Chapter 1:The Econometrics of Financial Returns

Contents

1 Introduction **2**

1.1 The Data . . . 2

1.2 The dimensions of the data . . . 4

2 The Challenges of Financial Econometrics **4**

3 Prof Wald and the missing bullet holes: identification matters **5**

4 The Traditional Model **6**

4.1 The view from the 1960s: Efficient Markets and CER . . . 6

4.1.1 Time-Series Implications . . . 7

4.1.2 Returns at different horizons . . . 7

4.1.3 The Cross-Section of Returns . . . 7

4.1.4 The Volatility of Returns . . . 8

4.1.5 Implications for Asset Allocation . . . 8

5 Empirical Challenges to the traditional model **8**

5.1 The time-series evidence on expected returns . . . 9

5.2 Anomalies . . . 13

5.3 The Cross-section Evidence on Expected Returns . . . 13

5.4 The behaviour of returns at high-frequency: non-normality and heteroscedasticity . . . 15

6 The Implications of the new evidence **15**

6.1 Asset Pricing with Predictable Returns . . . 15

7 Quantitative Risk Management and the behaviour of returns at high-frequency **17**

8 The Plan of the book. Predictive Models in Finance **18**

1. **Introduction**

Predicting the distribution of returns of financial assets is a task of primary importance for identifying desirable investments, performing optimal asset allocation within a portfolio, as well as measuring and managing portfolio risk. Optimal asset management depends on the statistical properties of returns at different frequencies. Portfolio allocation, i.e., the choice of optimal weights to be attributed to the different (financial) assets in a portfolio, is typically based on a long horizon perspective, while the measurement of risk of a given portfolio takes typically a rather short-horizon perspective. This means that a long-run investor decides her optimal portfolio allocation on the basis of the (joint) distribution of the returns of the relevant (i.e., from some pertinent asset menu from which to chose) financial assets at low frequency.[1] However, the monitoring of the daily risk of a portfolio normally depends on the statistical properties of the distribution of returns at high frequencies.

This book (project), in its characteristically applied nature, is designed to illustrate the statistical techniques to perform the analysis of time series of asset (often, financial) returns at different frequencies and its application to asset management and performance evaluation, portfolio allocation, and financial risk management.

The relevant concepts will be introduced and their application will be discussed by using a set of programmes written using mainstream econometric software (EVIEWS and MATLAB) specifically designed for each of the chapters. Draft codes for the solutions of the exercises, that are designed to allow the reader to understand how the different econometric techniques could be put at work, are made available in advance on the book webpage. Students are expected to work through them in specifically devoted computer lab sessions.

1.1. *The Data*

All empirical applications will be based on publicly available databases of US data observed at monthly (and therefore lower) frequency. They have been downloaded respectively from Robert Shiller's webpage

(http://www.econ.yale.edu/~shiller/)

and Ken French's webpage

(http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

The time series made available by Robert Shiller are saved in the successive columns of the EXCELworksheet DATA in the file **IE_DATA.XLS**

[1] There is emprical evidence that females outperform males as professional portfolio managers. One wonders whether this may be a reflection of typical decision horizons that may possibly differ across these two categories of investors.

The time-series in the IE_DATA.XLS files	
identifier	description
P	S&P composite index
D	S&P dividend (at annual rate)
E	S&P earnings
CPI	US consumer price index
GS10	YTM of 10-year US Treasuries
CAPE	cyclically adjusted PE ratio

As described in the section "Online Data" of the webpage these stock market data are those used in the book, Irrational Exuberance [Princeton University Press 2000, Broadway Books 2001, 2nd ed., 2005] and cover the period 1871-Present . This data set consists of monthly stock price, dividends, and earnings data and the consumer price index (to allow conversion to real values), all starting January 1871. The price, dividend, and earnings series are from the same sources as described in Chapter 26 of the book Market Volatility [Cambridge, MA: MIT Press, 1989], although they are observed at monthly, rather than annual frequencies. Monthly dividend and earnings data are computed from the S&P four-quarter totals for the quarter since 1926, with linear interpolation to monthly figures. Dividend and earnings data before 1926 are from Cowles and associates (Common Stock Indexes, 2nd ed. [Bloomington, Ind.: Principia Press, 1939]), interpolated from annual data. The CPI-U (Consumer Price Index-All Urban Consumers) published by the U.S. Bureau of Labor Statistics begins in 1913; for years before 1913 1 spliced to the CPI Warren and Pearson's price index, by multiplying it by the ratio of the indexes in January 1913. December 1999 and January 2000 values for the CPI-Uare extrapolated. See George F. Warren and Frank A. Pearson, Gold and Prices (New York: John Wiley and Sons, 1935). Data are from their Table 1, pp. 11–14.

The time series made available by Ken French are saved in the successive columns of the EXCELworksheet DATA in the file **FF_DATA.XLS.**

The time-series in the FF_Data_CH3.xls files	
identifier	description
EXRET_MKT	MKT excess ret
SMB	returns on SMB
HML	returns on HML
RF	returns on the risk-free asset
MOM	returns on MOM
RMW	returns on RMW
CMA	returns on CMA
PR(i,j)	returns on 25 FF portolios (i=1,...5,j=1,...,5)

The construction of the Fama French factors is described at http://mba.tuck.dartmouth.edu/pages/facu f_5_factors_2x3.html while the construction of the FF portfolios is described at http://mba.tuck.dartmouth.

1.2. *The dimensions of the data*

There are three relevant dimensions of the data on financial returns: time-series, cross-section and the horizon at which returns are defined. In general, we shall define $r^i_{t,t+k}$ as the returns realized by holding between time t and time $t+k$, the asset i. So the t index captures the time-series dimension, the i index the cross-sectional dimension, and the k index the horizon dimension.

