An Historycal Perspective

Carlo Favero

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

5 Predictive Models in Finance

An Historical Perspective

- 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 classical Constant Expected Returns (CER) model, when a simple econometric model serves the purpose of modelling returns at all horizon, to move to Time-Varying Expected Returns (TVER) model where different econometric models for returns are to be adopted according to the different horizon at which they are defined.

The view from the 1960: Efficient Markets and CER

The history of empirical finance starts with the "efficient market hypothesis"(Fama(1970):).

- expected returns are constant,
- 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 unpredictable: any predictability is a statistical artifact or cannot be exploited after transaction costs;
- volatility of returns is constant.

Time-Series Implications

$$r^{i}_{t,t+1} = \mu^{i} + \sigma^{i}\epsilon_{it} \qquad \epsilon_{it} \sim NID\left(0,1\right)$$

$$Cov\left(\epsilon_{it},\epsilon_{js}\right) = \left\{ \begin{array}{ll} \sigma_{ij} & t = s \\ 0 & t \neq s \end{array} \right. .$$

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).

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.

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

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

The cross-section of returns

Cross-sectional heterogeneity is related to a single factor, the market factor, and the CAPM determines all the cross-sectional variation in returns. The statistical model that determines all returns r_t^i and the market return r_t^m,is:

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

where r_t^{rf} is the return on the 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 $\mu_i = \beta_i \mu_m$.

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.

Implications for Asset Allocation

- Optimal asset allocation can be 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,
- 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.

Measuring Risk

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.

Measuring Risk

If the distribution of returns is normal, then α-percent VaR_α is obtained as follows (assume $\alpha \in (0,1)$):

$$\Pr\left(R^p < -VaR_\alpha\right) = \alpha \Longleftrightarrow \Pr\left(\frac{R^p - \mu_p}{\sigma_p} < -\frac{VaR_\alpha + \mu_p}{\sigma_p}\right) = \alpha$$

$$\Longleftrightarrow \Phi\left(-\frac{VaR_\alpha + \mu_p}{\sigma_p}\right) = \alpha,$$

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

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

and, given that μ_p and σ_p are constant over time, VaR_α is also constant over-time.

Measuring Risk

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*:

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

Empirical Challenges to the CER model

- 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.

Empirical Challenges to the CER model: the DDG model

- 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 is the excessive volatility of asset prices and returns which is clearly illustrated in Shiller(1981).

A simple log price-dividend ratio framework

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$(0 < \rho < 1)$.

A simple log price-dividend ratio framework

By forward recursive substitution one obtains:

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

which shows that the $(p_t - d_t)$ measures the value of a very long-term investment strategy (buy and hold) which is equal to the stream of future dividend growth discounted at the appropriate rate, which reflects the risk free rate plus risk premium required to hold risky assets.

Rejection of the CER

- Under the null of the CER the price-dividend ratio should be completely determined by the process generating dividends and determining their expectations. But stock prices are too volatile to be determined only by expected dividends (see Shiller(1981) and Campbell-Shiller(1987)).
- The price-dividend ratio is useful to predict returns, the more so the longer the horizon at which returns are defined.

The empirical evidence

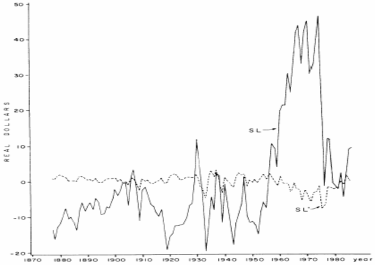
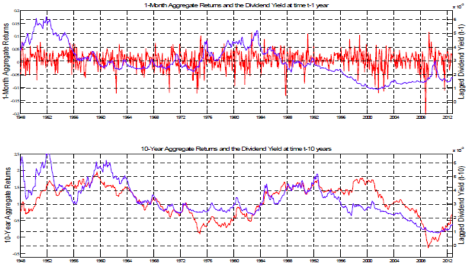


FIG. 2.—Stock market: deviations from means of actual spread (SL_t = Price$_t$ − θ · Dividend$_{t-1}$) and theoretical spread SL_t', θ = 12.195.

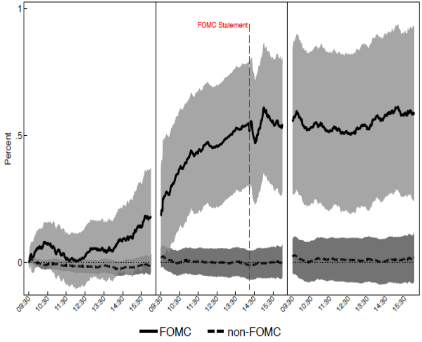
The empirical evidence



Aggregate Stock Market Returns and Lagged Dividend-Yield

Anomalies:an example

- 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.



Multiple Factor Models

The CAPM has implication for the cross section 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)

- The different exposure to a single factor model cannot explain the observed cross-sectional behaviour of returns.
- This evidence 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) dis-covered 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.
- 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 aditional factors. However, momentum provides direct evidence against the CER model as it indicates that the conditional expections of future returns is not constant.

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 :

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

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 $\mathcal{N}(0,1)$

The implications of the new evidence

The new evidence has brought about importnat implications both for asset pricing and risk management

Asset Pricing with Predictable Returns

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:

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

- $m_{t+1}(s)$ 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 $1 + R_{s,t+1} = \frac{X_{i,t+1}}{P_{i,t}}$.

Asset Pricing with Predictable Returns

For the safe asset, whose payoffs do not depend on the state of nature, we have:

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

so we can write:

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

Asset Pricing with Predictable Returns

For risky assets

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

- Consider now the case where the period t is made by two points in time very close to each other (a short holding horizon), in this case m_{t+1} can be safely approximated by a constant (very close to one) and excess returns are not predictable.
- As the point in time that define the period becomes further and furthes separated, then time variation in m cannot be discounted anymore and future excess returns becomes predictable if their covariance with m is predictable.

In fact, 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.

Risk Management

Once the portfolio weights (**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:

$$R^p \sim \mathcal{D}\left(\mu_p, \sigma_p^2\right)$$
$$\mu_p = \boldsymbol{\mu}'\mathbf{w} \qquad \sigma_p^2 = \mathbf{w}'\circ\mathbf{w}$$

Having solved the portfolio problem and having committed to a given allocation described by **w,** there is a different role that econometrics can play at high frequencies: measuring volatility and providing information on portfolio risk. Altough noise is not predictable, its volatility is.

Predictive Models in Finance

Predictive models are special cases of this general specification:

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

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 (that 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

Predictive Models in Finance

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.
When k is small and high-frequency returns

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

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 $\omega / \left(1 - \alpha - \beta\right)$. No assumption of normality is made for the innovation innovation in the process generating returns.

Predictive Models in Finance

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 feauter predictability in mean, constnat variance and the innovations are normally distributed.