# The CAPM and Regression Tests

Elmar Mertens*

October 12, 2002

## Contents

**1 A warm-up in three equations** 3

**2 The CAPM** 4
- 2.1 Derivation . . . . . 4
  - 2.1.1 Setting . . . . . 4
  - 2.1.2 Assumptions and Derivation . . . . . 4
  - 2.1.3 Key Insight of the CAPM . . . . . 7
- 2.2 Testable Implications . . . . . 9

**3 Regression tests of the CAPM** 11
- 3.1 Additional assumptions . . . . . 11
  - 3.1.1 Empirical Assumptions . . . . . 11
  - 3.1.2 Economic Assumption . . . . . 14
- 3.2 Cross-sectional regressions (Fama-MacBeth) . . . . . 15
  - 3.2.1 Average return regressions . . . . . 15
  - 3.2.2 The Fama-MacBeth method . . . . . 18
  - 3.2.3 Fama-MacBeth Recipe . . . . . 20
  - 3.2.4 SML estimated from Cross-Sectional Regressions . . . . . 23
- 3.3 Time-series regressions (GRS) . . . . . 24
  - 3.3.1 Cross-sectional restrictions, again . . . . . 24
  - 3.3.2 The GRS Test . . . . . 25
  - 3.3.3 GRS Recipe . . . . . 28
  - 3.3.4 SML estimated from Time-Series Regressions . . . . . 31

**A Notation** 32

---

*This note has been written for the course *Portfolio Theory and Capital Markets* at the University of Basel. For correspondence: Elmar Mertens, University of Basel, Wirtschaftswissenschaftliches Zentrum (WWZ), Department of Finance, Holbeinstrasse 12, 4051 Basel, Phone +41 (61) 267 3309, email `elmar.mertens@unibas.ch`

**List of Figures** 32

**References** 33

# 1 A warm-up in three equations

This note describes how the CAPM can be derived and how its implications can be tested using OLS regressions. The three equations below are at the heart of these issues. From previous sessions you should at least know the first two. For starters, let us have a look at them and recollect what we already know:

$$E\left[\tilde{R}_i\right] = R_f + \beta_{i,E}\left(E\left[\tilde{R}_E\right] - R_f\right) \tag{1}$$

$$\tilde{R}_{i,t} = a_{i,t} + \beta_{i,M}\tilde{R}_{M,t} + \tilde{\varepsilon}_{i,t},\ E\left[\tilde{\varepsilon}_{i,t}\right] = 0 \tag{2}$$

$$E\left[\tilde{R}_i\right] = R_f + \beta_{i,M}\left(E\left[\tilde{R}_M\right] - R_f\right) \tag{3}$$

In equation (2), please note that we call the intercept $a$ and not $\alpha$ in order to avoid confusion. Still, many books use also $\alpha$, for example Elton and Gruber 1995, p. 152.

---

**QUIZ:**

- Which of the above equations do you already know? Where do they come from?
- Characterize the above equations!
  - Are they just a statistical description of data, an economic model etc.?
  - What do they imply? Do they impose testable restrictions?
- Are the above equations mutually consistent?
- Suppose returns are not normal and investors care not about mean-variance. Define an arbitrage as having $E[R] > R_f$ and $\sigma = 0$ (this is actually a very strong form of arbitrage possibility, see Ingersoll 1987, Chapter 3 formalization of arbitrage as "some return for no risk".). What can you say about the above equations and their content in the absence of arbitrage?
- Is the notion of arbitrage described above consistent with the meaning of "arbitrage" employed by Elton and Gruber (1995, p.299) in their derivation of the CAPM?

---

# 2 The CAPM

As you may have already noted above, we assume the existence of a riskfree asset with return $R_f$. By "existence" of the riskfree asset we mean that there is unrestricted borrowing and lending at that rate, if not stated otherwise. We do so mostly as a matter of convenience for our notation and with an eye on the subsequent tests which we describe in terms of excess returns[1]. Please note that the derivation below applies both to the Sharpe-Lintner-Mossin CAPM (assuming unlimited lending and borrowing at $R_f$) and the Black CAPM (risky assets only, in this case replace $R_f$ with $\tilde{R}_{0\beta M}$). The regression tests described afterwards focus solely on the Sharpe-Lintner-Mossin CAPM.

## 2.1 Derivation

### 2.1.1 Setting

Investors face an investment decision over *one* period of time[2]. Investors *asses* the distribution of returns in the sense that they can act as if they knew their true parameters. This is not yet intended to preclude whether their expectations are subjective or objective, it is neither meant to understand that there is knowledge of some fixed distribution parameters[3]. The point is that they do not need to *estimate* them and do not need to account for estimation errors. We disregard production and consumption of goods (hence we will get a *partial* equilibrium model. For reasons unknown to us, Elton and Gruber (1995) speak of "general" equilibrium). There are multiple risky assets, their number is possibly finite[4]. There may or may not be a riskfree asset.

### 2.1.2 Assumptions and Derivation

**Perfect Market** This is about assuming away any kind of frictions in trading and holding assets: No transaction costs, infinitely divisible asset holdings, no taxes[5]. In addition, every investor is a price-taker, which means that he does

[1]The tests can also be conducted using raw returns. In any case they need to be real returns. The construction of real returns is not trivial given the scarcity of CPI data in high-frequencies, hence it is standard to proxy for real excess returns using nominal excess returns. See our remarks in section 3.1. For further readings please see Campbell, Lo, and MacKinlay (1997, p. 182).

[2]Still, some authors, e.g. Magill and Quinzii (1998), call the CAPM the "two-period" model. Indeed there are two dates, but between them, there is only one period.

[3]We consciously sidestep to say that they "know" expected values and the like. Later we will assume that every investor has the same expectations. Taking this together with a knowledge of *fixed* values of distribution parameters, this could lead us on a slippery slope: If expected returns are known in advance, what sense is there for a pricing model?

[4]Remember this when we come to the APT!

[5]Actually, taxes are permissible as long as they do not distort returns in that long/short returns will differ, or that it matters whether a return is accrued from dividends, interest payments or capital

not anticipate prices to change because of his behavior. In other words, no single investor has the power to move the market, nor does any particular group of investors. Please note that prices are still determined by the actions of investors, they are the result of the portfolio decisions by *all* investors not a single investor or a single group of investors.

Let us also assume under the label of the perfect market, that unlimited short-sales of risky assets (and if it exists, of the riskfree asset[6]) are possible.

Please note that the notion of a perfect market is different from that of a complete market.

$\Rightarrow$ The return on any asset is identical for every investor, regardless of whether he already owns it – which means it was part of his endowment – or whether he has to engage in a transaction to buy (or sell) it or whether his transactions have a particular size or whether he is particularly wealthy etc. Shorting an asset just changes the sign of its return, not its absolute value. Each asset can be hold long or short in every perceivable quantity (notwithstanding the constraint on total portfolio holdings imposed by each investor's budget constraint and the equilibrium constraint on total asset supply).

**Mean-Variance Portfolio Choice** Now we need some assumption which makes investor choose their portfolios based on mean – which they like – and variance – which they do not like. High means and low variances are sought after by risk-averse investors using Von-Neumann-Morgenstern expected utilities. To be sure that they base their decisions *only* on mean and variance we need either to assume a specific utility function (quadratic[7]) or a specific distribution. As a matter of convenience, we assume multivariate normality of returns, which ensures that every linear combination, i.e. portfolios, of returns is normally distributed. But returns need not be normally distributed for compatibility of mean-variance portfolio theory and portfolio choice of risk-averse expected-utility investors (see Ingersoll 1987, Chapter 4, in particular Appendix B, Owen and Rabinovitch 1983 for more general distributions).

Please note that we do not assume, that every investor has the same risk aversion (or parametrically the same type of utility function.)

$\Rightarrow$ Now we know that every investor will only hold a portfolio which he perceives as being mean-variance efficient (MVE).

**Complete Agreement** We want that every investor faces the same set of mean-variance portfolios. Hence they must all face the same set of means and

gains. Equally, transactions should not be taxed.

[6]Shorting the riskfree asset is more conventionally known as "borrowing money".

[7]See Elton and Gruber (1995, p. 219) for some of the undesirable features of such a utility function.

(co)variances. We assume that – at least in equilibrium – everybody's assessment of the return distribution is identical.

As an aside: This is also often referred to as "homogenous expectations". But if the latter is understood in the sense that there is one fixed set of expectations, it is clearly misleading, see our FN 3. There is room for disagreement and differing expectations in disequilibrium! Actually there is even more room for that as we would like, the CAPM has simply nothing to say about disequilibrium (like many other economic models). The pity is that the CAPM does not even try to specify how demand and supply curves would look like. It only tells us something about prices in *equilibrium*[8].

⇒ Every investor chooses from the same set of efficient portfolios – at least once there is complete agreement on the assessment of means and variances in equilibrium. (We can draw the same minimum-variance frontier in $(\mu^e, \sigma)$ space for every investor.)

**Market Equilibrium** In equilibrium, supply of assets equals demand. The market portfolio is defined as the portfolio of assets which are in positive net supply, weighted by their market capitalizations (which are a function of their prices). Usually it is assumed that the riskfree instrument is in zero net supply. On the demand side, the *net* holdings of all investors equal aggregate net demand.

⇒ In equilibrium, supply equals demand. The market portfolio is the portfolio of all the efficient portfolios held by the individual investors weighted by their relative wealth. Each investor's net wealth is positive, hence the market portfolio is a convex combination of efficient portfolios, from efficient set mathematics we know, that such a convex combination is also an *efficient* portfolio.

**Variations** See Elton and Gruber (1995, Chapter 14) or Fama (1976) for further variations on lending/borrowing at the riskfree rate and short-sales.

---

**QUIZ:**

- Relate each particular aspect of the perfect-market assumption on the consequences discussed above!
- Which role plays the riskfree rate in our derivation? Did we derive the Sharpe-Lintner-Mossin- (SLM) or the Black-CAPM?
- Why is it reasonable to assume that the riskfree rate (if it exists) is in zero net supply? What happens if it is in positive supply?
- How does the efficient set of all assets look like with and without riskfree asset?