2. **The Challenges of Financial Econometrics**

In general, financial data are not generated by experiments, what is available to the econometrician are observational data, which are given. To investigate the effect of a medicine an investigator can take a set of patients and attribute them randomly to a "treatment" group and a "control" group. The medicine is then administered to the members of the treatment group while a "placebo" is given to the members of the control group. The effect of the medicine can then be measured by the difference in the average health of the members of the two groups after the administration of the treatment.

If a researcher is interested in measuring the effect of monetary policy on stock market returns all she has are data on monetary policy indicators and the stock market returns which are given and not generated by a controlled experiments.

Special issues arise in routinely in financial data that are different in special days (say, for example, the days of the FOMC meetings), that are affected by seasonality, trends and cycles. Moreover rare-events affect financial returns and rare events are, by definition, not regularly observed. As Nassim Taleb forcefully stresses in his book Antifragile, absence of evidence in a given sample of data cannot be taken as evidence of absence.

Econometricians face questions of different nature: sometimes the interest lies in non-causal predictive modeling which can be handled by analyzing conditional expectations, while this is not sufficient to understand causation to which end correlation and conditional expectations are little informative. One issue is to evaluate if the monetary policy stance helps to predict stock market returns, which is very different from establishing a causation from monetary policy to the stock market, as the evidence of correlation between monetary policy and the stock market might very well reflect the response of monetary policy to stock market fluctuations taken as an indicator of (present and future) economic activity.

3. **Prof Wald and the missing bullet holes: identification matters**

Econometrics is about using the data. This is not as easy as it looks. There is an issue of fundamental importance that needs to be addressed when using the data, econometricians call it identification. To understand what this is about consider a nice story described by J.Ellenberg(2015) in his excellent book "How Not to be Wrong. The Hidden Maths in Everyday Life". The stroy is about Abraham Wald a famous statistician who was invited to join the Statistical Research Group (SRG). SRG was a group of statisticians employed strategically by the US Army in WWII to apply statistics to military issues. The SRG was faced with the problem of the optimally design of armoring military planes. The problem is interesting because it affected by a tradeoff: to prevent prevent planes from being shot down by enemy fighters your armor them, but armour makes the plane heavier and therefore they are handicapped in dogfights. So the question on the optimal level of amouring naturally arises. Data might be helpful to answer this question. When American planes came back from engagements over Europe, they were covered with bullet holes. Here are the data

Section of the Plane	Bullet Holes per sq. f.
Engine	1.11
Fuselage	1.73
Fuel System	1.55
Rest of the plane	1.8

Econometrics is about using the data to make decisions. So, in the light of the data, if you want to limit the armouring to the most relevant section of the plan to keep it light and effective where do you put the armoring ?

Before you answer let me tell you what was A. Wald choice. The armor, said Wald, does not go where the bullet holes are. It goes were the bullet holes are not:the engines.

To use the data it is important to identify how they are generated. These data are not taken unconditionally they are taken conditionally on one event: planes used for the observation came back. The sample is selected. So it informs us on the fact that planes that are hit on the engine are less likely to come back. That is why it is the engine that should be armoured.

4. **The Traditional Model**

The plan of our journey is determined by the evolution of the understanding and empirical modelling of asset prices and financial returns from the 1960s onwards. We shall start from the view from the sixties, based on the Constant Constant Expected Returns (CER) model and the CAPM, when a simple econometric model serves the purpose of modelling returns at all horizons and a one-factor model determines the cross-section, to illustrate its empirical failures and how it has been replaced by a Time-Varying Expected Returns (TVER) model where different econometric models for returns are to be adopted according to the different horizon at which returns are defined .

4.1. *The view from the 1960s: Efficient Markets and CER*

The history of empirical finance starts with the "efficient market hypothesis" (see Fama, 1970). This view, that dominated the field in the 1960s and 1970s, can be summarized as follows (see also the discussion in Cochrane, 1999):

- expected returns are constant and normally independently distributed;
- the CAPM is a good measure of risk and thus a good explanation of why some stocks earn higher average returns than others;
- excess returns are close to be unpredictable: any predictability is a statistical artifact or cannot be exploited after transaction costs are taken into account;
- the volatility of returns is constant.

Fama (1970) clearly stated:

> "... For data on common stocks, tests of 'fair game' (and random walk) properties seem to go well when conditional expected returns is estimated as the average return for the sample of data at hand. Apparently the variation in common stock returns about their expected values is so large relative to any changes in expected values that the latter can be safely ignored..."

4.1.1. **Time-Series Implications**

In practice, the traditional view can be recasted in terms of the simplest possible specification for the predictive models for returns, i.e., the constant expected returns model:

$$r^i_{t,t+1} = \mu^i + \sigma^i \epsilon_{it} \qquad \epsilon_{it} \sim NID\,(0,1)$$
$$Cov\,(\epsilon_{it}, \epsilon_{js}) = \begin{cases} \sigma_{ij} & t = s \\ 0 & t \neq s \end{cases}.$$

Note that the absence of predictability of excess returns is not a consequence of market efficiency per se but it instead results from a joint hypothesis: market efficiency plus some assumptions on the process generating returns (i.e., the Contant Expected Returns model).

