# Two Essays on General Equilibrium Foundation of Finance

Inaugural Dissertation
for a Doctoral Degree
of the Faculty of Economics
at the University of Bielefeld

submitted by

Thomas Siwik

Bielefeld, February 2002

Referee: Prof. Dr. Thorsten Hens
Institut für Empirische Wirtschaftsforschung
Universität Zürich

Co–Referee: Prof. Dr. Thomas Braun
Fakultät für Wirtschaftswissenschaften
Universität Bielefeld

With kind permission of the
Department of Economics and Business Administration,
University of Bielefeld
published by
Thomas Siwik, Cologne, Febuary 2002.
Contact by electronic mail: dissts@dcts.de.



Typeset with LaTeX and BibTeX.
Printed on permanent paper according to DIN–ISO 9706 and bound with an adhesive book binding by Deloitte & Touche, Düsseldorf.
Electronic copies in secured PDF–file format burned on CD–ROM according to ISO 9660 are available for downloading at the University of Bielefeld: http://archive.ub.uni-bielefeld.de/disshabi/wiwi.htm.
Hard and electronic copies coincide entirely in contents.

## Acknowledgments

This thesis definitely benefitted from the opportunities the Faculty of Economics at University of Bielefeld provided to me. The Faculty gave me free access to many resources, enabled me to join conferences and workshops, and to take part in their suspenseful economic research program, allowed me to work on a three month research project in a bank, and let me teach students in finance and economics.

My supervisor Thorsten Hens is the person who made those valuable opportunities actually possible for me. He also encouraged me to work in the field of "General Equilibrium Foundation of Finance". Especially he draws my attention to the problem of monotonicity in the CAPM. Throughout my arising thesis he gave very helpful hints and remarks. I am deeply indebted to Thorsten.

I am grateful to Andreas Löffler, with whom I had some fruitful discussions about Siegel's Paradox, which induced me to write about foreign exchange rates. Fortunately, a London bank fixed me up with the data for the empirical investigation. Moreover I would like to thank Dirk Ebmeyer for some helpful comments on this chapter.

For discussions in seminars and private talks about my arising thesis I feel much obliged to my second supervisor, Thomas Braun, as well to other professors, teachers, and former colleagues at the University of Bielefeld: Pu Chen, Dirk Ebmeyer, Peter Flaschel, Joachim Frohn, Anke Gerber, Nicole Gerling, Gangolf Groh, Sandra Güth, Klaus–Jochen Haake, Andreas Handel, Eckart Jäger, Carsten Köper, Sven Ludwig, Riza Öztürk, Andreas Park, Beate Pilgrim, Jan Pilgrim, Mathias Raith, Wolfgang Rohde, Klaus Reiner Schenk–Hoppé, Willi Semmler, Axel Weber, Jan Wenzelburger, Peter Wöhrmann, Christoph Wöster, and other participants of seminars.

Last but not least I would like to express my gratitude to Bruna Heavern for her proofreading of the entire document.

# Contents

# List of Figures

# 1 Preface

My thesis consists of two essays: *The CAPM – A General Equilibrium Foundation* and *The Foreign Exchange Rate in Financial Markets*. These essays are about two themes of financial markets: Optimal portfolio selection and arbitrage–free pricing. Portfolio selection and arbitrage pricing theory are the most prominent subjects in Finance and important issues in General Equilibrium Theory as well. This is due to the steadily rising importance of financial markets for consumers, firms, and states to invest their wealth, to raise capital, and to hedge their risks.

At a first glance both essays are thematically independent. But their common foundation is a General Equilibrium point of view. An important aspect of both essays is a General Equilibrium model sustaining the investigations. General Equilibrium models distinguish themselves in that all results are derived from rational individuals' interactive behaviour.

## 1.1 The Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is one of the most successful models in financial markets with regard to portfolio selection. The CAPM utility functions are assumed to depend positively on the mean and negatively on the variance of returns from an investment. This assumption allows for explicit solutions of equilibria in only one quantity, the harmonic mean of investors' absolute risk aversions. However, the actual value for practical applications lies in the very simple risk measure, the variance of returns, and the plausible and easy to handle properties of CAPM–equilibria: The Mutual Fund Theorem, the Beta–pricing rule, and the Efficient Frontier.

Those 'nice' results come just into effect at the expense of an ad–hoc assumption, the strict fulfillment of investors' budget identities in the investment as well as in the consumption period. In the General Equilibrium Theory of Incomplete Markets (GEI) this 'assumption' is in fact an implication of the existence of positively paying assets and the monotonicity of utility functions in state pay–offs. The CAPM is sometimes regarded as a special case of GEI–models in the theoretical finance literature (see (Geanakoplos & Shubik, 1990)), although this is only true for very special choices of mean/variance–utility functions and/or probability distributions of returns.

However, for mean/variance–utility functions monotonicity is not necessarily guaranteed. In standard CAPM–economies it is indeed possible that a strict positively paying asset increases the standard deviation of an investor's portfolio in relation to its mean in such a way that the investor values a bit more of this particular asset negatively. If the investor already owns this asset he would prefer to get rid of it rather than to consume it. Thus, investors may be satiated and would not voluntary keep their budget identity in the investment period. From the viewpoint of the General Equilibrium Theory this phenomenon is an unacceptable inconsistency with regard to the paradigm of individual rationality. The non–monotonicity can even lead to equilibrium prices permitting arbitrage opportunities, which one can hardly observe in real financial markets.

In the CAPM the riskless asset is the only investment in which consumers are always locally non–satiated. In the CAPM without a riskless asset, introduced by (Black, 1972), the problem of satiation becomes aggravated, because then the budget identity in the investment period is not a binding constraint just by monotonicity in the riskless asset. It is often argued that in financial markets riskless investments are obviously available, namely state guaranteed zero bonds. A more detailed look reveals that zero bonds are nominal riskless but bear the risk of inflation, i.e. are not riskless with regard to the real numéraire. Thus, the CAPM without a riskless asset is of realistic importance. The complexity of the CAPM increases furthermore if the common assumption is disposed that traders' endowments are marketed.

Hitherto the problem of non–monotonicity has been misinterpreted in that authors derived conditions to ensure positive asset prices. But these conditions do not exclude the possibility that positively paying investment opportunities are traded at a negative (arbitrage) price. In the two–period consumption based CAPM with homogenous expectations, but without a riskless asset and with non–marketed endowments, the problem of non–monotonicity is analysed. Conditions ruling out this problem already exist but are more restrictive and derived in a less general setup, see (Magill & Quinzii, 1996) and (Pilgrim, 1998). In the first essay this issue is embedded in a general overview of the theoretical literature about the CAPM. Different aspects, like the relationship between expected and mean/variance–utility, existence and uniqueness of equilibria, etc., are also addressed. This provides a compact and broad overview about the CAPM.