[8] As we will see in the next assumption, the quantity of assets supplied is purely exogenous.

- Why is it important that the market portfolio is a *convex* combination of efficient portfolios? Would equation (3) still hold, if it were *some* combination?
- Draw the minimum-variance frontier and the efficient set of all assets in $(\mu^e, \sigma)$ space when the riskfree asset ...
  - exists
  - does not exist
- From previous sessions you should know two types of two-fund separation: Black and Tobin. What are their differences? Are they mutually consistent?
- What can we say about the correlation of two efficient portfolios when the riskfree asset ...
  - exists
  - does not exist

  (Hint: Draw the efficient set and use the graph in your argument.) What if we look at two minimum-variance portfolios? Can we say something similar about their covariance?

---

### 2.1.3 Key Insight of the CAPM

The market portfolio is an efficient portfolio hence, equation (1) applies to the market portfolio and we can write it as in (3). It is tantamount to understand, that (1) holds in any case, ex-post or ex-ante and with any sort of return distributions[9] and investors. The only content of the CAPM is to identify the market as an efficient portfolio (Roll 1977)! This assertion holds with or without the existence of a riskfree asset, but we need unlimited short-sales so that we can apply the results from efficient set mathematics, in particular equation (1). Equation (3) is also known as the *Security Market Line* (SML).

The derivation presented here does not appeal to specific forms of utility function, in particular we do not explicitly mention the market level of risk-aversion as we did in our first class on the CAPM. Instead, we stress the link between the combination of efficient portfolios and the market portfolio. This prepares the ground for Critique of Roll (1977) and helps to view the CAPM tests as tests on the efficiency of a particular portfolio.

With an eye on the tests described in the remainder of this note, we will effectively be testing the efficiency of a particular portfolio. This implies that the methods described below can also be used in applications which are not concerned with the

[9] Ex-ante, finite (co)variances are required, though.

empirical validity of the CAPM. For instance, an asset manager might want to know whether a particular investment will increase the efficiency of his portfolio compared to this benchmark.

---

**QUIZ:**

- Discuss the assertion of Elton and Gruber (1995, p. 312, "Short Sales Disallowed"). Is it really possible to establish the efficiency of the market portfolio without short-sales? Does equation (1) still hold?
- Draw a sketch of the SML with five assets.
- Sketch a scatter plot of returns on (true) betas with five assets. Draw the SML. (Suppose the CAPM holds.)
- From previous classes you should know the Capital Market Line (CML) (Elton and Gruber 1995, p. 297):
  - What is the difference between CML and SML?
  - Is there a CML if there is no riskfree asset? If so, how does it look like?
- In your own words: What does the CAPM tell us about realized returns?
- What assets are priced by the CAPM?
- Over which return interval is the CAPM defined?
- $\beta_{i,M}$ is often called the "systematic risk" of stock $i$. This is often meant to imply that holding assets is compensated not in relation to their own variance, but only with respect to their "undiversifiable" risk, a.k.a. the "systematic" risk.
  - How would the CAPM look like if all assets were to have zero correlation? What is their systematic risk? Is the market premium zero?
  - What if there are just a few stocks, say $n = 3$, in the economy. Obviously their is not much to diversify. Would the CAPM work? How?
  - Take "systematic risk" as defined by $\beta_{i,M}$. List all its determinants! How is it influenced by properties of asset $i$?
- The CAPM says that the market portfolio is an efficient portfolio. Using mean-variance mathematics it can be shown, that for each asset $i$, there is one portfolio on the (minimum-variance) frontier where this asset has zero weight (Ingersoll 1987, p. 87). Could the market portfolio be such a portfolio? Under which conditions? Can you say anything about the price of asset $i$?

- In a previous session, you were introduced to the concept of shortfall probabilities and Roy-portfolios (Elton and Gruber 1995, Chapter 11). Suppose that investors minimize shortfall risk, possibly using different threshold values. Can you derive something like the CAPM from the aggregation of their behavior? If so, do you need additional assumptions, which (or do you need fewer assumptions)?

---

## 2.2 Testable Implications

The CAPM is above all a "model". Its assumptions are not intended to give an accurate description of reality, but to focus on what might be salient and tractable features of real-world markets yielding a solid description of asset prices. Like any other model, it should not be judged by the validity of its assumptions but by the accuracy of its predictions[10].

Before we discuss the testable implications of the CAPM. Let us rewrite the CAPM equation in excess return notation. Define excess returns as

$$\tilde{r}_{i,t} \equiv \tilde{R}_{i,t} - R_f$$

Then we write expected excess returns as $E[\tilde{r}_i] \equiv \mu_i^e$ and define the risk premium of the market portfolio $\lambda_M \equiv E[\tilde{r}_M]$. As we will see in the regression tests below, we will not always use the time-series average of the market portfolio as the estimate of $\lambda_M$. We can rewrite (3) so that

$$\mu_i^e = \beta_{i,M}\lambda_M \tag{4}$$

Hence, $\lambda_M$ is the expected excess return of a portfolio with a beta of one. Figure 1 graphs the SML in $(\mu^e, \beta)$ space.

**Efficiency of the market portfolio:**

1. There is a *linear* relationship between expected returns and market beta ($\beta_{i,M}$). In this linear relationship, beta explains fully the *cross-sectional* variation in expected returns. In addition to beta, there is no other variable (or "factor"[11]) which can explain expected returns

[10] One word at caution is order, though. Perhaps it is reasonable to say that assumptions need not be valid, but they should not be implausible either. What if we get an apparently accurate description of asset prices which is related to the birthdays of company CEO's? In search for parsimonious models we need to make simplifications, but we should not assume away anything important. This is indeed a thin line where we tread on in order to improve our models if they fail to yield a satisfactory description of the world.

[11] To be precise: The word "factor" is usually reserved for variables like the market or other time-series, usually portfolio returns, and not the sensitivity of expected returns to them. Actually, this is also a matter of whether we regress returns on betas or on the market. But we are getting ahead of ourselves. For the time being, just note that "factor" will later take a particular meaning. Hence the word "variable" is more appropriate here.

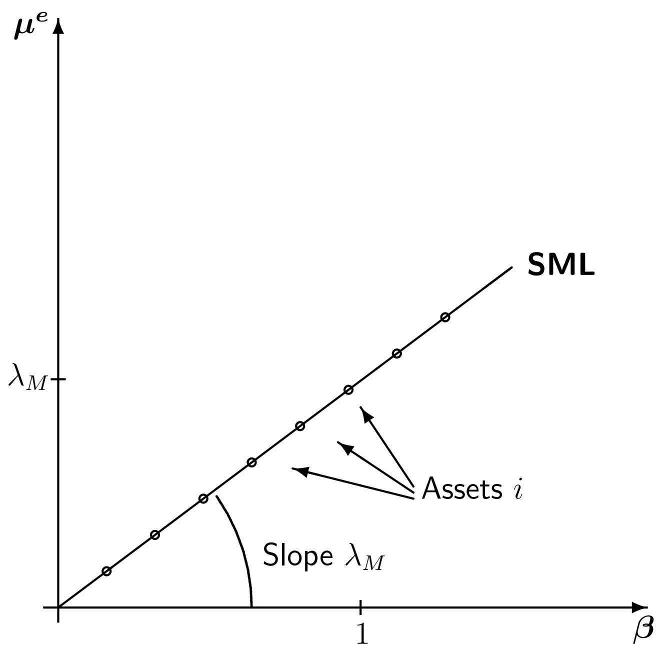


Figure 1: Security Market Line

2. The excess return[12] on a zero-beta portfolio is zero
3. The risk premium of the market is positive $\lambda_M > 0$

---

**QUIZ:**

- We already know from our first class on stylized facts that (expected) returns can be explained by many factors, for example firm size or scaled stock prices like P/E. Can we already refute the first implication, that beta explains fully the variation in expected returns?
- Which time-series restrictions are imposed by the CAPM? In particular, what factor structure of the "return generating process" is implied by the CAPM?
- Compare the CAPM (equation 3) with the Market Model (equation 2)

---

[12] If there is no riskfree asset, we would have to measure excess returns with respect to the market's zero-beta portfolio. This is not trivial as the latter needs to be estimated!

# 3 Regression tests of the CAPM

In this section we describe two types of regression tests: Cross-sectional and time-series regressions. In both cases, bear in mind that the CAPM implies a *cross-sectional* restriction and no time-series restrictions! Remember this when we come to devising a time-series regression test. But before we can get started with the empirical work, we have to make some more assumptions.

## 3.1 Additional assumptions

The CAPM is a model of expected returns in a one period-economy. In reality, we observe a history of asset prices and other variables from which we can compute their realized returns over various holding periods. In our tests we want to draw inferences from historical data on values the model's parameters, such as $\beta_{i,M}$ or $\lambda_M$. But before we can estimate these parameters, we need to assume some distribution theory which relates the realized values we observe to the parameters we want to estimate.

Please note, that these assumptions are not part of the CAPM, they are part of testing the CAPM. Consequently, we do not only test the CAPM, we always test whether the CAPM *and* these additional assumptions are true. This is generally known as a *Joint Hypothesis*. If the test does reject such a joint hypothesis, we cannot distinguish – without further testing – whether it failed because the CAPM does not hold or whether any of the other assumptions were wrong. In particular, we would like to discuss the following assumptions:

### 3.1.1 Empirical Assumptions

**Stationary return distribution** Investors in the CAPM know the return distribution over one particular investment period. In order to estimate the parameters of that distribution it is convenient, that it is stationary[13]. In addition, we assume that returns are drawn independently over time. Together this leads to the assumption of *identically and independently distributed* (i.i.d.) returns. Now we can use standard sampling theory and the central limit theorem to estimate the moments of return distributions, for example the time-series average of realized returns is our estimate of the expected value or returns. Please note that this amounts to assuming that *expected returns are constant over time*. This is clearly not implied by the CAPM itself! Modern tests of the CAPM[14] test conditional versions which weaken the assumption of constant expected-returns with considerable success (Jagannathan and Wang 1996).