4.1.2. **Returns at different horizons**

In this world, the horizon k does not matter for the prediction of returns because once μ_i and σ_i are estimated, expected retrurns at all horizons and the variance of returns at all horizon are derived deterministically.

$$E(r^i_{t,t+n}) = E(\sum_{k=1}^{n} r^i_{t+k,t+k-1}) = \sum_{k=1}^{n} E(r^i_{t+k,t+k-1}) = n\mu$$
$$Var(r^i_{t,t+n}) = Var(\sum_{i=1}^{n} r^i_{t+k,t+k-1}) = \sum_{i=1}^{n} Var(r^i_{t+k,t+k-1}) = n\sigma^2$$

4.1.3. **The Cross-Section of Returns**

The CER view allows for cross-sectional heterogeneity of returns but such cross-sectional heterogeneity is related to a single factor, the market factor, and the CAPM determines all the cross-sectional variation in μ^i. The statistical model that determines all returns r^i_t and the market return r^m_t,can be described as follows:

$$\begin{aligned}
\left(r^i_t - r^{rf}_t\right) &= \mu_i + \beta_i u_{m,t} + u_{i,t} \\
\left(r^m_t - r^{rf}_t\right) &= \mu_m + u_{m,t} \\
\begin{pmatrix} u_{i,t} \\ u_{m,t} \end{pmatrix} &\sim n.i.d. \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{ii} & \sigma_{im} \\ \sigma_{im} & \sigma_{mm} \end{pmatrix} \right]
\end{aligned}$$

where r^{rf}_t is the return onthe risk-free asset. We shall see that $\sigma_{im} = 0$ is a crucial assumption for the valid estimation of the CAPM betas, and that assumption that risk adjusted excess returns are zero (usually known as zero alpha assumption) requires that $\mu_i = \beta_i \mu_m$.

4.1.4. **The Volatility of Returns**

The volatility of returns is constant in the CER model which therefore is not capable of explaining time-varying volatility in the markets and the presence of alternating period of high and low volatility.

4.1.5. **Implications for Asset Allocation**

When the data are generated by CER optimal asset allocation can be derived by achieved by utility maximization that uses as inputs the historical moments of the distribution of returns, optimal portfolio weights are constant through the investment horizon. The optimal portfolio is always a combination between the market portfolio and the risk free asset. The risk associated to any given asset or portfolio of assets is constant over time. Think of measuring the risk of a portfolio with its Value-at-Risk (VaR). The VaR is the percentage loss obtained with a probability at most of α percent:

$$\Pr\left(R^p < -VaR_\alpha\right) = \alpha.$$

where R^p are the returns on the portfolio. If the distribution of returns is normal, then α-percent VaR_α is obtained as follows (assume $\alpha \in (0,1)$):

$$\begin{aligned} \Pr\left(R^p < -VaR_\alpha\right) &= \alpha \Longleftrightarrow \Pr\left(\frac{R^p - \mu_p}{\sigma_p} < -\frac{VaR_\alpha + \mu_p}{\sigma_p}\right) = \alpha \\ &\Longleftrightarrow \Phi\left(-\frac{VaR_\alpha + \mu_p}{\sigma_p}\right) = \alpha, \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative density of a standard normal. At this point, defining $\Phi^{-1}(\cdot)$ as the inverse CDF function of a standard normal, we have that

$$-\frac{VaR_\alpha + \mu_p}{\sigma_p} = \Phi^{-1}(\alpha) \Longleftrightarrow VaR_\alpha = -\mu_p - \sigma_p \Phi^{-1}(\alpha).$$

and, given that μ_p and σ_p are constant over time, VaR_α is also constant over-time. Consider the case of a researcher interested in the one per cent value at risk. Because $\Phi^{-1}(0.01) = -2.33$ under the normal distribution, we can easily obtain VaR if we have available estimates of the first and second moments of the distribution of *portfolio returns*:

$$\widehat{VaR}_{0.01} = -\hat{\mu}_p - 2.33\hat{\sigma}_p$$

5. **Empirical Challenges to the traditional model**

Over the course of time the traditional view has been empirically challenged on many grounds. In particular it has been observed that

- The tenet that expected returns are constant is not compatible with the observed volatility of stock prices. Stock prices in fact are "too volatile" to be determined only by expected dividends;

- there is evidence of returns predictability that increases with the horizon at which returns are defined.

- There are anomalies that make returns predictable on occasion of special events.

- The CAPM is rejected when looking at the cross-section of returns and multi factor models are needed to explain the cross-sectional variability of returns

- high frequency returns are non-normal and heteroscedastic, therefore risk is not constant over time.

5.1. *The time-series evidence on expected returns*

Practictioners implementing portfolio allocation based on the CER model experienced rather soon a number of problems that stressed limitations of this model but it was the work of Robert Shiller and co-authors that led te profession to go beyond the CER model. The basic empirical evidence against the CER model was the excessive volatility of asset prices and returns which is clearly illustrated in Shiller(1981).

We shall illustrate the excess volatility evidence by considering a simple model of stock market returns: the Dynamic Dividend Growth (DDG) model. As we shall discuss in detail in one of the next chapters total returns to a stock i can be satisfactorily approximated as follows:

$$r_{t+1}^{s} = \kappa + \rho\left(p_{t+1} - d_{t+1}\right) + \Delta d_{t+1} - \left(p_t - d_t\right)$$

where P_t is the stock price at time t and D_t is the dividend paid at time t, $p_t = \ln(P_t), d_t = \ln(D_t), \kappa$ is a constant and $\rho = \frac{P/D}{1+P/D}$, P/D is the average price to dividend ratio. In practice ρ can be interpreted as a discount parameter$(0 < \rho < 1)$. By forward recursive substitution one obtains:

$$(p_t - d_t) = \frac{\kappa}{1-\rho} + \sum_{j=1}^{m} \rho^{j-1}\left(\Delta d_{t+j}\right) - \sum_{j=1}^{m} \rho^{j-1}\left(r_{t+j}^{s}\right) + \rho^{m}\left(p_{t+m+1} - d_{t+m+1}\right)$$

which shows that the $(p_t - d_t)$ measures the value of a very long-term investment strategy (buy and hold). This value, in absence of bubbles, is equal to the stream of future dividend