## 1.2 Foreign Exchange Rates

The prognosis of foreign exchange rates is still an econometric challenge without convincing results hitherto. The empirical literature has shown that fundamental factors, for instance the change of the GNP, do not really contribute to the forecast of exchange rate changes in the short and middle run. Only the change of the money stock has a measurable impact, see e.g. (Frankel & Rose, 1995).

Since the beginning of the 1980s many empirical investigations about the information efficiency of foreign exchange markets came up, see for instance (Hansen & Hodrick, 1980). The hypothesis had been formulated that efficiency implies that the forward exchange rate is an unbiased predictor for the spot exchange rate (see (Fama, 1984)). To avoid Siegel's paradox the logarithms of the forward and the spot exchange rate are usually employed, see (Siegel, 1972). This is better known as the uncovered interest rate parity. However, empirically the uncovered interest rate parity yields not very promising results, see (Frankel & Rose, 1995), (Chiang, 1988), and (Hansen & Hodrick, 1980). The uncovered interest rate parity is also considered to be an arbitrage connection, e.g. in (Winters, 1999).

It will be shown that the uncovered interest rate parity is generally not a no–arbitrage condition and can be considered as an equilibrium condition only non–generically. This result obviously implies that any empirical analysis of the information efficiency relying on the hypothesis of the uncovered interest rate parity bears the risk of systematic errors. The true arbitrage relation, which the exchange rate has to satisfy, will be derived in a time continuous

arbitrage pricing model of Black–Scholes type.
Arbitrage–free prices can only be regarded as a necessary condition for an equilibrium. An economy having an equilibrium at these arbitrage–free prices can always be constructed. Presuming an immediate exploitation of possible arbitrage opportunities the no–arbitrage condition will be employed to determine a law of motion for the foreign exchange rate in the short run if markets are complete. Such an arbitrage–free correlation between the exchange rate, interest rates, and prices of stock indices is theoretically derived for complete asset markets. A reduced form equation of this relation will also be estimated with real data.
In the long run the economies' fundamentals seem to determine their exchange rates. Equilibria of abstract economies are usually not solvable explicitly, so that beyond existence and uniqueness not much more qualitative results can be obtained (see (Karatzas *et al.*, 1990)). For this reason the equilibrium of a quite extensive example economy is considered. In this economy two representative agents provide labor in each country as the only factor in production of two consumption goods. In fact, the relative price between both consumption goods, which are interpreted as the respective numéraires, is the real exchange rate. It will turn out that the exchange rate is equal to the ratio of the present values of all future net imports and exports.

# References

Black, F. 1972. Capital Market Equilibrium with Restricted Borrowing. *Journal of Business*, **45**, 444 – 455.

Chiang, Thomas C. 1988. The Forward Rate as a Predictor of the Future Spot Rate – A Stochastic Coefficient Approach. *Journal of Money, Credit, and Banking*, **20**(2), 212 – 232.

Fama, Eugene F. 1984. Forward and Spot Exchange Rates. *Journal of Monetary Economics*, **14**(3), 319 – 338.

Frankel, Jeffrey A., & Rose, Andrew K. 1995. Empirical Research on Nominal

Exchange Rates. *Chap. 33, pages 1689 – 1729 of:* Grossman, G., & Rogoff, K. (eds), *Handbook of International Economics*, vol. 3. Amsterdam: Elsevier Science B.V.

Geanakoplos, John, & Shubik, Martin. 1990. The General Asset Pricing Model as a General Equilibrium with Incomplete Markets. *The Geneva Papers on Risk and Insurance Theory*, **15**, 55 – 72.

Hansen, Lars Peter, & Hodrick, Robert J. 1980. Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis. *Journal of Political Economy*, **88**(5), 829 – 853.

Karatzas, I., Lehoczky, J. P., & Shreve, S. E. 1990. Existence and Uniqueness of Multi–Agent Equilibrium in a Stochastic, Dynamic Consumption / Investment Model. *Mathemathics of Operations Research*, **15**(1), 80 – 128.

Magill, Michael J. P., & Quinzii, Martine. 1996. *The Theory of Incomplete Markets*. Vol. 1. Cambridge, MA: MIT–Press.

Pilgrim, Jan. 1998. *Eine Numerische Analyse Evolutionärer Finanzmärkte*. M.Phil. thesis, Universität Bielefeld.

Siegel, Jeremy J. 1972. Risk, Interest, and Forward Exchange. *Quarterly Journal of Economics*, **86**(May), 303 – 306.

Winters, Bill. 1999. Masterclass – with JP Morgan: Wrong–Way Exposure. *Risk*, July, 52 – 55.

# 2 The CAPM – A General Equilibrium Foundation

**Abstract:** The Capital Asset Pricing Model (CAPM) is one of the cornerstones in Finance. The CAPM was the first theoretically sound and intuitive description of portfolio choice which met portfolio managers' approval in practice. Very early after its invention by Markowitz it has been extended to a full equilibrium model. In this section the classical two–period CAPM without spanned endowments and without a riskless asset is considered in the framework of the General Equilibrium Theory.

The main issue of this chapter is the monotonicity of mean/variance utility functions. The CAPM imposes the very restrictive assumption that agents' preferences depend only on mean and variance of returns on investment. This assumption may cause non–monotonicity of preferences in state pay–offs and is, thereby, responsible for arbitrage opportunities which equilibrium prices may permit. Explicit conditions are derived ruling out arbitrage and non–monotonicity in equilibrium.

Even though the CAPM had a shadowy existence apart from the General Equilibrium Theory of Arrow and Debreu many theoretical articles appeared. Beside monotonicity this chapter offers a broad overview about the theoretical (not the empirical) CAPM–literature. However, en passant some new proofs and results about well known aspects are also derived. For example a geometrical proof for the correspondence between mean/variance–utility and –preferences is given and the equivalence between absolute and relative risk aversion is shown.

Asset/Liability–Management is at a stage where new approaches are replacing $\mu/\sigma^2$–efficiency, but none of those have gained a similar prominence in equilibrium theory yet. The last issue of this chapter is about new developments in modelling the investment decision under risk. Exemplary two newly proposed risk measures will be considered which emerged from the criticism of variance and Value–at–risk.