In order to have a distribution theory of our regression estimators in a *finite-sample* – and also with an eye on the role of multivariate normality in the CAPM

[13]This means that the distributions parameter do not change over time. At each interval, returns are drawn from the same distribution.

[14]And extensions such as the ICAPM (Merton 1973).

derivation – it is convenient to assume that returns are *normally, identically and independently distributed* (n.i.i.d.). Please note that OLS estimates and their standard errors apply also for non-normally distributed returns. Neither does the normal distribution feature in the Gauss-Markov theorem[15] But without normality, we are left with an only asymptotically valid distribution theory for our regression estimators. Still, the GRS test discussed below seems to be reasonable robust to common levels of non-normality[16].

**Market Proxy** "The" market portfolio can be easily identified in theory. This is where the major contribution of the CAPM relies on. But in practice it is unobservable. Please remember that the CAPM covers all wealth and does not distinguish between different types of financial instruments[17]. This is the focus of Roll's Critique (Roll 1977). Effectively, the tests described below test only whether the proxy used for the market is an efficient portfolio. Still, the work of Stambaugh (1982) suggests that the results of such tests are largely invariant when "non-standard" assets like bonds, real estate and durable goods are included. Of course, this does not tell us whether they are good proxies in the sense that we would still get a similar result when using the real market portfolio.

Usually, broad, value-weighted (VW) stock market indices are used such as the CRSP value-weighted portfolio. Early tests, for example Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) had to go along with the equally weighted (EW) portfolio for lack of data and/or computational power. Please note that even "broad" commercial indices often do not cover the whole stock market[18] and their market-capitalization weights are often adjusted for free-float[19]. Another aspect is that full coverage, especially including tiny issues in illiquid markets, does not necessarily raise the quality of our data.

**Length of holding period** The CAPM does not tell us over which length of time the investors choose their portfolios. It could be a day, a month, a year or a decade. In the CAPM, there is no rebalancing during the holding period and there are no subsequent investments to be taken (at least they are of no

[15] The Gauss-Markov theorem establishes conditions under which OLS estimators are "optimal" in the sense of being "best linear unbiased estimators" (BLUE).

[16] Robustness of the estimation results in finite-sample can be checked by Monte-Carlo simulations (Kennedy 1998, Section2 2.8 and 2.10 ), so did for example Affleck-Graves and McDonald (1989) and MacKinlay (1985) concerning the test of Gibbons, Ross, and Shanken (1989) and established that the GRS test (described below) is pretty robust to the deviations from the normal which are usually encountered in financial market data. See also MacKinlay and Richardson (1991, FN 3), for improvements.

[17] As long as they are in positive net supply and as such included in the market portfolio, see above.

[18] For instance, MSCI indices cover usually about 60% of total market capitalization.

[19] This is currently pretty popular.

concern). Most tests use monthly data and hence they implicitly assume that one month is the appropriate holding period. We will not go into details here. The point to remember is that the choice of return period is also an assumption which has to be made for empirical testing[20].

From the perspective of theory, please note that considerations of multi-period investments with rebalancings lead to more general pricing models such as the *Intertemporal Capital Asset Pricing Model* (ICAPM) of Merton (1973). A simplified derivation of the ICAPM in discrete time is given by Fama (1996) who closely builds on the traditional CAPM derivation – such as the one presented here.

**Real return proxies** The CAPM is about real prices and real returns. But usually, we observe only nominal prices. What is worse, CPI data is usually available only at low frequencies and with substantial time-lag (and subject to ex-post revisions). Besides, we would need data on expected inflation, not realized inflation. The standard solution to this problem is to take nominal excess returns as proxies for real excess returns[21]. This is valid if we assume that the *nominal* riskfree rate covers also *the*[22] inflation premium. Then we can treat the excess returns as *real* excess returns, even though they are calculated using nominal prices. Hence we conduct our tests on the excess-return formulation of the Sharpe-Lintner-Mossin CAPM using data on excess returns.

Given that $R_f$ is not a constant this affects also second moments of return distributions! Still it is standard to test the CAPM using distributions which are in a (crude) sense "conditional" on the riskfree rate. (See Campbell, Lo, and MacKinlay 1997, p. 182, and Cochrane 2001, Chapter 12, for current practice of empirical tests and Merton 1980 on the sense of conditionality.)

**Equilibrium price data** As already mentioned, the CAPM is a theory of *equilibrium* asset prices. But are the prices we observe in our historical data really equilibrium prices? The answer is that we simply do not know. For instance consider whether high liquidity or low liquidity is an indicator of equilibrium prices[23]? We have to assume that our price data reflects information on equilibrium prices. This is likely to be reasonable over monthly or yearly intervals, where we can expect deviations from equilibrium to be relatively small and to wash out in

[20] At this point you also might want to think about multi-period return distributions of simple returns, horizon effects and how these affect the efficient set.

[21] See also our lecture note on "Return Calculations".

[22] That means there is one standard premium on all nominal assets. Specifically it should not co-vary with asset returns.

[23] At first you might want to go for zero liquidity because that is what equilibrium is all about. But in standard micro, there are just no trades at disequilibrium – remember the meaning of tâtonnement. Suppose there are trades at disequilibrium in reality, still the question is how to distinguish an equilibrium from a disequilibrium price. Do fewer trades mean that it took longer to reach equilibrium or did prices longer stay at equilibrium? You see, this is not trivial.

large samples. With weekly data this should also hold, although illiquid trading might be of concern for small issues. With daily or intra-day data you might want to consider this point more carefully.

### 3.1.2 Economic Assumption

**Efficient Markets** Recall the assumptions we have made so far concerning the return distributions: We stipulated that investors should completely agree in their assessment of the distributions and – for the purpose of testing – we required that returns have stationary distributions so that we can estimate their moments without great avail. One element is missing, though: We need a link between our time-series estimates and the assessment of the investors. That is where the efficient market hypothesis (EMH) comes in. For our purpose, the EMH requires that investor's expectations are rational, given our assumption of stationary distributions it requires that our statistical estimates will match investors' assessment of expected returns and their variance-covariance structure[24].

That link might seem obvious but it is worth stressing it. In general, for any test of an asset pricing model, we always have to make the joint-hypothesis that markets are efficient and that our pricing model holds.

---

**QUIZ:**

- In one of the previous lectures we discussed the Random-Walk model. How does that model fit in with the assumption of stationary distributions? Doesn't it imply that asset prices are non-stationary, that estimates based on historical prices are useless?
- In their description of CAPM tests Elton and Gruber (1995, Chapter 15) assume that the "Market Model holds". Relate their assumption to ours. Are they consistent? Are they identical? If not, which assumption is stronger?
- The CAPM is concerned with the VW market portfolio. Does this imply that testing with an EW (stock) market portfolio is always "inferior" to using its value-weighted counterpart?
- Major index providers, such as MSCI, are changing their index calculations in order to adjust the market weights for "free-float". Does this make them more or less amenable for use as proxies of the market portfolio?
- As a short digression on EW vs. VW, imagine two investors tracking the EW, respectively the VW market portfolio:

[24] We will cover the EMH in greater detail in another session.

  - Are both following passive "do-nothing" strategies?
  - Can you describe their pattern of buying and selling assets (or: increasing and decreasing their asset positions)?
  - Does the choice of EW vs. VW affect the systematic risk of a particular stock? If so, how? Can you sort stocks into groups so that each group is affected identically? (Hint: If there is an effect, identify two groups. One with rising, one with declining systematic risk when switching from EW to VW – or vice versa.)

- How can the riskfree rate be random?
- Can we test the Zero-Beta CAPM of Black (1972) with returns taken in excess of the riskfree rate?

---

## 3.2 Cross-sectional regressions (Fama-MacBeth)

Given that the CAPM imposes a cross-sectional restriction, namely that expected asset returns are linearly related to their market beta, it seems to be straightforward to test the CAPM by running a cross-sectional regression of expected returns on betas. If we want to test specific alternatives, we could just add additional explanatory variables to the regression. In a nutshell, that is already the basic idea of cross-sectional regressions tests. But we do not know expected returns and betas, we can only estimate their values. When running cross-sectional regressions, we also need to account for potential cross-sectional correlation of residual returns. While the assumption of zero serial correlation is generally a good approximation for monthly returns, abnormal returns do co-vary considerably[25] across assets. That is where the classic method of Fama and MacBeth (1973) comes in. As we will see, it does not account for errors in the estimation of betas.

This is not only a historically important method. It is also very intuitive and can be easily extended to time-varying betas or it can make use of adjusted beta estimates (Elton and Gruber 1995, Chapter 7) as discussed in homework 2 of this class. For further readings please consult Campbell, Lo, and MacKinlay (1997, p. 215f.) or Cochrane (2001, Section 12.3). Cochrane embeds the method of Fama and MacBeth as a special case in a wider framework of Generalized Method of Moments (GMM) tests which combine cross-sectional and time-series estimation methods.

### 3.2.1 Average return regressions

Before we come to the method of Fama and MacBeth let us explore our initial idea of regressing expected returns on betas. We do not know their true values, but given

[25] See also Fama (1973) for why Market Model residuals *must* be cross-sectionally correlated.

the assumptions of stationary return distributions we can use estimate of time-series averages, variances and covariances of returns. Estimates are now denoted by a "hat" and we use as inputs for our cross-sectional regression (7):

$$\hat{\mu}_i^e = \frac{1}{T}\sum_{t=1}^{T} r_{i,t} \tag{5}$$

$$\hat{\beta}_{i,M} = \frac{\hat{\sigma}_{i,M}}{\hat{\sigma}_M^2} = \frac{\sum_{t=1}^{T}(r_{i,t} - \hat{\mu}_i^e)(r_{M,t} - \hat{\mu}_M^e)}{\sum_{t=1}^{T}(r_{M,t} - \hat{\mu}_M^e)^2} \tag{6}$$

where the latter is equivalent to the slope of a univariate time-series regression of $r_i$ on $r_M$ ($\tilde{r}_i = \alpha_{i,M} + \beta_{i,M}\tilde{r}_M + \tilde{\varepsilon}$). Alternatively, we could use any other extraneous estimation of beta as mentioned above. Henceforth, we will always assume that betas are estimated from full-sample time-series regressions if not stated otherwise. As our input variables are based on prior (regression) estimations, such regressions are also known as "Two-Pass Regression Tests".