COINTEGRATION 1083

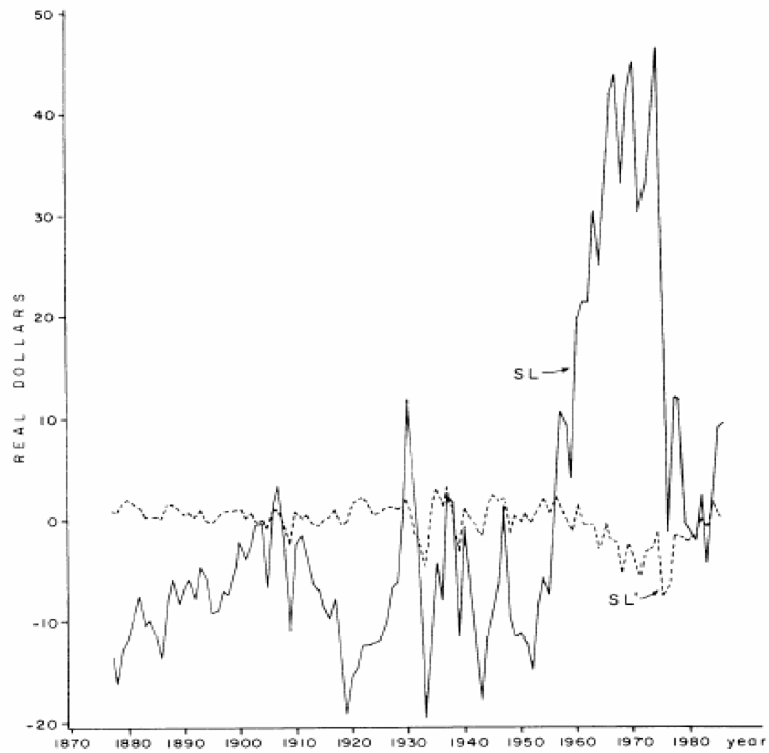


FIG. 2.—Stock market: deviations from means of actual spread (SL_t = Price$_t$ − θ · Dividend$_{t-1}$) and theoretical spread SL_t^*, θ = 12.195.

Figure 1:

growth discounted at the appropriate rate, which reflects the risk free rate plus risk premium required to hold risky assets.

By introducing uncertainty we have:

$$(p_t - d_t) = \frac{\kappa}{1-\rho} + \sum_{j=1}^{m} \rho^{j-1} E_t \left(\Delta d_{t+j}\right) - \sum_{j=1}^{m} E_t \rho^{j-1} \left(r_{t+j}^{s}\right) + \rho^m E_t \left(p_{t+m+1} - d_{t+m+1}\right)$$

Two considerations are relevant here. First, note that under the CER and no bubbles the price dividend ratio should reflect only expected dividend growth. The empirical evidence is strongly against this prediction (see the Shiller(1981) and Campbell-Shiller(1987)). Stock prices are too volatile to be determined only by expected dividends. The following figure, taken from Campbell-Shiller(1987) illustrates the point by reporting the observed price-dividend ratio and a counterfactual price-dividend ratio which is obtained by assuming constant future expected returns and by using a Vector Autoregressive Model to predict future dividend-growth:

The volatility in the price-dividend ratio is clear much higher than that predicted by the

CER model.

Second, once the hypothesis of CER is rejected, the DDG model has interesting implications for predictability of returns at different horizons. If we decompose future variables into their expected component and the unexpected one (an error term) we can write the relationship between the dividend-yield and the returns one-period ahead and over the long-horizon as follows:

$$
\begin{aligned}
r^s_{t+1} &= \kappa + \rho E_t \left(p_{t+1} - d_{t+1}\right) + E_t \Delta d_{t+1} - \left(p_t - d_t\right) + \rho u^{pd}_{t+1} + u^{\Delta d}_{t+1} \\
\sum_{j=1}^{m} \rho^{j-1} r^s_{t+j} &= \frac{\kappa}{1-\rho} + \sum_{j=1}^{m} \rho^{j-1} E_t \left(\Delta d_{t+j}\right) - \left(p_t - d_t\right) + \rho^m E_t \left(p_{t+m} - d_{t+m}\right) + \\
& \quad \rho^m u^{pd}_{t+m} + \sum_{j=1}^{m} \rho^{j-1} u^{\Delta d}_{t+j}
\end{aligned}
$$

These two expressions illustrate that when the price dividends ratio is a noisy process, such noise dominates the variance of one-period returnsand the statistical relation between the price dividend ratio and one period returns is weak. However, as the horizon over which returns are defined gets longer, noise tends to be dampened and the predictability of returns given the price dividend ratio increases.