## 2.1 Introduction

Although the Capital Asset Pricing Model imposes highly controversial assumptions, no other model has sharpened researches' as well as practitioners' intuition of investors' portfolio decision and the pricing of risks in as much. Its popularity rose due to three major advantages; agents' preferences depend only on mean and variance of their investment, while variance is easily understood by many people as a measure of risk; the main results agree with our intuition; and, equilibria are explicitly solvable in only one endogenous quantity, the harmonic average of absolute risk aversions. In comparison to the General Equilibrium (GE) model (Merton, 1973) noted:

> *While more general and elegant than the capital asset pricing model in many ways, the general equilibrium model of Arrow and Debreu has not had the same impact, principally because of its empirical intractability and rather restrictive assumptions that there exist as many securities as states of nature.*[1]

The idea of ordering distributions according to their means and standard deviations is indeed much older than the CAPM. Its roots go back into the history of famous acturiates like (Tetens, 1789).[2] Into portfolio selection the mean/variance criterion was introduced by (Markowitz, 1952), (Tobin, 1958) and (Markowitz, 1959), who also calculated the efficient set of portfolios in terms of minimal variance for a given level of mean. But the major development towards a full equilibrium model has been undertaken by (Sharpe, 1963), (Sharpe, 1964), (Lintner, 1965b), (Mossin, 1966), and (Treynor, 1961)[3] (see also (Lintner, 1965a), (Lintner, 1970), (Merton, 1972), and (Sharpe, 1970)). In those contributions the well known equilibrium properties of the CAPM were established: The Efficient Frontier, the Beta–pricing formula and the Mutual Fund Separation. They all assumed homogeneous expectations, the

[1] Only complete markets require this assumption. However, empirical intractability is still an issue of GE models.

[2] This hint is due to (Borch, 1969).

[3] The last reference is a hint from (Copeland & Weston, 1979).

existence of a riskless asset and that investors' endowments are given as asset shares to them. (Black, 1972) relaxed the assumption of a marketed riskless asset and (Oh, 1992) allowed for endowments not spanned by existing assets. (Lintner, 1969), (Rabinovitch & Owen, 1978) and (Gonedes, 1976) introduced heterogeneous beliefs into the CAPM, since heterogeneity of priors can be regarded as a major incentive for trade. Critical new insights were not derived, while the model looses its simplicity as individual expectations enter as complex weighted averages. (Güth *et al.*, 2000) show in a CAPM with heterogeneous beliefs that it is not always advantageous to be an insider having a 'better' prior. (Merton, 1973) made a further big step by extending the CAPM to an intertemporal version.

### 2.1.1 Main Issues

This chapter's main purpose is a deeper investigation of a certain criticism of the CAPM: non–monotonicity of preferences. In the CAPM agents' utility functions depend on mean and variance of returns only. Whereas utility is increasing in mean it is decreasing in variance. Moreover, variance is a symmetric measure of the dispersion of returns around mean. Thus, utility depending negatively on variance punishes the pure distance of pay–offs from mean regardless of their sign, even positive pay–offs. Thereby, under realistic circumstances it might be that agents' utility functions are decreasing in pay–offs of certain states. This clearly contradicts the natural understanding that more of the numéraire is desirable. Moreover, monotonicity of preferences is a central axiom in the state–preference approach which implies the existence of positive state prices in any equilibrium of financial markets.

A consequence of non–monotone preferences is that a trader may prefer one of two feasible portfolios although the preferred one is first–order stochastically dominated by the other. For example, suppose an extreme variance averse agent likes a portfolio with lower mean and variance better compared to another portfolio. However, the distribution of the disliked pay–offs may well stochastically dominate the distribution of the preferred pay–offs. This 'contradiction' is better known as the 'Mean–Variance Paradox' in business economics, see [(Copeland & Weston, 1979) p. 85]. Since first–order stochastic dominance of pay–offs preserves under strictly monotonic transformations only, the Mean–Variance Paradox corresponds one–to–one to non–monotonicity of preferences.

Non–monotonicity can even lead to non–positive prices of positive pay–offs in an equilibrium (implying negative state prices), which establishes arbitrage opportunities in the conventional sense.[4] A price system providing arbitrage opportunities would never be viable in real financial markets with selfish investors. This would prove a CAPM to be completely artificial.

In the literature the question of positive asset prices has been addressed [see (Allingham, 1991), (Nielsen, 1992) and (Hiroshi & Hiroshi, 1995)]. It will be shown, that positive prices for assets with positive mean do not rule out arbitrage, while the condition of positive prices for consumption streams with positive mean is unnecessarily strong. Minimal conditions are derived, which are sufficient to rule out arbitrage as well as satiation in state pay–offs.

A related aspect is the critique that preferences depending only on mean and variance are insensitive to the risk of shortfall or of fat tails, measured by skewness and kurtosis, respectively. While non–monotonicity can cause inconsistencies, like arbitrage opportunities, this weakness is only a limitation. It is also a strength, because just two quantities are regarded instead of three or four factors. There are several attempts to include skewness into the Efficient Frontier, see for instance (Jurczenko & Maillet, 2000). A more general approach is to develop consistent axioms a risk measure should satisfy from a normative point of view. The last section points into that direction as two different axiomatics are introduced, which characterize investors' portfolio decision with less controversy but with similar ease. The risk measures are applied to a typical situation an institutional investor faces in the management of assets and liabilities.

[4] However, negative asset prices are well compatible with the no–arbitrage condition if returns are negative with positive probability.

### 2.1.2 Outline of the Chapter

Although it is not the main focus of this essay some space is devoted to general issues of the CAPM. This should not only provide a thorough understanding of the CAPM literature. Indeed, all the calculations leading to the equilibrium allocation are necessary for those mentioned conditions about arbitrage and monotonicity. The starting point of the investigation is the static, two–period CAPM proposed by (Black, 1972). Since non–monotonicity may occur only with regard to pay–offs in certain events, the CAPM is analysed within the state–preference approach. Unrestricted short selling will be permitted and investors shall share a common belief about the likelihood of future events. An agent's decision problem is regarded as purely investment based, i.e. consumption in the decision period is not considered. This view is justified by the empirical observation that consumers often separate a certain share of their total wealth for investments in securities and insurances. Moreover, it is neither assumed that investors' endowments are marketed nor that a riskless asset exists.

This chapter is organized as follows. In the next section the CAPM–assumptions are introduced. Then three issues arising with mean/variance utility are briefly reviewed; (Löffler, 1996b) answered the question how preferences on the consumption space and mean/variance utility functions are related; (Chamberlain, 1983) and (Agnew, 1971) characterized the distributions which imply mean/variance utility if those were derived from expected utility; and finally, (Lajeri & Nielsen, 1994) provided an extensive analysis of the notion of risk aversion in the CAPM.