Now suppose the following regression model where we drop the "hat" symbols on $\mu^e$ and $\beta$ now as a matter of convenience:

$$\mu_i^e = \gamma_0 + \gamma_1\beta_{i,M} + \xi_i \tag{7}$$

Please note that the betas enter here as explanatory variables in order to estimate the coefficients $\gamma_0$ and $\gamma_1$. The CAPM restrictions are tested by standard t-tests on the regression coefficients: $\gamma_0 = 0$ and $\gamma_1 > 0$ where we use the coefficients' standard deviation from the regression.

In addition, we want to ensure, that there are no other parameters next to beta which explain the cross-section of expected returns. For instance, we could test whether expected returns are also determined by variance. Particularly it would be interesting whether "idiosyncratic" or residual risk gets priced. Given the variance decomposition of the Market Model (2) $\sigma_i^2 = \sigma_M^2\beta_{i,M}^2 + \sigma^2(\varepsilon_i) = \sigma_{i,M}\beta_{i,M} + \sigma^2(\varepsilon_i)$ we could specify

$$\mu_i^e = \gamma_0 + \gamma_1\beta_{i,M} + \gamma_2\sigma(\varepsilon_i) + \xi_i \tag{8}$$

Alternatively, the relationship between expected returns and beta might be non-linear. So we try

$$\mu_i^e = \gamma_0 + \gamma_1\beta_{i,M} + \gamma_2\beta_{i,M}^2 + \xi_i \tag{9}$$

as a first approximation to potential non-linearities. Obviously, there are not many limits to devising specific alternatives. Combining the above equations we get

$$\mu_i^e = \gamma_0 + \gamma_1\beta_{i,M} + \gamma_2\beta_{i,M}^2 + \gamma_3\sigma(\varepsilon_i) + \xi_i \tag{10}$$

which captures already the idea of the alternative hypothesises tested by Fama and MacBeth (1973) (albeit with a more elaborate methodology as we will see shortly).

Before we come to the contribution of Fama-MacBeth in the next section, let us recap pros and cons of the approach described sofar. The method of Fama and MacBeth retains all of the pros while it tackles some of the cons.

**Pros:**

- We can test the CAPM against specific alternative hypothesises (equations 8 – 10)
- We can input betas obtained from various estimation methods (Adjustments as discussed in Elton and Gruber 1995, Chapter 7, and the like, "fundamental" betas etc.)

**Cons:**

- No account for estimation error in our variables
- Assumes away cross-sectional correlation of residual returns (between the $\tilde{\xi}_i$), this may lead to standard errors which are "off by a factor of 10" (Cochrane 2001, p. 247) (Hint: In your homework the difference will be much smaller.)
- Disregards serial correlation of residuals
- No regard for time-variation in betas and expected returns

The issues mentioned under the cons will affect standard errors of our estimates, which in turn affects the t-stats of our CAPM tests. Unfortunately, the necessary weakening of the underlying assumptions will tend to widen the standard errors, hence standard errors from the regression in (7) will be understated and the t-stats will reject the test hypothesis too often. That is why it is important to account for these issues as good as possible.

---

**QUIZ:**

- Explain how the error terms enter in equation (7). An unexpected component of expected excess return $\mu_i^e$?
- What about testing whether $E[\xi] = 0$?
- Suppose a regression like $\mu_i^e = \gamma_0 + \gamma_1 \kappa_{i,M} + \xi_i$. Depending on what $\kappa_i$ stands for, what can such a regression tell us about the CAPM if ...
  - $\kappa_i =$ Firm size?
  - $\kappa_i = \sigma_i$?
  - $\kappa_i = \sigma_{i,M}$?
- What does "cross-sectional correlation of residual return" actually mean? Is it like a market crash where every stock goes down? Can you give an example?
- What is your estimate of expected (excess) return on asset $i$ in equations (7) – (10)?

---

### 3.2.2 The Fama-MacBeth method

The Fama-MacBeth method allows to incorporate time-varying betas (see the quiz below for time-varying expected returns). When using the same constant, full-sample estimates of the $\beta$'s we get numerically identical estimates for expected $\gamma$'s as in the average return regression. But the standard errors get corrected for cross-sectional correlation of the residuals. With the method of Fama and MacBeth, we run cross-sectional regressions similar to those discussed above but now we use the realized returns for each observed period instead of their time-series averages. For each period $t$, we run a regression like

$$r_{i,t} = \gamma_{0,t} + \gamma_{1,t}\beta_{i,M|t} + \xi_{i,t} \tag{11}$$

over all assets $i$. To clarify our notation, there are N assets ($i = 1 \dots N$) and T observed periods ($t = 1 \dots T$). This means we run $T$ regressions, each one using $N$ observations. We get a time series of estimated regression coefficients $\hat{\gamma}_{j,t}$ where $j = 0, 1$. Please note that we have added a time subscript $t$ to $\beta_{i,M}$ to allow for the possibility to use different values for $\beta_{i,M}$ in each regression.

**Interpreting Fama-MacBeth regression coefficients** Fama (1976, p. 326f.) shows that the coefficients $\gamma_{0,t}$ and $\gamma_{1,t}$ of the cross-sectional regression (11) have a very specific economic content: They are the returns on two portfolios with least-possible variance *while having* the following properties: $\gamma_{0,t}$ is the return in excess of the riskfree rate on a least-variance portfolio whose weights sum to one ("fully invested")[26] and which has a market beta of zero. In addition,$\gamma_{1,t}$ is the return on a zero-investment[27], least-variance portfolio with a beta of one (see Fama and MacBeth 1973, FN 10, and Fama 1976, p. 326f. and p. 334f.). Please note that these portfolios are not necessarily on the (unconstrained) minimum-variance frontier you know from efficient set mathematics. To be precise, these portfolios are constructed as having the least possible variance *while being subject to* additional constraints on their beta and – in multivariate regressions – on other variables like the squared beta. They are only on "the" minimum-variance frontier if this is possible under the additional constraints. Of course, under the CAPM, $\gamma_0$ and $\gamma_1$ are the excess returns on minimum-variance frontier portfolios. Please note that the term "least variance" is non-standard and synonym to "minimum-variance". We use it here only to highlight the use of least-squares regressions as a mean of constructing portfolios. Please note further that we consciously distinguish between a return in excess of the riskfree rate and the return on a zero-investment portfolio.

This assertion can be easily extended to multivariate regressions. For example in

$$r_{i,t} = \gamma_{0,t} + \gamma_1\beta_{i,M|t} + \gamma_{2,t}\sigma(\varepsilon_{i,t}) + \xi_{i,z}$$

[26] The excess return is again the return on a zero-investment portfolio!

[27] Weights sum to zero.

$\gamma_{0,t}$ is the return in excess of the riskfree rate on a fully-invested, least-variance portfolio where the effects of other variables (beta and idiosyncratic volatility $\sigma(\varepsilon_{i,t})$) are zero, $\gamma_{1,t}$ is the return of a zero-investment, least-variance portfolio with a beta of one and zero $\sigma(\varepsilon_{i,t})$ and equivalently $\gamma_{2,t}$ is the return on a zero-investment, least-variance portfolio with a $\sigma(\varepsilon_{i,t})$ of one and zero beta. In general, for

$$\tilde{r}_{i,t} = \gamma_{0,t} + \sum_{j=1}^{J} \gamma_{j,t}\kappa_{i,j|t} + \tilde{\varepsilon}_{i,t}$$

we can say that $\gamma_{0,t}$ is the return in excess of the riskfree rate on a fully-invested, least-variance portfolio where the effects of other variables are zero, and $\gamma_{j,t}$ (for $j > 1$) is the return on a zero-investment, least-variance portfolio where the variable $\kappa_j$ is one and the other $\kappa$'s are zero. The least-variance property follows directly from the Gauss-Markov theorem, in particular the efficiency of OLS estimators (here: $\gamma_{j,t}$, $j \geq 0$). The other constraints are imposed by the regression equations[28].

**Fama-MacBeth Coefficient Tests** From the above it follows that the time series of coefficient estimates $\hat{\gamma}_{0,t}$ and $\hat{\gamma}_{1,t}$ are time series of portfolio returns with particular features: The first being the least-variance return in excess of the riskfree of a fully-invested zero-beta portfolio, the second being the least-variance return of a zero-investment portfolio with a beta of one. If the CAPM holds, the expected return of $\gamma_{0,t}$ should be zero, the expected return of $\gamma_{1,t}$ – which is the risk premium on the market – should be positive. Using time-series averages as estimates of expected values we can test whether these are significantly different from zero with standard t-tests.

In sum, for $j = 0, 1 \ldots$, we use

$$\bar{\gamma}_j = \frac{1}{T}\sum_{t=1}^{T} \hat{\gamma}_{j,t} \tag{12}$$

and

$$\hat{\sigma}(\bar{\gamma}_j) = \sqrt{\frac{1}{T(T-1)}\sum_{t=1}^{T} \left(\hat{\gamma}_{j,t} - \bar{\gamma}_j\right)^2} \tag{13}$$

and form the t-stat

$$\hat{t}(\bar{\gamma}_j) = \frac{\bar{\gamma}_j}{\hat{\sigma}(\bar{\gamma}_j)} \tag{14}$$

That is the Fama-MacBeth procedure!

[28]Hint: In matrix notation the OLS estimator $\vec{\beta}$ of $\vec{y} = \vec{\beta}\boldsymbol{X} + \vec{\varepsilon}$ is defined as $\vec{\beta} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\vec{y}$. Define $\boldsymbol{W} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'$ and you see that the $\vec{\beta}$ are "portfolios" of the $\vec{y}$ with weights defined by $\boldsymbol{W}$. Further analysis of $\boldsymbol{W}$, using $\boldsymbol{W}\boldsymbol{X} = \boldsymbol{I}$, shows that the desired conditions on portfolio weights and portfolio values of the other variables, such as betas, are imposed. See Fama (1976, p. 326f.).