The DDG model predicts a tighter relation between aggregate stock market returns and the price-dividend ratio as the horizon at which returns are defined increases. A first evidence of the increasing explanatory power of the dividend-yield as the investment horizon increases is reported in Table (1). Here we report the slopes, the adjusted R^2, as well as the adjusted t-stats as in Valkanov (2003), of the following predictive regression

$$
r_{t:t+k} = \alpha_k + \beta_k \log\left(D_t / P_t\right) + \sigma \varepsilon_{t+k} \qquad \varepsilon_{t+k} \sim N(0,1)
$$

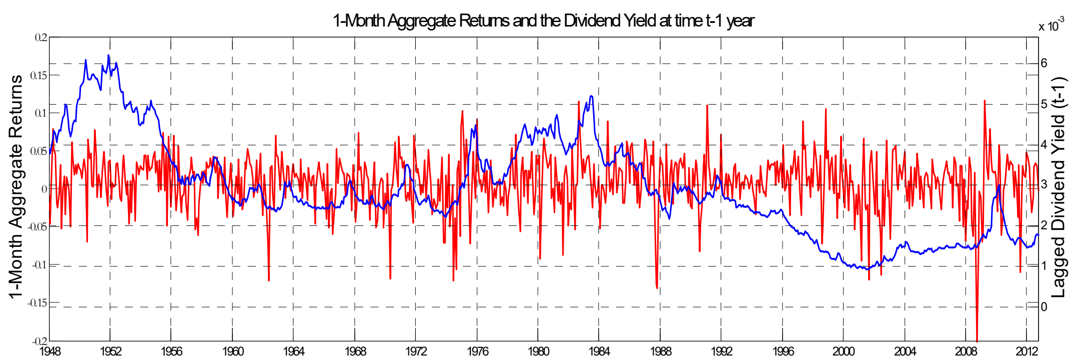
where $r_{t:t+k}$ the aggregate US stock market returns from t to $t+k$, D_t the aggregate dividend, P_t the index, ε_{t+k} an idiosyncratic error component and σ its corresponding risk.

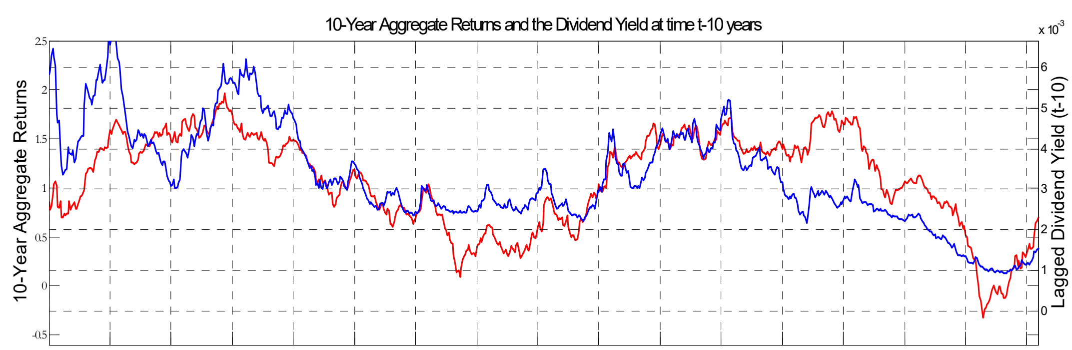
The sensitivity of the aggregate cumulative returns on the log dividend-yield β_k increases with the investment horizon. The same is true for the adjusted R^2, meaning, the longer the forecasting term, the higher the predictive power of the value-weighted dividend-yield This increasing monotonic relationship is visually confirmed in Figure (2), which reports the 1-year returns as well as the 10-year returns together with the dividend-price ratio. The top panel reports the lagged dividend yield (t-1 year) and the annual aggregate stock market returns in the US. On the other hand the bottom panel reports the lagged dividend yield (t - 10 years) the 10-year aggregate stock market returns in the US.

Table 1: The Predictive Power of the Dividend-Yield

This table reports the OLS estimates of the aggregate US stock market returns on the value-weighted dividend-price ratio. The sample is monthly and goes from 1946:01 to 2012:12. The first column reports the forecasting horizon. The second column reports the slope coefficients while the third the adjusted t-stats, i.e. $t/\sqrt{T}$ as in Valkanov (2003). The last column reports the adjusted R^2.

Horizon k	$\hat{\beta}$	$t/\sqrt{T}$	R^2
1	0.726	0.092	0.007
4	3.369	0.187	0.032
8	7.105	0.269	0.066
16	15.96	0.412	0.144
24	23.59	0.523	0.214
60	54.69	0.976	0.487





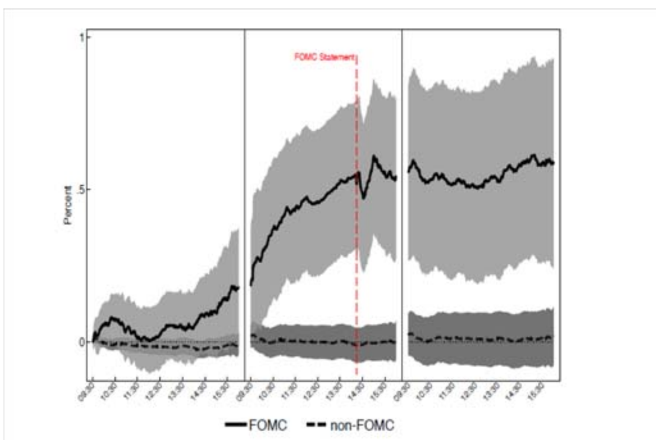
Aggregate Stock Market Returns and Lagged Dividend-Yield

5.2. *Anomalies*

Are Stock Market Returns Unpredictable ?

- Lucca Moench(2014)
  - Document large average excess returns on U.S. equities in anticipation of monetary policy decisions made at scheduled meetings of the Federal Open Market Committee (FOMC) in the past few decades.
- Cieslak et al. (2015)
  - Document that since 1994 the US equity premium follows an alternating weekly pattern measured in FOMC cycle time, i.e. in time since the last Federal Open Market Committee meeting.