The next section addresses the problem of satiation. Satiation in income may arise without a marketed riskless asset, which prevents an equilibrium to settle down. In the fourth section the Capital Market Line and traders' optimal demand is derived, followed by a complete characterization of equilibrium properties in the fifth section. Thereafter, in an intermediate section the economy is transformed into the two–goods economy proposed by (Dana, 1999), where consumption takes place in mean and variance only. It prepares the review of results concerning existence and uniqueness of equilibria, notably associated with (Nielsen, 1990b) and (Nielsen, 1988). The next section focuses on the main subject, monotonicity of preferences and arbitrage in CAPM–equilibria. Since the risk measure variance is responsible for the lack of monotonicity of preferences, in the concluding section an outlook about two alternative risk measures is presented. Those have been proposed recently by (Artzner *et al.*, 1998) and (Aspandiiarov *et al.*, 1998), who suggest to replace in the portfolio choice problem 'Value–at–Risk' by 'Expected Shortfall' and variance by 'Weighted Value–at–Risk', respectively.

## 2.2 The CAPM – Preliminaries

This section introduces the assumptions necessary for subsequent investigations. Regarding the CAPM, those assumptions might appear too voluminous to the reader. But this originates from the intention to analyse the CAPM in the framework of the General Equilibrium Theory of Financial Markets. Nevertheless, most assumptions, definitions and notational conveniences are standard in the finance literature (consult Section I.11 in (Duffie, 1988) and (Nielsen, 1987)). Broad remarks are spared therefore.

In their investment decision agents face a risky environment:

**Assumption (ENV):** Let the risky environment be represented by a probability space $(\Omega, \mathcal{F}, P)$. $\Omega$ is interpreted as the set of states $\omega$ of the world, $\mathcal{F}$ is the $\sigma$-algebra of measurable events and $P$ the common subjective probability measure on $(\Omega, \mathcal{F})$. The consumption space $L$ consists of all real valued random variables with finite variance, that is $L := L^2(\Omega, \mathcal{F}, P)$. $\mathcal{F}$ is assumed to be countably generated and $P$ to be $\mathcal{F}$–finite.[5]

**Notation:** An event $\mathbb{E} \in \mathcal{F}$ is called essential if $P(\mathbb{E}) > 0$. Let $x, y \in L$. The relations $=, \neq, \shortparallel, \nshortparallel$ on $L$ are applied in their probabilistic versions,

[5]Those two assumptions allow to carry over the algebra of Euclidean spaces [see (Billingsley, 1995) p. 249]. Both hold for the Borel $\sigma$–algebra and distributions with finite density for example. The latter assumption is a prerequisite for the existence of a Radon–Nikodym derivative.

i.e. $x \nparallel y :\Leftrightarrow \forall \lambda \in \mathbb{R} : \; P(\lambda x = y) < 1$ and likewise for $=, \neq, \parallel$. $L$ is endowed with the $L^2(P)$–inner–product, $x \bullet_P y := E_P(xy) = \int_\Omega xydP$, where $E_P(\cdot)$ denotes the expectation operator with respect to the measure $P$. Orthogonality refers to this inner–product, i.e. $x \perp_P y :\Leftrightarrow x \bullet_P y = 0$. The probability measure on operators is generally suppressed except when $P$ is not meant. $Var(x)$ is the variance of $x$ and $Cov(x, y)$ the covariance of $x$ and $y$. Because formulas would get very lengthy by the brackets of operators (Greek) letters abbreviate the moments with its arguments appearing as indices: $\mu_x \equiv E(x)$, $M_{x,y} \equiv E(xy)$, $\sigma_x \equiv \sqrt{Var(x)}$, $\sigma_{x,y} \equiv Cov(x, y)$ and $\rho_{x,y} \equiv \frac{\sigma_{x,y}}{\sigma_x \sigma_y}$, respectively. Let $L_+ = \{x \in L \mid P(x \geq 0) = 1\}$ and $L_{++} = \{x \in L_+ \mid P(x > 0) > 0\}$ be the non–negative and the positive consumption space, respectively. $\text{car}(L)$ denotes the cardinality of $L$, which is equal to the number of orthogonal, non–zero random variables in $L$. Let $b(X)$ stand for an orthogonal basis of a linear subspace $X \subseteq L$. This basis is defined as a possibly infinite set of non–zero, orthogonal random variables. Whenever a basis is finite it will also be used as a line–vector, otherwise as a net. If a constant is in $X$ it shall be regarded as the first element of the basis or the net.

Financial markets allow to some extend the trading of risks:

**Assumption (MARKET):** Security markets are described by a finite subset, $A \subset L$, of traded securities $\{A^1, \ldots, A^J\}$, where each random variable $A^j$ stands for the random pay–off of asset $j \in \{1, \ldots, J\} := \mathbf{J}$. The span, $X = \langle A \rangle$, of $A$ denotes the linear closed subspace of marketed consumption plans. Competitive and frictionless markets are assumed, which also permit unlimited short selling. Therefore it can be assumed without loss of generality that assets are linearly independent, i.e. $\dim(X) = J$.[6] The column vector $q : A \to \mathbb{R}^J$ assigns to every security a price. Portfolios are denoted by the column vector $\theta \in \mathbb{R}^J$ and

[6] In a frictionless world redundant assets are needless. However, they are not dispensable if intermediation costs (see (Bettzüge *et al.*, 2000)) or short selling constraints (see (Elsinger & Summer, 1998)) are effective.

deliver a random income of $\sum_{j\in\mathbf{J}} A^j\theta_j$. Positive and negative entries in $\theta$ correspond to purchases and sales respectively.

A riskless asset is traded only if the following assumption is imposed:

**Assumption (BOND):** A riskless asset is marketed:

$$\mathbf{1} := \mathbf{1}_\Omega(\omega) \in X.$$

**Remark 1** *Since it is not critical for the analysis, the consumption space is not specified to represent either a physical commodity or paper money. Note that real financial markets admit at best* monetary *riskless securities. This implies a riskless asset based on* commodities *only if the risk of inflation can be completely hedged away.*

**Notation:** $z_X$ stands for the orthogonal projection of $z \in L$ onto $X$. Thence $X^\perp$ denotes the orthogonal space of $X$. The set $A$ is also used as a row vector, i.e. $A = \left(A^1, \ldots, A^J\right)$. Thus, $A\theta$ denotes portfolio pay–offs. Without loss of generality let $A^1 = \mathbf{1}$ if (BOND) holds. Although '**1**' is merely a scalar the bold one **1** as well as the bold zero **0** is used as such in expressions involving random variables.[7] For a finite state space this allows for the interpretation of those expressions as vectors. If not stated otherwise any vector is a column–vector. The upper index '$T$' marks the transpose of a matrix or a vector. Moreover $\mu_A$ denotes the vector of component wise means of $A$, $\sigma_A^2$ and $M_A$ the covariance matrix and the matrix of second moments of $A$ respectively. $\left(\sigma_A^2\right)^{-1}$ is the inverse of $\sigma_A^2$ as well as the Moore–Penrose[8] inverse of $\sigma_A^2$ in case of (BOND), which then reads:

$$\begin{pmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \left(\sigma^2_{A\backslash 1}\right)^{-1} \end{pmatrix},$$

[7] Remind that $x = 0$ $P-a.s.$ is abbreviated with $x = \mathbf{0}$, etc.!