At first, it seems to good to be true. Simple OLS regressions and a t-stat on some time-series averages are all we need. Given the economic interpretation above it looks straightforward, but does it really hold up against more technically-oriented derivations? A detailed and thoughtful account on how the Fama-MacBeth method can be embedded in a more general estimation context is given by Cochrane (2001, Section 12.3) who reaffirms that it is indeed a valid improvement on the "average-return" regression of equation (7). Cochrane shows that under certain conditions (identical and constant estimates of betas) the estimated regression coefficients are numerically the same but depending on the correlation structure of the residuals, the accuracy of standard errors is considerably improved by the Fama-MacBeth method (cross-sectionally correlated residuals).

Of course, we still input estimated beta coefficients and we do not consciously account for their estimation error. Unless we can reasonably expect that these errors are small – that is precisely what Fama and MacBeth (1973) did by estimating the betas of carefully constructed portfolios – this is the strongest drawback of any Two-Pass Regression test. The estimators are biased, but consistent. At least, the necessary adjustments to the estimations' standard errors (while retaining the biased coefficient estimates) can be computed directly (Shanken 1992), otherwise a simultaneous estimation of betas and gammas is at order (Cochrane 2001, Chapter 11f.). See also Kan and Zhang (1999) for a recent discussion of two-pass regressions with an eye on the decrease in power if the pricing model is misspecified. In addition, the Fama-MacBeth assumes that there is no serial correlation in residual returns. If we want to comprehensively tackle serially correlated errors and the errors-in-variables problem, we should indeed use GMM, see Cochrane (2001) for an excellent treatment.

### 3.2.3 Fama-MacBeth Recipe

- Specify a cross-sectional regression model of the type

$$\tilde{r}_{i,t} = \gamma_{0,t} + \gamma_{1,t}\beta_{i,M|t} + \sum_{j=2}^{J} \gamma_{j,t}\kappa_{i,j|t} + \tilde{\varepsilon}_{i,t}$$

  where the additional $J-1$ explanatory variables $\kappa_{i,j|t}$ are optional, for example $\kappa_{i,2} = \beta_{i,M}^2$ as discussed above

- Estimate the inputs $\kappa$. For $\hat{\beta}_{i,M}$ you usually take estimates from market model regressions

- Estimate the cross-sectional regressions for each date in your sample $t = 1 \ldots T$, you get time series $\hat{\gamma}_{j,t}$ with $T$ observations for each coefficient $j = 0, 1 \ldots J$.

  Hint: When you do this in Excel, try `LINEST`[29] (for multivariate regressions – in the univariate case it is easier to work with `SLOPE` and `INTERCEPT`) and make

[29] The German name of this function is `RGP`.

appropriate use of relative references[30]. Depending on how you structure the spreadsheet, `LINEST` might appear clumsy because it returns a matrix. This can be alleviated by nesting `LINEST` in a call of `INDEX`, this allows you to place one single value of the matrix returned by `LINEST` in a single cell – be sure to call `LINEST` with intercept! (Please consult an Excel reference for proper syntax and further details. Be particularly careful with the way `LINEST` orders the regression output in the return matrix!).

- Calculate the t-stats as in equation (14) and test whether

$$\bar{\gamma}_0 = 0$$

$$\bar{\gamma}_1 > 0$$

and for $j > 1$

$$\bar{\gamma}_j = 0$$

- When it comes to evaluating the t-stats it is natural to test all restrictions with two-sided tests, except for $\bar{\gamma}_1 > 0$ where a one-sided test is appropriate.

---

**QUIZ:**

- In the interpretation of the Fama-MacBeth regression coefficients we say that under the CAPM $\gamma_0$ and $\gamma_1$ are the excess returns on frontier portfolios. Which portfolios do we mean?
- What is the difference between a return in excess of the riskfree rate (shorthand: "excess return") and the return on a zero-investment portfolio?
- How would a plot of $\hat{\mu}_i^e$ on $\hat{\beta}_{i,M}$ typically look like? How does the SML look like if it is estimated from a cross-sectional regression ?
- Black, Jensen, and Scholes (1972) find that their estimate of the SML is too flat (see also Elton and Gruber 1995, p. 349f., Table 15.2). What does that mean for the cost-of-capital of a company which has issued a high (low) beta stock and calculates its cost-of-capital with the Sharpe-Lintner-Mossin CAPM?
- In their valuation handbook, Copeland, Koller, and Murrin (2000) advocate to use the 10 year T-Bond instead of the 30 day T-Bill as the riskfree rate when implementing the CAPM to gauge a company's cost-of-capital. (Note that this involves also calculating the market risk premium as the market's expected return in excess of that 10 year T-Bond rate)

[30] This means using the $ – or not ! – in cell references so that you only have to copy the `LINEST` function appropriately in order to repeat the cross-sectional regression for each observation.

- Given the findings of Black, Jensen, and Scholes (1972) quoted above, do you expect that this yields a better fit of expected returns with the SML?
- Copeland, Koller, and Murrin (2000) argue that the T-Bond should be used because its duration (the average period of time the capital is invested) matches the duration of stocks (which is in principle infinite) better than the short maturity T-Bill. Are they right? If so, would their reasoning be consistent with the Sharpe-Lintner-Mossin CAPM or perhaps the Black CAPM? How?

- Based on a true story: Suppose you are advised by your bank, called "Shylock & Co.", which uses the SML with a 30 day T-Bill to estimate stock returns. You are also advised by your management consultant, "MacBeth and Friends" who follows Messrs Copeland, Koller, and Murrin. You are afraid of using a market premium which is too aggressive (too high) and now you listen to both of them. Whose advice would look more sound to you?

- In the Fama-MacBeth regression (11), at which levels of $R^2$ would you reject the CAPM?
  - What answer would you give with regard to equation (7)?
  - What would you think if the $R^2$ of (15) were higher than in (11)? (Hint: You need not worry about adjusted $R^2$ here[31], the same issue applies when comparing $\bar{R}^2$.)

- Suppose you estimate the original model of Fama and MacBeth (1973, their equation 7)

$$r_{i,t} = \gamma_{0,t} + \gamma_1 \beta_{i,M|t} + \gamma_{2,t} \sigma(\varepsilon_{i,t}) + \xi_{i,z} \tag{15}$$

  adapted to our excess-return setup and you find that $E[\gamma_1] > 0$, $E[\gamma_2] = 0$, $E[\gamma_3]$ but $E[\gamma_0] \neq 0$. What does that tell us about the Sharpe-Lintner-Mossin CAPM? What about the Black (Zero-Beta) CAPM?

- Could you incorporate time-varying expected returns in the Fama-MacBeth procedure? For example suppose the market premium varies with the term spread[32] $\text{TERM}_t$ like in $\lambda_{M,t} = \omega\ \text{TERM}_t + \xi_t$

- From the quiz above: What is your estimate of expected excess return on asset $i$ in the alternative equations (7) – (10)? What is your estimate of expected excess return on $i$ from the Fama-MacBeth estimations of (11) and (15)?

---

[31] Adjusted $R^2$ accounts for differences in the degrees of freedom between regressions where the number of explanatory variables differ, see for example Greene (2000, p. 240).

[32] That is the difference in interest rates between a long and a short bond – please note that "long" and "short" are to be understood in terms of maturity here.

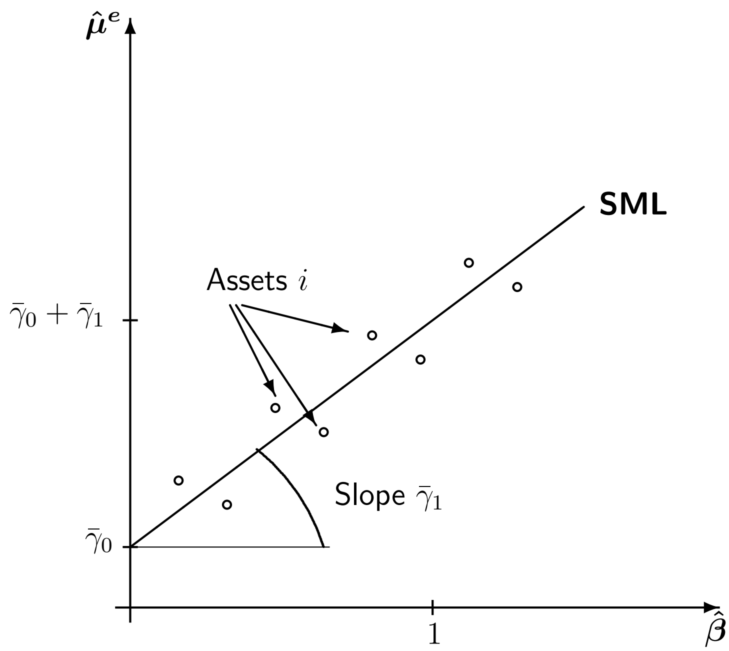


Figure 2: Security Market Line estimated from Cross-Sectional Regressions

- When testing whether $\bar{\gamma}_1 > 0$, do you reject the CAPM for values above or below the critical t-value? Would a two-sided test (instead of the appropriate one-sided test) err on the side of conservatism or not (reject too few or too often)?

---

### 3.2.4 SML estimated from Cross-Sectional Regressions

In cross-sectional regressions we fit the SML on the cross-section of expected excess returns $\mu_i^e$ and $\beta$'s. Hence we estimate the slope $\bar{\gamma}_1$ and intercept $\bar{\gamma}_0$ by minimizing the pricing errors of all assets in a least-squares sense. The estimated slope of the SML, $\bar{\gamma}_1$, is our estimate of the market risk premium $\lambda_M$. Figure 2 shows how a plot of expected excess returns, betas and the SML would look like when using estimates from cross-sectional regressions[33].