The Data



5.3. *The Cross-section Evidence on Expected Returns*

The CAPM has important empirical implications for the cross sections of assets. If

$$E\left(r^i - r^f\right) = \beta_i E\left(r^M - r^f\right)$$

then heterogeneity in excess returns to different assets should be totally explained by the different exposure to a single common risk factor, the market excess returns.

Given a sample of observations on r_t^i, r_t^f, r_t^M,the β_i can be estimated first by OLS regression over the time series of returns, then the following second-pass equations can be estimated over the cross-section of returns:

$$\bar{r}_i = \gamma_0 + \gamma_1 \beta_i + u_i$$

Where $\bar{r}_i$are the average returns in the period over which the β_i have been computed.

If the CAPM is valid , then γ_0 and γ_1 should satisfy:

$$\gamma_0 = \bar{r}^f, \gamma_1 = \bar{r}^M$$

where $\bar{r}^M$ is the mean market excess return.

When the model is estimated with appropriate methods, the restrictions are strongly rejected (Fama-French(1992), Fama-McBeth). This evidence has paved the way to the estimation of multi-factor models of returns. Fama-French(1993) introduced a three-factor model based on the integration of the CAPM with a "small-minus-big" market value (SMB) and "high-minus-low" book-to-market ratio (HML).These factors are equivalent to zero-cost arbitrage portfolio that takes a long position in high book-to-market (small-size) stocks and finances this with a short position in low book-to-market (large-size) stocks. Jegadeesh and Titman(1993) discovered the importance of a further additional factor in explaining excess returns: momentum(MOM). An investment strategy that buys stocks that have performed well and sells stocks that have performed poorly over the past 3-to 12-month period generates significant excess returns over the following year. More recently Fama- French(2013) have extended the standard factors model based on the Market, SMB, HML and MOM, to include two more factors: RMW and CMA. RMW (Robust Minus Weak) is the return on a protfolio long on robust operating profitability stocks and short on weak operating profitability stocks, while CMA (Conservative Minus Aggressive) is the average return on a position long on conservative investment portfolios and short on aggressive investment. It is interesting to note that augmenting the CAPM with SMB and HML, does not challenge per se the CER model, which still hold as valid if the constant expected return model can be applied to the two additional factors. However, momentum provides direct evidence against the CER model as it indicates that the conditional expections of future returns is not constant.

5.4. *The behaviour of returns at high-frequency: non-normality and heteroscedasticity*

At small horizon (i.e. when k is small: infra-daily, daily, weekly or at most monthly returns) the following framework is supported by the data :

$$\begin{aligned} R_{t,t+k} &= \sigma_{k,t} u_{t+k} \\ \sigma_{k,t}^2 &= f(\mathcal{I}_t) \qquad u_{t+k} \sim IID\ \mathcal{D}(0,1). \end{aligned}$$

The following features of the model at high frequency are noteworthy:

1. The distribution of returns is centered around a mean of zero, and the zero mean model dominates any alternative model based on predictors.
2. The variance is time-varying and predictable, given the information set, $\mathcal{I}_t$, available at time t.
3. The distribution of returns at high frequency is not normal, i.e., $\mathcal{D}(0,1)$ may often differ from $\mathcal{N}(0,1)$

6. **The Implications of the new evidence**

6.1. *Asset Pricing with Predictable Returns*

Empirical work based on the DDG model has shown that the CER model does not provide the best representation of the data. This evidence opens a very interesting question on the determinants of time-varying expected returns. Different approaches have been used in finance to model time-varying expected returns, they are all understood within the context of a basic model that stems from the assumption of the absence of "arbitrage opportunities" (i.e. by the impossibility of making profits without taking risk). Consider a situation in which in each period k state of nature can occur and each state has a probability $\pi(k)$, in the absence of arbitrage opportunities the price of an asset i at time t can be written as follows:

$$P_{i,t} = \sum_{s=1}^{k} \pi_{t+1}(s) m_{t+1}(s) X_{i,t+1}(s)$$

where $m_{t+1}(s)$ is the discounting weight attributed to future pay-offs, which (as the probability π) is independent from the asset i, $X_{i,t+1}(s)$ are the payoffs of the assets (we have seen that in case of stocks we have $X_{i,t+1} = P_{t+1} + D_{t+1}$), and therefore returns on assets are defined as $1 + R_{s,t+1} = \frac{X_{i,t+1}}{P_{i,t}}$.For the safe asset, whose payoffs do not depend on the state of nature, we have:

$$
\begin{aligned}
P_{s,t} &= X_{i,t+1}\sum_{s=1}^{k}\pi_{t+1}(s)m_{t+1}(s) \\
1+R_{s,t+1} &= \frac{1}{\sum\limits_{j=1}^{m}\pi_{t+1}(s)m_{t+1}(s)}
\end{aligned}
$$

In general, we can write:

$$
\begin{aligned}
P_{i,t} &= E_t\left(m_{t+1}X_{i,t+1}\right) \\
1+R_{s,t+1} &= \frac{1}{E_t\left(m_{t+1}\right)}
\end{aligned}
$$

consider now a risky asset :

$$
\begin{aligned}
E_t\left(m_{t+1}\left(1+R_{i,t+1}\right)\right) &= 1 \\
Cov\left(m_{t+1}R_{i,t+1}\right) &= 1-E_t\left(m_{t+1}\right)E_t\left(1+R_{i,t+1}\right) \\
E_t\left(1+R_{i,t+1}\right) &= -\frac{Cov\left(m_{t+1}R_{i,t+1}\right)}{E_t\left(m_{t+1}\right)}+\left(1+R_{s,t+1}\right)
\end{aligned}
$$