[8] For a definition of the Moore–Penrose inverse see the last appendix of the next chapter.

where $\sigma^2_{A\backslash 1}$ is the covariance matrix without the first asset and $\mathbf{0}$ is a vector filled with zeros. For any $x \in X$ the unique generating portfolio[9] $\theta_x$ is defined by $A\theta_x = x$.

An important distinction between financial market economies is how much of the entire risks are allowed for trading.

**Definition 1 ((In)complete Markets)** *Markets are called either complete or incomplete if either $X = L$ or $X \subsetneq L$ holds, respectively. A market $\hat{X}$ is less complete than another market $X$ if and only if $\hat{X} \subsetneq X$.*

**Example 2** *If the number of essential events in $\mathcal{F}$ is finite then completeness is equivalent to* $\dim(X) = J = \mathrm{car}(L)$. *Since the state space may be infinite, markets can only be complete if the set of essential events is finite. Only then any risk can be insured by a finite number of assets. If for example $\Omega = \mathbb{R}, \mathcal{F} = \mathbb{B}(\mathbb{R})$ and $P$ is the standard normal distribution of a real random variable $x$ the Arrow–security $\mathbf{1}_\Omega (x \in A)$ is for any essential $A \in \mathcal{F}$ a measurable, square integrable random variable, which is linearly independent of $x$. The set of such Dirac–measures is infinite.*

Agents are endowed with preferences defined on the entire consumption space:

**Assumpion (AGENTS):** Agents[10], $i \in \{1, \dots, I\} =: \mathbf{I}$, have preferences $\precsim_i$ on $L$, which can be represented by a utility function $U^i$ on $L$ :

$$\forall i \in \mathbf{I} : \exists v^i : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}, \text{ such that}$$
$$(\mu_x, \sigma^2_x) \longmapsto v^i(\mu_x, \sigma^2_x) =: U^i(x).$$

[9] Since the relationship between $x$ and $\theta_x$ is one–to–one, the differentiation will not be strict between assets, portfolios, pay–offs, consumption streams, consumption plans, returns etc., which are merely expressions of the same thing. With "returns" total returns and not rates of return are meant, because the latter involves prices determined endogenously.

[10] Agents, investors, consumers, traders, individuals, etc. are not distinguished.

The $\mu/\sigma^2-$utility function $v^i$ is assumed to be strictly increasing in mean and strictly decreasing in variance[11], strictly quasi–concave in $(\mu,\sigma)$ and twice differentiable[12] in $(\mu,\sigma^2)$. The derivative for $\sigma^2$ at $(\cdot,0)$ is taken to be the derivative from the right. To express utility in mean and standard deviation the abbreviation $\hat{v}^i(\mu_x,\sigma_x) \equiv v^i(\mu_x,\sigma_x^2)$ is used whenever it is appropriate. Agents' endowments are denoted by $w^i \in L$. Their decision problem is to maximize utility by choosing a consumption plan and a portfolio for a given price vector $q$ :

$$D^i\left(q\right) := \arg \max_{(x,\theta)\in\mathbb{B}^i(q)} U^i(x), \tag{$D^i$}$$

where the budget set reads:

$$\mathbb{B}^i\left(q\right) := \left\{ (x,\theta) \in L \times \mathbb{R}^J \middle| \; x = A\theta + w^i_{X^\perp} \text{ and } q^T\theta \leq q^T\theta_{w^i_X} \right\}.$$

Two periods are associated with the decision problem: Today agents choose their optimal portfolio and tomorrow pay–offs are realized and consumption takes place. Agent $i$'s consumption of state pay–offs, mean, and standard deviation is denoted by $x^i$, $\mu^i \equiv \mu_{x^i}$, and $\sigma^i \equiv \sigma_{x^i}$ respectively, and by $x^i_X = A\theta^i$, $\mu^i_X \equiv \mu_{x^i_X}$, and $\sigma^i_X \equiv \sigma^2_{x^i_X}$ respectively, if only the consumption in $X$ is concerned.

**Remark 2** *Agents are forced to consume their second period income. So free disposal in the future period is not permitted, although they have not necessarily monotone preferences. This problem will be considered later on.*

**Notation:** Sometimes an agent's index is suppressed, if not a particular investor is meant. For any variable $y^i$ let $y^{\mathbf{I}} := (y^i)_{i\in\mathbf{I}}$ and $\bar{y} := \sum_{i\in\mathbf{I}} y^i$ (if defined).

[11] Instead of variance some authors use standard deviation. Except of some mathematical regularities, e.g. differentiability, both measures of risk can be used equivalently since they are monotone transformations from each other.

[12] This assumption allows to use differential calculus and to simplify notation thereby. Differentiability in $(\mu,\sigma)$ is weaker in some pathological cases than in $(\mu,\sigma^2)$.

A consumption plan $x \in L$ is said to be 'marketed' or 'spanned' if it is also in $X$. The assumption that all endowments are marketed simplifies considerably the notation.

**Assumption (SPANNING):** The endowments of each agent lie in the marketed subspace:

$$\forall i \in \mathbf{I} : w^i \in X.$$

The importance of $\mathbf{1}_X$ and $\bar{w}_X$ in the subsequent analysis, i.e. the orthogonal projection of the riskless asset $\mathbf{1}$ and the aggregated endowments $\bar{w}$ onto the marketed subspace $X$, justifies their own definition:

**Definition 3** *In the following* $\mathbf{1}_X$ *is called the* ***Quasi–Bond*** *and* $\bar{w}_X$ *the* ***market asset****.*

The properties of the Quasi–Bond are developed in the next section. If the (BOND) is not marketed the Quasi–Bond shall take up its place in the asset vector $A$ and the orthogonal basis $b(X)$.
Not all assumptions are essential to a CAPM–economy:

**Definition 4 (CAPM)** *A tuple* $\left(L, (U^i, w^i)_{i\in\mathbf{I}}, A\right)$ *is called a* ***CAPM–economy*** *or simply a* ***CAPM*** *if it satisfies the assumptions* ***(ENV), (MARKET)*** *and* ***(AGENTS)****.*

The equilibrium concept corresponds to the definition of a financial market equilibrium in the General Equilibrium Model with Incomplete Markets (abbreviated with GEI), see (Magill & Quinzii, 1996):

**Definition 5 (EQU)** *For a given CAPM–economy a CAPM–equilibrium is a tuple* $\left((x^{*i})_{i\in\mathbf{I}}, (\theta^{*i})_{i\in\mathbf{I}}, q^*\right) \in L^I \times \mathbb{R}^{J\cdot I} \times \mathbb{R}^J$ *ensuring that:*

*i. Markets for consumption plans clear:* $\sum_{i\in\mathbf{I}} x^{*i} = \sum_{i\in\mathbf{I}} w^i$.