Please note that we get the same SML using either an average return regression or the Fama-MacBeth method when we use the same, constant estimates of the $\beta$'s.

[33] You will notice that the expected excess return on a portfolio with $\beta = 1$ equals $\bar{\gamma}_0 + \bar{\gamma}_1$. This does not contradict the interpretation of the Fama-MacBeth coefficients: $\bar{\gamma}_1$ is the expected return of a zero-investment portfolio with $\beta = 1$, which need not be the expected return on the market in excess of the riskfree rate! For instance, if the Sharpe-Lintner-Mossin CAPM with unrestricted borrowing and lending does not hold, but some variant of the Black CAPM (Elton and Gruber 1995, Chapter 14) where the expected return on the market's zero-beta portfolio is higher than the riskfree rate!

## 3.3 Time-series regressions (GRS)

### 3.3.1 Cross-sectional restrictions, again

We have already stressed above that the CAPM implies a cross-sectional restriction. In particular, the SML states that the *cross-sectional* variation in expected (excess) returns $\mu_i^e$ is fully explained by a linear relationship with $\beta_{i,M}$. Formalizing this restriction with a cross-sectional regression model is straightforward, in the previous section we spent most of our attention on the econometric implementation of the SML. But how do we express the CAPM restriction in a time-series regression?

Our candidate model for time-series regressions is the Market Model regression (2). It contains beta and the return on the market which looks already pretty close to the CAPM (equation 3). Let us take expectations of (2) and write

$$E\left[\tilde{R}_{i,t}\right] = a_{i,t} + \beta_{i,M} E\left[\tilde{R}_{M,t}\right],\ \forall\, i = 1 \dots N \tag{16}$$

Again, please note that we have consciously chosen to denote the regression intercept $a$ and not (yet) $\alpha$. Equivalently we can rewrite the CAPM (equation 3) to get something similar to (16)

$$E\left[\tilde{R}_{i,t}\right] = R_f\left(1 - \beta_{i,M}\right) + \beta_{i,M} E\left[\tilde{R}_{M,t}\right],\ \forall\, i = 1 \dots N \tag{17}$$

Obviously, the CAPM implies $a_{i,t} = R_f\left(1 - \beta_{i,M}\right)$, it constrains the intercepts in the Market Model.

In excess returns, the CAPM (equation 4) implies that the intercepts $\alpha_i$ in time-series regressions for each asset $i$

$$\tilde{r}_{i,t} = \alpha_i + \beta_{i,M}\tilde{r}_{M,t} + \tilde{\varepsilon}_{i,t},\ \forall\, i = 1 \dots N \tag{18}$$

are all zero. We call $\alpha_i$ the *pricing error* of asset $i$, that means $\alpha_i$ is the difference in the expected return of asset $i$ estimated from its time-series average with what is predicted by the CAPM[34].

The cross-sectional character of this restriction is that *all* $\alpha_i$ need to be zero, not just some of them and not only the sum $\sum_i \alpha_i$ should be zero (which implies that the average intercept is zero). In empirical tests, we can only use estimates $\hat{\alpha}_i$ and they will of course be different from zero. In fact, the distribution theory of the univariate t-tests from the regressions ($t(\hat{\alpha}_i) = 0$) tells us that a certain percentage of our estimators will even be significantly different from zero in repeated samples (under the CAPM's Null Hypothesis). Looking at univariate t-stats alone, does not do the trick. We need a joint test whether all the $\alpha_i = 0$ which accounts also for correlation in the estimation errors of $\alpha$'s. That is where the GRS test comes in.

[34] Using the time-series average of $r_M$ as estimate of the market risk premium.

**QUIZ:**

- The CAPM implies zero pricing errors $\alpha_i = 0$ in (18).
  - What is the difference between equation (18) and the Market Model?
  - Does the CAPM imply that intercepts in equation (16) are equal?
  - How many free parameters does the CAPM leave in equation (16)?
- What are the pricing errors in the cross-sectional regressions such as (11)?
- How do the pricing errors influence the estimation of the cross-sectional regressions (like 11)? How do they affect the estimation of the time-series regressions (18)?
- How would a plot of $\hat{\mu}_i^e$ on $\hat{\beta}_{i,M}$ typically look like? How does the SML look like if it is estimated from a time-series regression?
- Say, we run regressions like (18) on 20 assets and find that 2 of the $\alpha_i$ are statistically different from zero at the 10% level using univariate regression t-stats. Do we reject the CAPM?
- At which levels of $R^2$ in a time-series regression $\tilde{r}_i = \alpha_i + \beta_{i,M}\tilde{r}_M + \tilde{\varepsilon}$ would you reject the CAPM? Why (not)?
- What is the estimator of the expected market premium $\lambda_M$ in ...
  - a time-series regression as described above?
  - the Fama-MacBeth method?

  Are they consistent?
- What is the alternative hypothesis in a time-series test? Can you imagine a multivariate counterpart to the cross-sectional regression of (15)?
  (Hint: Compare the interpretation of the explanatory variables in a cross-sectional and in a time-series regression. Can we really answer the second question – based on what we have discussed so far?)

---

### 3.3.2 The GRS Test

The contribution of Gibbons, Ross, and Shanken (1989) is twofold:

1. A *finite sample* test on whether the estimated $\alpha$'s ($\hat{\alpha}_i$) in (18) are *jointly* zero, we stack the $\hat{\alpha}_i$ in a vector $\vec{\alpha}$ and write in vector notation $H_0: \ \vec{\alpha} = \vec{0}$
2. They show that $\vec{\alpha} = \vec{0}$ is *equivalent to* testing differences in (ex-post) Sharpe Ratios or shifts in the efficient set

**Econometric Perspective** From Maximum-Likelihood theory, tests on joint coefficient values are known as Wald, Lagrange Multiplier (LM) or Likelihood Ratio (LRT) tests (Davidson and MacKinnon 1993, page 243). Often you find such tests as pre-packaged functions in econometric software like EViews. The problem is that their distribution is usually only valid asymptotically[35].

From OLS theory you might know the F-Test which can be applied to multiple coefficient restrictions (remember that the F-test is equivalent to the square of the t-test if the restriction is imposed that a single coefficient is zero). Under the assumption of normally distributed residuals, the F-Test is valid in finite samples. Gibbons, Ross, and Shanken derive an F-statistic on $\vec{\alpha} = \vec{0}$ *across the regressions for each asset $i$ in (18)* which is valid in finite samples (under normally distributed returns). It can be shown that their test is robust under common deviations from normality found in financial data (Affleck-Graves and McDonald 1989, MacKinlay 1985, see also MacKinlay and Richardson 1991, FN 3, for improvements.).

Britten-Jones (1999) shows how to back out the GRS Test from a *single* OLS regression. His paper is particularly insightful because it implies a distribution theory on the *weights* of efficient portfolios. We will not follow that lead here but suggest his paper for further reading on the link between least-squares regressions, projections and mean-variance theory.

**Economic Intuition** In their paper, Gibbons, Ross, and Shanken (1989) show that in the Sharpe-Lintner-Mossin CAPM the following problems are equivalent:

1. Are the intercepts $\alpha_i$ in the $N$ regressions (18) *jointly* zero?

2. Is the market portfolio ("test factor") efficient? Does the market portfolio "span" the efficient set?

3. Are two Sharpe Ratios significantly different? In the context of the CAPM: Does the market portfolio have the maximum Sharpe Ratio? In particular, is the Sharpe Ratio of the market portfolio significantly *smaller* than the Sharpe Ratio of the efficient combination of the market with the test assets $i = 1 \dots N$?

You can confirm the intuition behind these equivalences by looking at our results from mean-variance mathematics.

Before we present a recipe for implementing the GRS test, let us review shortly its derivation, in particular with regard to the last assertion from above, namely that the GRS test is equivalent to a comparison of two Sharpe-Ratios.

Similar to a univariate F-test (which is equal to a squared t-test), its multivariate analogue looks at the ratio between the square of the estimated coefficients and their

[35] In the context of CAPM tests see the Wald statistic $J_0$ of Campbell, Lo, and MacKinlay 1997, p. 191, the GRS test described below corresponds to their $J_1$.

variance, in matrix notation this involves the quadratic expression

$$\vec{\alpha}' \left[\text{Var}[\vec{\alpha}]\right]^{-1} \vec{\alpha} = T \frac{\vec{\alpha}' \mathbf{\Sigma}^{-1} \vec{\alpha}}{1 + \left(\frac{\hat{\mu}_M^e}{\hat{\sigma}_M}\right)^2}$$

where $\mathbf{\Sigma}$ is the variance-covariance matrix of the residuals in (18). By the way, this is the asymptotic Wald test on $\vec{\alpha} = \vec{0}$. It has a chi-square distribution with $N$ degrees of freedom (Campbell, Lo, and MacKinlay 1997, p. 192).

The first contribution of GRS is the finite sample validity of their test, which is achieved by an appropriate adjustment of the above expression for degrees of freedom (see our recipe). Here we will focus on their second contribution, how to re-write the expression in terms of Sharpe Ratios. GRS show that for the sample estimates (see the recipe for how to calculate them) we can write

$$\vec{\alpha}' \hat{\mathbf{\Sigma}} \vec{\alpha} = \left(\frac{\hat{\mu}_q^e}{\hat{\sigma}_q}\right)^2 - \left(\frac{\hat{\mu}_M^e}{\hat{\sigma}_M}\right)^2$$

where the portfolio $q$ is the ex-post tangency portfolio[36] constructed from the $N$ assets plus the test factor $M$. The GRS test statistic $W_N$ (described in the recipe) involves just a re-scaling of this expression. Its content can be interpreted in terms of the difference in squared Sharpe Ratios between $q$ and $M$. For ease of notation, we let $q$ denote the log portfolio, whose Sharpe Ratio is identical to the tangency portfolio (see below for details).