Turning now to excess returns we can write:

$$
E_t\left(R_{i,t+1}-R_{s,t+1}\right)=-\left(1+R_{s,t+1}\right)cov\left(m_{t+1}R_{i,t+1}\right)
$$

Assets whose returns are low when the stochastic discount factor is high (i.e. when agents values payoffs more) require an higher risk premium, i.e. an higher excess return on the risk-free rate. Turning to predictability at different horizon, if you consider the case in which t is defined by taking two points in time very close to each other the safe interest rate will be approximately zero and m will not vary too much across states. The constant expected return model (with expected returns eqaul to zero) is compatible with the no-arbitrage approach at high-frequency. However, consider now the case of low frequency, when t is defined by taking two very distant points in time; in this case safe interest rate will be diffferent from zero and m will vary sizeably across different states. The constant expected return model is not a good approximation at long-horizons. Predictability is not a symptom of market malfunction but rather the consequence of a fair compensation for risk taking, then it should reflect attitudes toward risk and variation in market risk over time. Different theories on the relationship between risk and asset prices should then be assessed on the basis of their ability of explaining the predictability that emerges from the data.

Also, different theories or return predictability can be interpreted as different thoeries of the determination of m. On the one hand we theories of m based on rational investor behaviour, on the other hand we have alternative approaches based on psycological models of investor behaviour. Our main interest is to show how the predictability of returns can be used for optimal portfolio allocation purposes, rather than on discriminating between the possible sources of predictability.

7. Quantitative Risk Management and the behaviour of returns at high-frequency

Once the portfolio weights ($\hat{\mathbf{w}}$) are chosen, possibly exploiting the predictability of the distribution of the relevant future returns, the distribution of a portfolio returns can be described as follows:

$$\begin{aligned} R^p &\sim \mathcal{D}\left(\mu_p, \sigma_p^2\right) \\ \mu_p &= \boldsymbol{\mu}'\hat{\mathbf{w}} \qquad \sigma_p^2 = \hat{\mathbf{w}}'\boldsymbol{\Sigma}\hat{\mathbf{w}} \end{aligned}$$

Having solved the portfolio problem and having committed to a given allocation described by $\hat{\mathbf{w}}$, there is a different role that econometrics can play at high frequencies: measuring volatility and providing information on portfolio risk. As our simple specification of the previous section shows, noise is not predictable but its volatility is. The role of econometrics in applied risk management is best seen through a different statistical model of high frequency returns. When k is small (i.e., when one is considering infra-daily, daily, weekly or at most monthly returns) the following framework is normally referred to:

$$\begin{aligned} R_{t,t+k} &= \sigma_{k,t} u_{t+k} \\ \sigma_{k,t}^2 &= f(\mathcal{I}_t) \qquad u_{t+k} \sim IID\ \mathcal{D}(0,1). \end{aligned}$$

The following features of the model at high frequency are noteworthy:

1. The distribution of returns is centered around a mean of zero, and the zero mean model dominates any alternative model based on predictors.

2. The variance is time-varying and predictable, given the information set, $\mathcal{I}_t$, available at time t.

3. The distribution of returns at high frequency is not normal, i.e., $\mathcal{D}(0,1)$ may often differ from $\mathcal{N}(0,1)$

Given these features of the data, econometrics can still be used at high frequency to assess the risk of a given portfolio. In particular, we shall investigate the role of econometrics for deriving the time-varying Value-at- Risk (VaR) of a given portfolio.

8. The Plan of the book. Predictive Models in Finance

We shall begin our journey by considering asset allocation under the constant expected returns model. We shall then discuss the limitations of this model and consider alternatives that will be based on different specifications of the relevant predictive model. In particular we shall consider in turn asset allocation at different horizon with models featuring predictability of expected returns, and Risk Management with models featuring the predictability of the distribution of returns.Given that financial decisions are based on the predicted distribution of returns they require a model of future behaviour of the variables of interest. Predictive models are statistical models of future behaviour in which relations between the variables to be predicted and the predictors are specified as functional relation determined by parameters to be estimated. Predictive models can be univariate, when there is only one variable of interest, or multivariate when we have a vector of variables of interest.

Predictive models considered in this book will be special cases of this general specification:

$$\begin{aligned}
\mathbf{r}_{t,t+k} &= f\left(X_t^{\mu}, \Theta_t^{\mu}\right) + \mathbf{H}_{t+k}\boldsymbol{\epsilon}_{t+k} && (1)\\
\boldsymbol{\Sigma}_{t+k} &= \mathbf{H}_{t+k}\mathbf{H}'_{t+k}. \\
\boldsymbol{\Sigma}_{t+k} &= g\left(X_t^{\sigma}, \Theta_t^{\sigma}\right) + + \sum\nolimits_{j=1}^{q} \mathbf{B}_j \boldsymbol{\Sigma}_{t+k-j}\mathbf{B}'_j, && (2)\\
\boldsymbol{\epsilon}_{t+k} &\sim \mathcal{D}\left(\mathbf{0}, \mathbf{I}\right)
\end{aligned}$$

where $\mathbf{r}_{t,t+k}$ is the vector of returs between time t and time t+k in which we are interested, X_t^{μ} is the vector of predictors for the mean of our returns that we observe at time t, f specifiies the functional relation (thant is potentially time-varying) between the mean returns and the predictors that depends also on a set of parameters Θ_t^{μ}, the matrix $\mathbf{H}_{t+k}$ determines the potentially time varying variance-covariance of the vector of returns..The process for the variance is predictable as there is a functional relation determining the relationship between $\mathbf{H}_{t+k}$ and a vector of predictors X_t^{σ} that is driven by a vector of unknown parameters Θ_t^{σ}.

