) companies on the basis of the results of variance ratio (VR) tests. [b] Small, medium and large firms are in size deciles 1 to 4, 5 to 7, and 8 to 10 respectively, based on the average value of equity in each year.

(A-15) $AR_{jt} = b_0 + b_1 D_{jt} + b_2 D_{j,t-1} + u_{jt}$

(A-16) $AR_{jt} = b_0 + b_1 (D_{jt} - \hat{\delta}_j) + b_2 D_{j,t-1} + b_3 e_{j,t-4} + v_{jt}$

	b_1	b_2	b_3	R^2	# obs.
Panel A: all size groups					
All firms					
Equation A-15	0.010	-0.002		2.05	41260
	(29.188)	(-7.501)			
Equation A-16	0.010	-0.003	0.001	2.10	36296
	(27.076)	(-8.297)	(3.446)		
NRW firms					
Equation A-15	0.011	-0.003		2.61	26795
	(26.385)	(-7.197)			
Equation A-16	0.011	-0.003	0.001	2.64	23579
	(24.145)	(-7.534)	(2.437)		
RW firms					
Equation A-15	0.008	-0.002		1.26	14465
	(13.586)	(-3.681)			
Equation A-16	0.008	-0.002	0.001	1.34	12713
	(12.910)	(-4.291)	(2.540)		
Panel B: large firms					
All large firms					
Equation A-15	0.004	-0.001		1.00	12317
	(11.219)	(-4.510)			
Equation A-16	0.004	-0.001	0.000	0.83	10907
	(9.467)	(-3.857)	(0.965)		
NRW large firms					
Equation A-15	0.005	-0.002		1.12	7675
	(9.320)	(-3.765)			
Equation A-16	0.004	-0.002	0.000	1.01	6807
	(8.004)	(-3.695)	(0.346)		
RW large firms					
Equation A-15	0.003	-0.001		0.85	4642
	(6.200)	(-2.751)			
Equation A-16	0.003	-0.000	0.001	0.68	4100
	(5.270)	(-1.617)	(1.481)		
Panel C: small firms					
All small firms					
Equation A-15	0.015	-0.002		3.09	16547
	(22.556)	(-3.377)			
Equation A-16	0.016	-0.003	0.002	3.31	14493
	(21.432)	(-4.618)	(3.352)		
NRW small firms					
Equation A-15	0.017	-0.003		4.07	10957
	(20.958)	(-3.820)			
Equation A-16	0.018	-0.004	0.002	4.22	9593
	(19.442)	(-4.384)	(2.777)		
RW small firms					
Equation A-15	0.012	-0.001		1.75	5590
	(9.959)	(-1.188)			
Equation A-16	0.013	-0.002	0.002	2.03	4900
	(9.842)	(-2.225)	(1.846)		

T-statistics in parentheses. R^2 is expressed in percentage points. Firms are grouped in random-walk (RW) companies and non-random-walk (NRW) companies on the basis of the results of variance ratio (VR) tests. Small, medium and large firms are in size deciles 1 to 4, 5 to 7, and 8 to 10 respectively, based on the average value of equity in each year.

Table 3.4: Test of restrictions imposed by unbiased expectations (NRW firms)

(A-18) $AR_{jt} = b_0 + b_1(D_{jt} - \hat{\delta}_j) + b_2(\hat{\phi}D_{j,t-1}) + b_3e_{j,t-4} + w_{jt}$				
Panel A: estimation results				
	b_1	b_2	R^2	# obs.
All NRW firms	0.012	-0.009	2.79	23579
	(24.956)	(-9.906)		
Large NRW firms	0.005	-0.005	1.09	6807
	(8.293)	(-4.333)		
Small NRW firms	0.019	-0.011	4.33	9593
	(19.614)	(-5.584)		
Panel B: Wald Test of H_0: $b_1 + b_2 = 0$				
	All NRW firms	Large NRW firms	Small NRW firms	
P-value	0.001	0.854	0.000	

T-statistics in parentheses. R^2 is expressed in percentage points. Firms are grouped in random-walk (RW) companies and non-random-walk (NRW) companies on the basis of the results of variance ratio (VR) tests. Small, medium and large firms are in size deciles 1 to 4, 5 to 7, and 8 to 10 respectively, based on the average value of equity in each year.

Table 3.5: Test 2

(A-19) $AR_{jt} = \beta_0 + \beta_1 D_{j,t-1} + +\beta_2 e_{j,t-4} + v_{jt}$

(A-20) $AR_{jt} = \beta_0 + \beta_1 D_{j,t-1} + w_{jt}$, where D_{t-5} is IV for D_{t-1}