*ii. Financial markets clear:* $\sum_{i\in\mathbf{I}} \theta^{*i} = \sum_{i\in\mathbf{I}} \theta_{w^i_X}$.

iii. *For all* $i \in \mathbf{I}$ *their equilibrium consumption* $x^{*i}$ *and asset demand* $\theta^{*i}$ *solve* $D^i$ *at prices* $q^*$, *i.e.* $\forall i \in \mathbf{I} : \left(x^{*i}, \theta^{*i}\right) \in D^i\left(q^*\right)$.

Note that under the equilibrium condition *iii.* both market clearing conditions are equivalent. This is not the case if free disposal is permitted in the second period. Then an exchange equilibrium *i.* implies the clearing of financial markets *ii.* but not vice versa.

**Lemma 6** *Suppose Condition iii. of Definition 5 holds for a particular allocation. Then the two market clearing conditions i. and ii. are equivalent. It is incompatible with q being an equilibrium asset price vector that* $q^T\theta^{*i} < q^T\theta_{w_X^i}$ *for some* $i \in \mathbf{I}$.

**Proof.** The Lemma follows by some simple equivalent transformations of the market clearing conditions:

$$\begin{aligned}
\sum\nolimits_{i\in\mathbf{I}} x^{*i} &= \sum\nolimits_{i\in\mathbf{I}} w^i && \Leftrightarrow \\
\sum\nolimits_{i\in\mathbf{I}} \left(x^{*i} - w^i_{X^\perp}\right) &= \sum\nolimits_{i\in\mathbf{I}} w^i_X && \overset{\text{by } \left(x^{*i},\theta^{*i}\right)\in\mathbb{B}^i(q^*) \text{ and } w^i_X=A\theta_{w^i_X}}{\Longleftrightarrow} \\
A\sum\nolimits_{i\in\mathbf{I}} \theta^{*i} &= A\sum\nolimits_{i\in\mathbf{I}} \theta_{w^i_X} && \overset{\text{by } rg(A)=J>0}{\Longleftrightarrow} \\
\sum\nolimits_{i\in\mathbf{I}} \theta^{*i} &= \sum\nolimits_{i\in\mathbf{I}} \theta_{w^i_X} && \overset{\text{by } q^T\theta^{*i}\leq q^T\theta_{w^i_X}}{\Longrightarrow} \\
\forall i \in \mathbf{I} : q^T\theta^{*i} &= q^T\theta_{w^i_X}.
\end{aligned}$$

■

In the sequel some further assumptions are imposed on the CAPM–economy. But most of them must hold only on a certain range of feasible allocations, especially all those which come into question for an equilibrium allocation. The following definition specifies this range:

**Definition 7** *The set on which equilibrium prices, portfolios and allocations can be restricted to without knowledge of endogenous variables of equilibria* *is called the* relevant equilibrium range*, abbreviated* RER*.*

$$RER := \Big\{ \left(x^{\mathbf{I}}, \mu^{\mathbf{I}}, \sigma^{\mathbf{I}}, \theta^{\mathbf{I}}\right) \in L^I \times \mathbb{R}^I \times \mathbb{R}^I_+ \times \mathbb{R}^{I \cdot J} \big| \sum\nolimits_{i \in \mathbf{I}} \theta^i = 0, \forall i \in \mathbf{I} : \\ x^i = w^i + A\theta^i, U^i\left(x^i\right) \geq U^i\left(w^i\right), \mu^i = E\left(x^i\right), \sigma^i = \sqrt{Var\left(x^i\right)} \Big\}$$

*By writing $x^{\mathbf{I}} \in RER$ or $\mu^{\mathbf{I}}, \sigma^{\mathbf{I}} \in RER$ for notational simplicity the existence of the rest of the tuple is demanded such that the entire tuple is in $RER$.*

This is treated as a *dynamic* definition in the sense that the restrictions to $RER$ will be tightened and the tuple will be enlarged, e.g. to cover also asset prices. Thus, the definition should not be taken too definitely.
The notation of the differential calculus is shortened by the usual abbreviation of derivatives:

**Notation:** As subscripts the positions $k, l$ or the names $x, y$ of the arguments of a function $f$ indicate partial derivatives, i.e. $f_{k,l} \equiv f_{x,y} := \frac{\partial^2}{\partial x \partial y} f$.[13]

Generally the examples of this chapter are given in one of the two most popular simplifications of the CAPM in the literature, the case of quadratic utility functions and of $\mu/\sigma^2$–linear utility functions:

**Definition 8 (QUADU)** *Agents have 'Quadratic Utility Functions'* :⇔

$$\forall i \in \mathbf{I} : U^i(x) = E\left(x - \frac{\varphi^i}{2}x^2\right) = \mu_x - \frac{\varphi^i}{2}\left[\sigma_x^2 + \mu_x^2\right],\ \varphi^i > 0.$$

**Definition 9 (LINU)** *Agents have '$\mu/\sigma^2$–Linear Utility Functions'* :⇔

$$\forall i \in \mathbf{I} : U^i(x) = \mu_x - \frac{\varphi^i}{2}\sigma_x^2,\ \varphi^i > 0.$$

Quadratic utility functions do not satisfy the assumption (AGENTS) because utility is not increasing in mean for $\mu_x \geq \frac{1}{\varphi^i}$. Thus, (QUADU) must be accompanied by the additional assumption that for all agents $\mu^i \geq \frac{1}{\varphi^i}$ are

[13] There should be no danger of mixing up operators and derivatives, e.g. $\mu_x$ and $f_x$. Indices are either numbers or scalars on functions or random variables on operators.

not in $RER$. $\mu/\sigma^2$–linear utility functions can be derived from expected utility with an exponential utility function and normally distributed pay–offs.[14]

## 2.3 Mean–Variance Preferences

$\mu/\sigma^2$–utility functions are regarded as an alternative to the state preference approach used in General Equilibrium Theory of financial markets. While the later claims the agents to be in a kind extremely rational as they must foresee all essential events, the former approach just involves two reference numbers: mean and variances. Although it is not necessary to model the states of the world explicitly in a CAPM, this is done here for reasons of comparison to the GEI–model; the reader might think of arbitrage for instance.

This section focuses on CAPM–preferences. In the first subsection the question is answered which characterization of preferences on $L$ corresponds one–to–one to mean/variance utility functions. In the second the requirement is discussed which has to be imposed onto the distribution of assets such that expected utility implies mean/variance preferences. The purpose of the last subsection is a descriptive analysis of the indifference curves and to introduce a measure of risk aversion for mean/variance preferences.