Our first look at the data clearly show that the appropriate specification of the general predicitive model depends on the horizon at which returns are defined. Consider, for example, the problem of univariate modelling of stock market returns. When k is small and high-frequency returns On the one hand, in the simple asset allocation model, the econometric framework considered for returns is as follows:[2]

$$\begin{aligned}
r_{t,t+k} &= 0 + \sigma_{t+k} u_{t+k} \qquad u_{t+k} \sim IID\ \mathcal{D}(0,1), \\
\sigma_{t+k}^2 &= \omega + \alpha\sigma_{t+k-1}^2 + \beta u_{t+k-1}^2, \quad |\alpha+\beta| < 1
\end{aligned}$$

[2]During the lectures, it is possible that the sum of IIDness of returns and of normality has also been denoted as $u_{t+k} \sim n.i.d.(0,1)$. Note that IID $N(0,1)$ and n.i.d.(0, 1) have identical meaning.

This is a model that feautres no predictability in the mean of r returns (the expected future return at any horizon is constant at zero), but there is predictability in the variance of returns that it is mean reverting towards a long-term value of $\omega/(1-\alpha-\beta)$. No assumption of normality is made for the innovation innovation in the process generating returns. Consider now the case of large k, i.e. long-horizon returns (note that in the continuously compounded case, $r_{t,t+k} \equiv \sum_{j=1}^{k} r_{t,t+j}$), in this case the relevant predictive model can be written as follows:

$$r_{t,t+k} = \alpha + \boldsymbol{\beta}'\mathbf{X}_t + \sigma u_{t+k} \qquad u_{t+k} \sim IID\ \mathcal{N}(0,1),$$

where $\mathbf{X}_t$ is a set of predictors observed at time t. In this case we have that returns feature predictability in mean, constant variance and the innovations are normally distributed. As the horizon k increases, predictability increases and therefore the uncertainty related to the unexpected components of returns decreases (i.e., the annualized variance of returns is a downward sloping function of the horizon). Moreover—as we have already discussed—the dependence of $\sigma_{t,k}$ on time (i.e., its time-varying nature) declines and long-horizon returns can be described as a (conditional) normal homoskedastic processes. In the short-run noise dominates and modelling returns on the basis of fundamentals is very difficult. However, as the horizon increases fundamentals become more important to explain returns and the risk associated to portfolio allocation based on econometric models is reduced. The statistical model becomes more and more precise as k gets large.

References

[1] Campbell J.Y. and R.J. Shiller (1987) "Cointegration and Tests of Present Value Models", Journal of Political Economy, 95, 5, 1062-1088

[2] Campbell, John Y., and Robert Shiller, 1988, Stock Prices,Earnings, and Expected Dividends, *Journal of Finance*, 43, 661-676.

[3] Campbell, John Y., and Robert Shiller, 1988, The Dividend-Price Ratio and Expectations of future Dividends and Discount Factors, *Review of Financial Studies*, 1:195-228

[4] Campbell, John Y., and Robert Shiller, 1998 Valuation Ratios and The Long-Run Stock Market Outlook, 1998, *Journal of Portfolio Management*

[5] Campbell, John Y., and Robert Shiller, 2001 Valuation Ratios and The Long-Run Stock Market Outlook, an update, Cowles Foundation DP 1295

[6] John Y. Campbell and Tuomo Vuolteenaho (2004), Inflation Illusion and Stock Prices (Cambridge: NBER Working Paper 10263)

[7] Cochrane, J. H. The Dog that Did Not Bark: A Defense of Return Predictability. *Review of Financial Studies*, 21 (2008), 4, 1533-1575.

[8] Cochrane J.(1999) New Facts in Finance

[9] Ellenberg J.(2015) "How Not to be Wrong. The Hidden Maths in Everyday Life"

[10] Fama E.(1970)

[11] Fama, Eugene and Kenneth R. French, 1988, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics*, 22, 3-26.

[12] Fama, Eugene and Kenneth R. French, 2013, A Five-Factor Asset Pricing Model

[13] Ferreira and Santa Clara (2010)

[14] Lander J., Orphanides A. and M.Douvogiannis(1997) Earning forecasts and the predictability of stock returns: evidence from trading the S&P, *Board of Governors of the Federal Reserve System*, http:www.bog.frb.fed.org

[15] Lettau, Martin, and Sydney Ludvigson, 2005, Expected Returns and Expected Dividend Growth, *Journal of Financial Economics*, 76, 583-626

[16] Lettau, Martin, and Stijn Van Nieuwerburgh, 2008, Reconciling the Return Predictability Evidence, *Review of Financial Studies*, 21, 4, 1607-1652

[17] Littermannn B.(2003), Modern Investment Management. An Equilibrium Approach, *Wiley Finance*

[18] Meznly, Lior; Tano Santos, and Pietro Veronesi, Understanding Predictability, *Journal of Political Economy, 112 (2004),1,1-47.*

[19] Franco Modigliani and Richard Cohn (1979), Inflation, Rational Valuation, and the Market, *Financial Analysts' Journal.*

[20] Newey, W. K. and K. D. West. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix., *Econometrica*, 55 (1987), 3, 703-08.

[21] Newey, W. K., and D. K. West. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61 (1994), 631-653.

[22] Robert J. Shiller. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71 (June 1981), 421-436.

[23] Taleb N.(2012) *Antifragile. Things that Gain from Disorder*, Random House

[24] Valkanov, R. Long-Horizon Regressions: Theoretical Results and Applications. *Journal of Financial Economics*, 68 (2003), 201–232. 33