	β_1	β_2	R^2	# obs.
Panel A: all size groups				
All firms				
Equation A-19	0.001	-0.002	0.12	36292
	(3.402)	(-5.964)		
Equation A-20	0.004			36292
	(2.117)			
NRW firms				
Equation A-19	0.002	-.002	0.22	23579
	(5.003)	(-6.097)		
Equation A-20	0.010			23579
	(3.284)			
RW firms				
Equation A-19	-0.000	-0.001	0.04	12713
	(-0.838)	(-2.162)		
Equation A-20	-0.002			12713
	(-0.986)			
Panel B: large firms				
All large firms				
Equation A-19	0.000	-0.001	0.04	10907
	(0.279)	(-2.127)		
Equation A-20	-0.004			10907
	(-0.376)			
NRW large firms				
Equation A-19	0.000	-.001	0.08	6807
	(0.475)	(-2.370)		
Equation A-20	-0.025			6807
	(-0.544)			
RW large firms				
Equation A-19	-0.000	-0.000	0.00	4100
	(-0.121)	(-0.263)		
Equation A-20	-0.007			4100
	(-1.715)			
Panel C: small firms				
All small firms				
Equation A-19	0.002	-0.003	0.24	14493
	(4.020)	(-4.759)		
Equation A-20	-0.002			14493
	(-1.196)			
NRW small firms				
Equation A-19	0.004	-.004	0.44	9593
	(5.235)	(-4.616)		
Equation A-20	0.005			9593
	(1.853)			
RW small firms				
Equation A-19	-0.000	-0.003	0.09	4900
	(-0.038)	(-2.154)		
Equation A-20	-0.001			4900
	(-0.451)			

T-statistics in parentheses. R^2 is expressed in percentage points. Firms are grouped in random-walk (RW) companies and non-random-walk (NRW) companies on the basis of the results of variance ratio (VR) tests. Small, medium and large firms are in size deciles 1 to 4, 5 to 7, and 8 to 10 respectively, based on the average value of equity in each year.

Bibliography

[1] Andreassen, P., "On the Social Psychology of the Stock Market: Aggregate Attributional Effects and the Regressiveness of Prediction", *Journal of Personality and Social Psychology*, 53, No. 3, 1987, pp. 490-496.

[2] Andreassen, P., and S. J. Krauss, "Judgmental Extrapolation and the Salience of Change", *Journal of Forecasting*, 9, 1990, pp. 347-372.

[3] Ball, R., "Anomalies in Relationships between Securities' Yields and Yield-Surrogates", *Journal of Financial Economics*, June-September 1978, pp. 103-126.

[4] Ball, R., and P. Brown, "An empirical evaluation of accounting income numbers", *Journal of Accounting Research*, 1968, pp. 159-178.

[5] Barberis, N., A. Shleifer, and R. W. Vishny, "A Model of Investor Sentiment", NBER Working Paper No. 5926, February 1997.

[6] Bernard, V., "Stock Price REactions to Earnings Announcements", in R. Thaler ed., *Advances in Behavioral Finance*, 1992, Russel Sage Foundation, New York.

[7] Bernard, V., and J. Thomas, "Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium?" *Journal of Accounting Research*, Supplement 1989, pp. 1-36.

[8] Bernard, V., and J. Thomas, "Evidence that Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings", *Journal of Accounting and Economics*, December 1989, pp. 305-341.

[9] Box, G. E. P., and G. M. Jenkins, *Time-Series Analysis: Forecasting and Control*, 2nd edition, 1976, Holden-Day, San Francisco, CA.

[10] Brown, L. D., P. Griffin, R. Hagerman, and M. Zmijewski, "An evaluation of alternative proxies for the market's assessment of unexpected earnings", *Journal of Accounting and Economics*, 1987 , vol. 9, pp. 159-194.

[11] Brown, L. D., and J. C. Y. Han, "Do stock prices fully reflect the implications of current earnings for future earnings for AR1 firms?" *Journal of Accounting Research*, forthcoming.

[12] Brown, L. D., and M. S. Rozeff, "Univariate time-series models of quarterly accounting earnings per share: A proposed model." *Journal of Accounting Research*, 1979 , vol. 17, pp. 179-189.

[13] Campbell, J. Y., Lo, A. W., and MacKinlay, A. C., *The Econometrics of Financial Markets*, 1997 , Princeton University Press, Princeton, NJ.

[14] Fama, E., "Efficient capital markets: A review of theory and empirical work", *Journal of Finance*, 1970 , vol. 25, pp. 383-417.

[15] Fama, E., "Market Efficiency, Long-Run Returns, and Behavioral Finance", 1997 , Working Paper.

[16] Flavin, M., "The adjustment of consumption ot changing expectations about future income", *Journal of Political Economy*, 1981 , vol. 89, pp. 974-1009.

[17] Foster, G., "Quarterly Accounting Data: Time Series Properties and Predictive-Ability Results", *The Accounting Review*, January 1977 , pp. 1-21.

[18] Foster, G., C. Olsen, and T. Shevlin, "Earnings Releases, Anomalies, and the Behavior of Security Returns" *The Accounting Review*, October 1984 , pp. 574-603.

[19] Grossman, S. J., and J. E. Stiglitz, "On the impossibility of informationally efficient markets", *American Economic Review*, June 1980, vol. 70(3) , pp. 393-408.

[20] Hong, H., and J. C. Stein, "A unified theory of underreaction, momentum trading and overreaction in asset markets", NBER Working Paper No. 6324, December 1997.

[21] Lo, A. W., and MacKinlay, A. C., "Stock prices do not follow a random walk: evidence from a simple specification test", *Review of Financial Studies*, 1988 , vol. 1, pp. 41-66.