### 2.3.1 Preferences and Utility Functions

It is theoretically more profound to start with axioms about preferences on consumption opportunities and to deduce utility functions from those axioms. A formal definition of mean–variance preferences was proposed by Duffie [see (Duffie, 1988) p. 95 D.]:

**Definition 10 ($\mu/\sigma^2$–Preferences)** *A $\mu/\sigma^2$–preference relation is a complete continuous pre–order $\succ$ on $L$ with two properties:*

[14]Let $x \sim N\left(\mu, \sigma^2\right)$. Then $E\left[\exp\left(-\alpha x\right)\right] = \exp\left(\alpha\mu - \frac{\alpha^2}{2}\sigma^2\right)$, which is a monotone transformation of a $\mu/\sigma^2$–linear utility function. This kind of $\mu/\sigma^2$–linear utility functions has been considered by (Lintner, 1970) and intensively analysed by (Nielsen, 1990a).

1. $\succ$ *is 'strictly variance averse'* $:\Leftrightarrow$

$$\forall x, y \in L,\ \sigma_y^2 > 0 \ \ \text{and} \ \ \sigma_{x,y} = \mu_y = 0 \ \ \text{implies} \ \ x \succ x + y,$$

2. $\succ$ *is 'strictly monotone in the riskless asset'* $:\Leftrightarrow$

$$\forall \lambda \in \mathbb{R}_{++}, x \in L : x + \lambda \mathbf{1} \succ x.$$

Duffie regarded those axioms to be weaker than the assumption of $\mu/\sigma^2$–utility functions. However, (Löffler, 1996b) proved the equivalence between the $\mu/\sigma^2-$preference and the $\mu/\sigma^2$–utility approach:

**Theorem 11** *Consider an arbitrary set* $\hat{A} \in L$ *such that* $\hat{X} = \left\langle \hat{A} \right\rangle$, $\infty > \dim\left(\hat{X}\right) \geq 3$ *and (BOND) holds for* $\hat{X}$. *Then a* $\mu/\sigma^2-$*preference relation on* $\hat{X}$ *implies a continuous* $\mu/\sigma^2-$*utility function on* $\hat{X}$ *and vice versa.*

Löffler proves the theorem analytically, while the following proof gives his arguments a geometric interpretation.
**Proof.** Obviously a $\mu/\sigma^2-$utility function induces a $\mu/\sigma^2-$preference relation. The proof of the reverse direction is the difficult part:
Let $b\left(\hat{X}\right)$ denote an orthonormal basis of $\hat{X}$ in which the first asset is the bond. Without loss of generality let $\hat{A} = b\left(\hat{X}\right)$. All portfolios inducing mean zero and a certain variance $\sigma^2$ lie on a circle in $\mathbb{R}^{\dim(X)-1}$ with radius $\sigma$. This is because all risky assets in the basis have zero mean, a standard deviation of one and in pairs zero correlation. Since there is a one–to–one correspondence between generating portfolios and consumption streams in $\hat{X}$, it is sufficient to consider the indifference surfaces implied by a $\mu/\sigma^2-$preference relation in the space of portfolios.

The first dimension of a portfolio fixes the mean. Neglect the first dimension, i.e. suppose an arbitrary $\mu$. Consider the graph of the indifference relation in $\mathbb{R}^{\dim(X)-1}$. By variance aversion the upper–contour set contains all line–segments between points on the indifference surface and zero because of diminishing variance towards zero. Thus, the upper–contour set is topological invariant to the ball by continuity. In $\mathbb{R}^{\dim(X)-1}$ any non–zero portfolio orthogonal to a portfolio on a circle (thus, any tangential deviation) worsens this portfolio, since it has zero mean and zero covariance with it (see also the following Figure (1)). This holds because the circle is the only geometric object, for which the tangential and the orthogonal hyperplane coincide. If the graph of the indifference relation deviates form the circle at any point, it is possible to draw a circle through that point and to find there a tangent–vector, which points into the upper–contour set. This contradicts the deviation of the indifference curves from a circle. Thus, indifference holds for portfolios with equal mean and equal variance.[15]

This shows that any $\mu/\sigma^2-$preference relation on $L$ is indeed a preference relation on $\mathbb{R} \times \mathbb{R}_+$. Since it is assumed to be continuous, complete and transitive the rest of the proof carries over from the usual representation of preferences by utility functions, see [(Mas–Colell *et al.*, 1995) Proposition 3.C.1 on p. 47]. ■

Differentiability is an additional property of $\mu/\sigma^2-$utility functions, which corresponds to the differentiability of the indifference surfaces of $\mu/\sigma^2-$preference relations.

**Example 12** *The axes of Figure (1) measure the demand for two uncorrelated assets both of which having zero mean and unit standard deviation. There are two indifference contours in the plot, one which belongs to possible $\mu/\sigma^2-$preferences, called 'AGENTS', and one, 'not AGENTS', which does not satisfy variance aversion. The upper contour sets are enclosed by the indifference surfaces, whereas the lower contour sets are outside, since more of one asset is worse. At portfolio 'A' a deviation in the orthogonal direction towards portfolio 'B' would increase variance. But this deviation points into the upper contour set of 'not AGENTS', showing their relative higher variance aversion coming from Asset 1 than from Asset 2. But $\mu/\sigma^2-$preferences demand agents to be neutral with respect to the source of variance. Only for*

[15](Löffler, 1996b) proves the theorem for positive convex cones as the marketed subspaces. In this case the corresponding segments of the circles have to be considered.

*a circle an orthogonal deviation is also tangential at the indifference surface, which is mirrored in the preference surface denoted by AGENTS.*