If $M$ is *ex-post* efficient, it has the highest Sharpe Ratio of all portfolios which can be constructed and it will have the same Sharpe Ratio as $q$ (where $M$ is included!). Otherwise, $q$ can only have a higher Sharpe Ratio (because together with the $N$ test assets we can only do better than with $M$ alone). Suppose $M$ is ex-ante efficient. Unfortunately, it need not be ex-post efficient in this case. By luck, it might happen that some of the $N$ test assets in $q$ do better than expected (and than $M$ does) and $q$ will have a higher *realized* Sharpe Ratio of than $M$ – *ex-post*. But $q$ should not be much more efficient than $M$, at least not if $M$ has the highest *expected* Sharpe Ratio[37]. The big question is how large can the difference in realized, squared Sharpe Ratios get, while we can still attribute this to luck and ex-post sampling error? At what point should we seize to believe in the efficiency of $M$? That is the question the GRS F-statistic $W_N$ answers.

---

**QUIZ:**

[36] That means the weights in $q$ are proportional to $\boldsymbol{V}^{-1}\vec{\mu}^e$, where $\boldsymbol{V}$ is the variance-covariance matrix of (excess) returns and $\vec{\mu}^e$ is the vector of mean excess returns. Note that they include the $N$ test assets as well as the test factor $M$.

[37] Note, that this is equivalent to the presumed ex-ante efficiency of $M$.

- Why can the market portfolio's Sharpe Ratio only be smaller (or equal) than the Sharpe Ratio of the (ex-post) efficient set obtained from the market portfolio and the test assets?

- Bear in mind that the GRS test is applied to the Sharpe-Lintner-Mossin CAPM: What particular shape has the efficient set of all assets? Can you now explain the equivalence in Problem 2. and 3. above?

- In the paper of Gibbons, Ross, and Shanken (1989, p. 1141) it is written

  > The hypothesis that $\alpha_{ip} = 0 \; \forall \, i$ is violated if and only if some linear combination of the $\alpha$'s is zero ...

  What do you think?
  (Hint: $\alpha$'s are excess returns on zero-beta portfolios. A linear combination of $\alpha$'s is the excess return on a zero-beta portfolio of the test assets. Hence the sentence above says that the CAPM is violated if some zero-beta portfolio of the test assets has an excess return of zero.)

---

### 3.3.3 GRS Recipe

- You have $N$ test assets and one test factor (your proxy of the market portfolio, denoted $M$) with $T$ observations of realized excess returns

- Run time-series regressions of each test asset's excess returns on the test factor

  $$r_{i,t} = \hat{\alpha}_i + \hat{\beta}_{i,M} r_{M,t} + \hat{\varepsilon}_{i,t}$$

  Stack the estimated intercepts $\hat{\alpha}_i$ in the vector $\vec{\alpha}$

- For the test factor, calculate sample mean $\hat{\mu}_M^e$ and variance $\hat{\sigma}_M^2$. Do not adjust the sample variance for degrees of freedom, but use the maximum-likelihood estimate:

  $$\hat{\mu}_M^e = \frac{1}{T} \sum_{t=1}^{T} r_{M,t}$$

  $$\hat{\sigma}_M^2 = \frac{1}{T} \sum_{t=1}^{T} \left( r_{M,t} - \hat{\mu}_M^e \right)^2$$

- As discussed above, there are two ways of constructing the GRS Test making use of

  $$\vec{\alpha}' \hat{\mathbf{\Sigma}} \vec{\alpha} = \left( \frac{\hat{\mu}_q^e}{\hat{\sigma}_q} \right)^2 - \left( \frac{\hat{\mu}_M^e}{\hat{\sigma}_M} \right)^2$$

  Now you can choose to compute either of the following:

– Estimate the variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ of the residuals $\hat{\varepsilon}_{i,t}$. Gibbons, Ross, and Shanken (1989, p. 1124) stress that only an unbiased estimate should be used[38]. Following Cochrane (2001, p. 230f.)[39] and Campbell, Lo, and MacKinlay (1997, p. 191f) use the (biased) maximum-likelihood estimate instead. Hence, estimate the covariance between assets $i$ and $j$ as[40]

$$\hat{\sigma}_{i,j} = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{i,t}\hat{\varepsilon}_{j,t}\ \forall\, i, j$$

The necessary adjustment in the degrees of freedom will be done when calculating the statistic $W_N$ below.

In Excel you can simply take the results of the function `COVAR`. (Hint: To construct the full covariance matrix in Excel with `COVAR` you should make judicious use of relative cell-references and making copies of your initial cell. You cannot create the whole matrix in one step, but at least you can create each line by making copies of the first element, using relative references.)

Be careful when using the covariance function of the Analysis ToolPak in the Options menu! First of all, it scales the covariation by $\frac{1}{T-1}$[41] which gives an unbiased estimate between two series of given data, but not for our regression residuals[42]. Secondly, the ToolPak yields only the lower triangular covariance matrix. You still need to fill up the upper triangular by yourself.

– Alternatively, we could construct the ex-post tangency portfolio $q$ and calculate its Sharpe Ratio. For ease of computation, we will look at another efficient portfolio, known as the "log-portfolio" instead of the tangency portfolio. Both have the same Sharpe Ratio, but the log-portfolio is easier to calculate[43]. We can (but need not, see below) calculate the weights of $q$ as

$$\vec{w}_q = \hat{\boldsymbol{V}}^{-1}\vec{\mu}$$

Where $\hat{\boldsymbol{V}}$ is the variance-covariance matrix and $\vec{\mu}$ is the vector of estimated

[38] That requires to scale the co-variations by $\frac{1}{T-2}$ in this case. We use $T-2$ and not the usual $T-1$ because our computations are based on the residuals from a regression where two parameters had to be estimated in advance.

[39] Cochrane does actually not specify which estimator of $\boldsymbol{\Sigma}$ to use.

[40] For $i = j$ we get the variance, $\hat{\sigma}_{i,i} = \hat{\sigma}_i^2$.

[41] The German Excel 97 help says it scales by $\frac{1}{T}$ but it does use a scaling of $\frac{1}{T-1}$ instead.

[42] In a univariate regression you estimate two parameters! Now you would have to re-scale by $\frac{T-1}{T}$.

[43] The log-portfolio is the portfolio preferred by an investor with logarithmic utility, i.e. he has unit relative risk aversion. Recall that the weights of an efficient portfolio can be computed as $\eta\hat{\boldsymbol{V}}^{-1}\vec{\mu}$ where $\eta$ is the coefficient of relative risk aversion.

means of the $N$ excess returns *plus* the test factor[44].

Calculate $\hat{\boldsymbol{V}}$ using the COVAR function as discussed above for $\hat{\boldsymbol{\Sigma}}$. In Excel you will also need the functions MINVERSE, TRANSPOSE and MMULT. Actually, we might want to scale the weights so that they sum to one[45] $\vec{w}_q^* = \frac{1}{\mathbf{1}'\vec{w}_q}\vec{w}_q$ (Campbell, Lo, and MacKinlay 1997, p. 198). But please note, that this does not affect the Sharpe Ratio of $q$.

In order to get the Sharpe Ratio $\frac{\hat{\mu}_q^e}{\hat{\sigma}_q}$ you need not even construct the time series of returns of $q$ and estimate its moments[46] but make use of $\vec{\mu}$ and $\hat{\boldsymbol{V}}$ in

$$\frac{\hat{\mu}_q^e}{\hat{\sigma}_q} = \frac{\vec{w}_q'\vec{\mu}}{\sqrt{\vec{w}_q'\hat{\boldsymbol{V}}\vec{w}_q}} = \frac{\vec{\mu}'\hat{\boldsymbol{V}}^{-1}\vec{\mu}}{\sqrt{\vec{\mu}'\hat{\boldsymbol{V}}^{-1}\hat{\boldsymbol{V}}\hat{\boldsymbol{V}}^{-1}\vec{\mu}}} = \sqrt{\vec{\mu}'\hat{\boldsymbol{V}}^{-1}\vec{\mu}}$$

As the last transformation shows, we need not even calculate the weights $\vec{w}_q$ in order to compute the Sharpe Ratio[47] of $q$

- Calculate the W-statistic

$$W = \frac{\vec{\alpha}'\hat{\boldsymbol{\Sigma}}^{-1}\vec{\alpha}}{1 + \left(\frac{\hat{\mu}_M^e}{\hat{\sigma}_M}\right)^2} = \frac{\left(\frac{\hat{\mu}_q^e}{\hat{\sigma}_q}\right)^2 - \left(\frac{\hat{\mu}_M^e}{\hat{\sigma}_M}\right)^2}{1 + \left(\frac{\hat{\mu}_M^e}{\hat{\sigma}_M}\right)^2}$$

  Again, in Excel you will need the matrix functions MINVERSE, TRANSPOSE and MMULT

- In order to get the F-statistic, normalize[48] $W$

$$W_N = \frac{T-N-1}{N}\, W \sim F(N)(T-N-1)$$

[44] Hence their dimensions are $(N+1) \times (N+1)$ respectively $(N+1) \times 1$.

[45] Remember what the "tangency" portfolio is about in the CAPM!

[46] If you should do so, please use the same estimators as for the test factor. In particular use the maximum-likelihood estimator for the variance $\hat{\sigma}_q^2 = \frac{1}{T}\sum_{t=1}^{T}\left(r_{q,t} - \hat{\mu}_q^e\right)^2$.

[47] The fact that we can compute the square of an optimal Sharpe Ratio by generalizing from the square of scalars $\frac{\mu^2}{\sigma^2}$ to the quadratic form $\vec{\mu}'\boldsymbol{V}^{-1}\vec{\mu}$ offers also a new perspective on the quadratic form involving the regression $\alpha$'s, $\vec{\alpha}'\boldsymbol{\Sigma}^{-1}\vec{\alpha}$. We can interpret it as a squared Sharpe Ratio, too. It is the Sharpe Ratio of an optimal portfolio formed by combining $N$ portfolios obtained from hedging each test asset against the market: $\tilde{r}_i - \beta_{i,M}\tilde{r}_M = \alpha_i + \tilde{\varepsilon}_i$. These portfolios have means equal to $\alpha_i$ and (co-)variances corresponding to $\sigma(\tilde{\varepsilon}_i, \tilde{\varepsilon}_j)$ which are captured by $\vec{\alpha}$ and $\boldsymbol{\Sigma}$. If the Sharpe-Lintner CAPM holds, such market-hedged portfolios cannot have non-zero expected returns and their optimal Sharpe Ratio should not be significantly different from zero. Now we see that Gibbons, Ross, and Shanken (1989) have shown, that asking whether an optimal portfolio of the *hedged* test assets has a (squared) Sharpe Ratio of zero is equivalent to asking whether an optimal portfolio of the $N$ test assets *plus* the test factor has a squared Sharpe Ratio which is only insignificantly greater than the Sharpe Ratio of the test factor alone.

[48] If we had used the unbiased estimate of $\hat{\boldsymbol{\Sigma}}$ we should now use the original normalization done by Gibbons, Ross, and Shanken (1989), $W_N = \frac{T}{N}\,\frac{T-N-1}{T-2}\,W \sim F(N)(T-N-1)$, yielding the same result.

Under the Null-Hypothesis ($\vec{\alpha} = \vec{0}$), $W_N = 0$

- To find the percentile of the F-distribution associated with $W_N$ you need to call the function `FDIST(W_N, N, T-N-1)`. Low p-values reject, high p-values fail to reject

---

**QUIZ:**

- When calculating the covariance matrix $\hat{\boldsymbol{\Sigma}}$, why don't we account for the mean of the *estimated* residuals?
- Why are Sharpe Ratios of tangency portfolio and log-portfolio equal?
- The above recipe for the GRS test contains formulas for mean, variance and Sharpe Ratio of the (ex-post) log-portfolio $q$: CHECK

$$\hat{\mu}_q^e = \vec{\mu}'\hat{\boldsymbol{V}}^{-1}\vec{\mu}$$
$$\hat{\sigma}_q^2 = \vec{\mu}'\hat{\boldsymbol{V}}^{-1}\vec{\mu}$$
$$\frac{\hat{\mu}_q^e}{\hat{\sigma}_q} = \sqrt{\vec{\mu}'\hat{\boldsymbol{V}}^{-1}\vec{\mu}}$$

  Verify the following statements, relate them to our results from mean-variance mathematics and try to give also an intuitive explanation:

  – The expected return on $q$ (and hence also the Sharpe Ratio) is positive. (Hint: $\boldsymbol{V}$ is positive definite.)

  – The expected return on $q$ equals its variance. Hence the efficient set in $(\mu, \sigma^2)$ space is the 45 degrees line.

---

### 3.3.4 SML estimated from Time-Series Regressions

In time-series regressions, we actually do not estimate the SML in that we fit its intercept and slope on the cross-section of expected excess returns and $\beta$'s. Rather, the SML is given with a slope equal to the time-series average of our market proxy's excess returns ($\hat{\mu}_M^e$) and an intercept of zero (Sharpe-Lintner-Mossin CAPM). $\hat{\mu}_M^e$ is our estimate of the market risk premium.

Hence we disregard pricing errors when we construct the SML in a time-series estimation. Figure 3 shows how a plot of expected excess returns, betas and the SML would look like when using estimates from time-series regressions. Depending on the set of test-assets and the market proxy, it may even happen that we have only positive (or negative) pricing errors as depicted in figure 3 (suggesting a rather

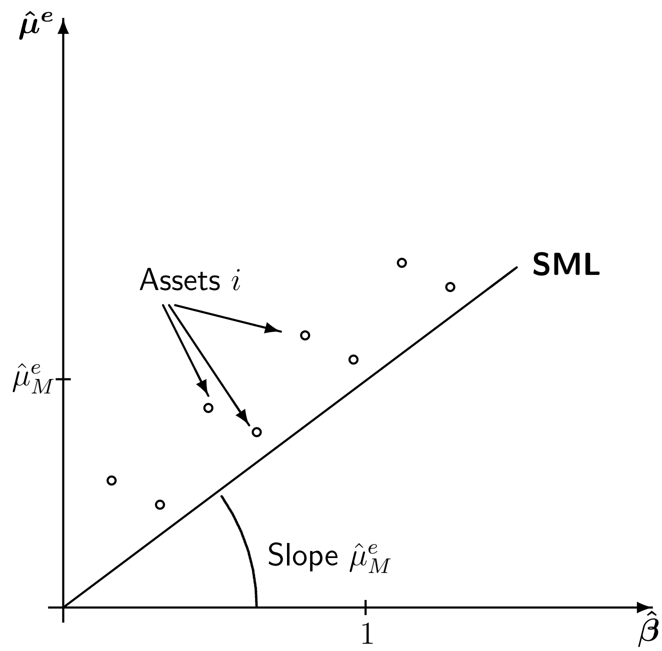


Figure 3: Security Market Line estimated from Time-Series Regressions

unfortunate construction of test assets and market proxy). Instead of seeking for a "best-fit" SML, we rather ask ourselves whether the pricing errors $\alpha_i$ are still in line with the SML – "given" by $\hat{\mu}^e_M$ as discussed above – and the correlation structure of test assets and market proxy (see the interpretation of the GRS test).

# A Notation

| | |
|---|---|
| $\tilde{R}$ | Random rate of return |
| $R_{i,t}$ | Return on asset / portfolio $i$ in period $t$ |
| $R_{E,t}$ | . . . on efficient portfolio $E$ . . . |
| $R_{M,t}$ | . . . on the market portfolio . . . |
| $R_f$ | Constant, riskfree rate of return |
| $r_{i,t}$ | Excess return: $\tilde{r}_{i,t} \equiv \tilde{R}_{i,t} - R_f$ |
| $\alpha_{i,P}$ | Intercept of (univariate) regression of $i$ on $P$ |
| $\beta_{i,P}$ | Slope of (univariate) regression of $i$ on $P$ |
| $\mu^e_i$ | Mean of excess return on $i$. $\mu^e_i \equiv E[r_i]$ (True value! No estimate.) |
| $\sigma^2_i$ | Variance of return on $i$ |
| $\sigma_{i,j}$ | covariance between returns on $i$ and $j$ |
| $\hat{\gamma}$ | Estimate of parameter $\gamma$ |

## List of Figures


1 Security Market Line . . . . . . . . . . . . . . . . . . . . . . . . . 10
2 Security Market Line estimated from Cross-Sectional Regressions . . 23
3 Security Market Line estimated from Time-Series Regressions . . . . 32


## References


Affleck-Graves, John, and Bill McDonald. 1989. "Nonnormalities and Tests of Asset Pricing Theories." *The Journal of Finance* 44 (4): 889 – 908 (September).

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *The Journal of Business* 45 (3): 444–455 (July).

Black, Fischer, M. Jensen, and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*, edited by M. Jensen, 79–121. New York: Praeger.

Britten-Jones, Mark. 1999. "The Sampling Error in Estimates of Mean Variance Efficient Portfolio Weights." *The Journal of Finance* 54 (2): 655–671 (April).

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Cochrane, John H. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Copeland, Tom, Tim Koller, and Jack Murrin. 2000. *Valuation – Measuring and Managing the Value of Companies*. 3rd edition. Wiley Finance. New York, NY: John Wiley & Sons, Inc.

Davidson, Russell, and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York, NY: Oxford University Press. 6th printing.

Elton, Edwin J., and Martin J. Gruber. 1995. *Modern Portfolio Theory and Investment Analysis*. 5th edition. New York, NY: John Wiley & Sons, Inc.

Fama, Eugene F. 1973. "A Note on the Market Model and the Two-Parameter Model." *The Journal of Finance* 28 (5): 1181 – 1185 (December).

———. 1976. *Foundations of Finance*. Basic Books.

———. 1996. "Multifactor Portfolio Efficiency and Multifactor Asset Pricing." *The Journal of Financial and Quantitative Analysis* 31:441–465.

Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *The Journal of Political Economy* 81 (3): 607–636 (May – June).

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57 (5): 1121–1152 (September).

Greene, William H. 2000. *Econometric Analysis*. 4th edition. Upper Saddle River, NJ: Prentice Hall.

Ingersoll, Jonathan E., Jr. 1987. *Theory of Financial Decision Making*. Studies in Financial Economics. Savage, MD: Roman & Littlefield.

Jagannathan, Ravi, and Zhenyu Wang. 1996. "The Conditional CAPM and the Cross-Section of Expected Returns." *The Journal of Finance* 52 (1): 3 – 53 (March).

Kan, Raymond, and Chu Zhang. 1999. "Two-Pass Tests of Asset Pricing Models with Useless Factors." *The Journal of Finance* 54 (1): 203 – 235 (February).

Kennedy, Peter. 1998. *A Guide to Econometrics*. 4th edition. Cambridge, MA: The MIT Press.

MacKinlay, A. Craig. 1985. "An analysis of multivariate financial tests." Unpublished dissertation, University of Chicago.

MacKinlay, A. Craig, and Matthew P. Richardson. 1991. "Using Generalized Method of Moments to Test Mean-Variance Efficiency." *The Journal of Finance* 46 (2): 511 – 57 (June).

Magill, Michael, and Martine Quinzii. 1998. *Theory of Incomplete Markets*. Volume 1. Cambridge, MA: The MIT Press.

Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5): 867–887 (September).

———. 1980. "On Estimating the Expected Return on the Market." *The Journal of Financial Economics* 8:323 – 361.

Owen, Joel, and Ramon Rabinovitch. 1983. "On the Class of Elliptical Distributions and their Application to the Theory of Portfolio Choice." *The Journal of Finance* 38 (3): 745 – 752 (June).

Roll, Richard. 1977. "A Critique of the Asset Pricing Theory's Tests, Part I: On Past and Potential Testability of the Theory." *The Journal of Financial Economics* 4 (2): 129–176 (March).

Shanken, Jay. 1992. "On the Estimation of Beta-Pricing Models." *The Review of Financial Studies* 5 (1): 1 – 33.

Stambaugh, Robert F. 1982. "On the exclusion of assets from tests of the two parameter model." *The Journal of Financial Economics* 10:235–268.