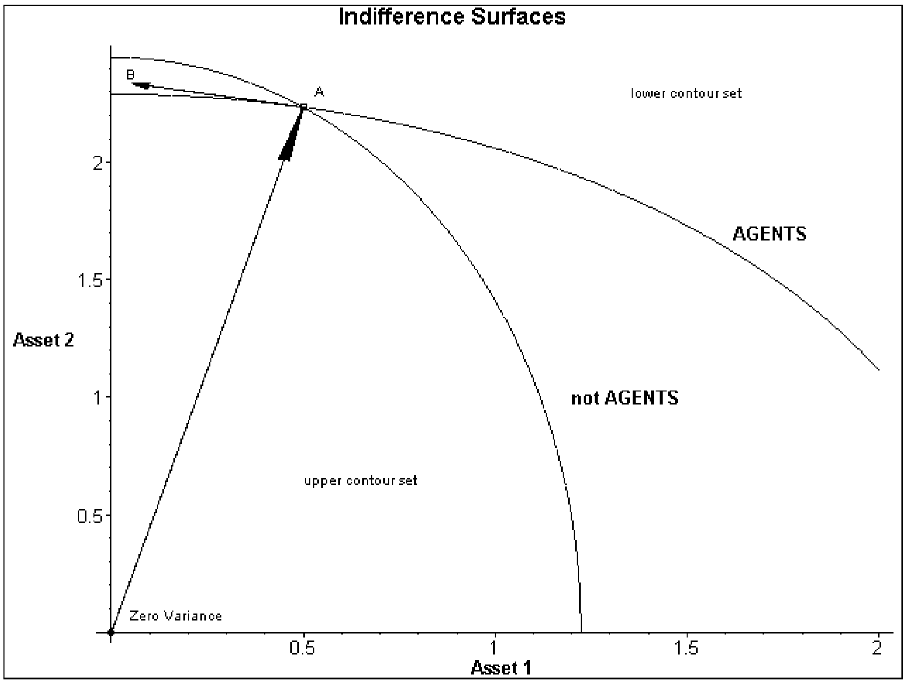


*Figure 1: Indifference Surfaces Concerning Risky Assets*

This line of reasoning does not work if there is only one risky asset, because a circle cannot be constructed in $\mathbb{R}$. Since the proof holds for arbitrary subspaces of $\hat{X}$ with cardinality 3 the finite cardinality of $\hat{X}$ is dispensable as long as the bond is contained (or without the bond for risky assets with zero mean). The restriction of $\prec$ on $\hat{X}$ is a slight generalization over preferences on $L$, since agents have only to be aware of the consumption plans in $X$. This restriction does not mean that the subspace $\hat{X}$ on which preferences are defined is actually the true marketed subspace $X$. Only $X \subseteq \hat{X}$ and (SPANNING) with respect to $\hat{X}$ must hold. Therefore (BOND) for $\hat{X}$ is hardly a restriction since it must not hold for $X$.

**Example 13** *Let* $(\mathbf{1}, b_2, b_3, b_4)$ *be an orthonormal basis of* $L$*. The utility* *function* $U : L \to \mathbb{R}$*, defined by*

$$U(x) = \mu_x - \alpha_2 |\sigma_{x,b_2}| - \alpha_3 |\sigma_{x,b_3}| - \alpha_4 |\sigma_{x,b_4}| - \alpha_5 |\sigma_{x,b_4}| \, 1_{\{\sigma_{x,b_4} \geq 0\}}$$

*for some* $\alpha_1, \ldots, \alpha_5 > 0$ *implies* $\mu/\sigma^2$*-preferences on* $\hat{X}$ *only if* $b_4 \in \hat{X}^{\perp}$*.*

### 2.3.2 Expected Utility and CAPM–Preferences

In contrast to an arbitrary $\mu/\sigma^2$–utility function monotonicity is assured for a von–Neumann–Morgenstern (abbreviated by 'vNM') representation of a $\mu/\sigma^2$–utility function: $U(x) = E(u(x))$. Thus, if $\mu/\sigma^2$–preferences were founded by a suitable choice of state dependent preferences, this would guarantee the monotonicity of $\mu/\sigma^2$–preferences in pay–offs of essential events by the monotonicity of the Bernoulli utility function $u(\cdot)$. This foundation is generally not possible. Note that for (QUADU) the mean/variance utility function results from an expected quadratic utility function, whose Bernoulli utility function is not strictly monotone. Expected utility and mean/variance utility are two independent sights of how agents might evaluate risk. In the special case in which the probability distribution of consumption plans $F$ is parameterized in mean and variance only, $\mu/\sigma^2$–preferences could be derived from expected utility. Actually, the condition

$$\forall x \in X + w^i : v^i(\mu_x, \sigma_x^2) = \int_{\mathbb{R}} u(x) \, F\left(dx, \mu_x, \sigma_x^2\right)$$

must hold for a strictly increasing vNM utility function $u : \mathbb{R} \to \mathbb{R}$ to represent a $\mu/\sigma^2$–preference relation. Suppose, for the moment, that (SPANNING) holds; then $X + w^i = X$ is valid. Moreover, (BOND) shall hold until further notice since risky assets will be normalized to have zero mean.

Since any consumption is a linear combination of random variables in $X$, namely of the assets in $A$, not only these particular random variables must have the *same* two parameter distribution $F(\mu, \sigma^2)$ in mean and variance but also all linear combinations. Thus, the distribution of assets must satisfy the Identity in Law:

$$\exists y \in X : \forall x_1, x_2 \in X, \lambda_1, \lambda_2 \geq 0 : \exists \gamma_1 \geq 0, \gamma_2 \in \mathbb{R} :$$
$$\lambda_1 x_1 + \lambda_2 x_2 \overset{d}{\sim} \gamma_1 y + \gamma_2 \mathbf{1}.$$

The Identity in Law implies a symmetric distribution of consumption streams around their mean, i.e.

$$\exists x \in X : \mu_x \mathbf{1} - x \overset{d}{\sim} x - \mu_x \mathbf{1}.$$

This is necessary because short sales are permitted. Arbitrary but symmetric distributions would be sufficient if there is only one risky asset (including discrete distributions). For at least two *stochastically independent* risky assets only the class of stable distributions is closed under variations of scale and location. All stable distributions have a certain, four parameter characteristic function in common, see for instance [ (Embrechts *et al.*, 1997) Ch. 2.2.]. And from those stable distributions *only* the normal distribution has finite variance and is not skewed, as was already pointed out by (Cass & Stiglitz, 1970) in the context of two fund separation and expected utility. Therefrom comes the popular but wrong belief that the CAPM is only compatible with normally distributed returns in the vNM–utility context.

(Agnew, 1971) gave a counter–example if assets are *not* stochastically independent. A slight modification of his example is the following:

$$A_1 = \mathbf{1} \text{ and } \forall j = 2, \ldots, J : A_j = x_j y,$$

where all risky assets have finite variance and a mean of zero, and the elements of all risky assets are in pairs independent. Thus, assets themselves are uncorrelated but not independent. His specification follows in the next example.

The way he constructed his example describes the class of characteristic functions, such that invariance of distribution due to scale and location is assured. Agnew points out: Since any portfolio $\theta \in \mathbb{R}^J$ is only valued by its mean and variance there has to be a symmetric distribution $F$ with mean zero and variance one such that:

$$\forall\theta \neq \mathbf{0} : (A - \mu_A)\left(\sigma_A^2\right)^{-\frac{1}{2}}\theta\frac{1}{\|\theta\|} \overset{d}{\sim} F. \qquad \text{(Concentric)}$$

This allows for a parameterized $F$ but whose parameters apply to all portfolios in the same fashion. Let $\phi : \mathbb{R}^J \to \mathbb{C}$ and $\psi : \mathbb{R} \to \mathbb{C}$ be the characteristic functions of the distribution of assets and $F$, respectively. Then the characteristic functions must satisfy

$$\begin{aligned}
\phi(\theta) &= \exp(i\mu_A\theta)\,\psi\left(\left(\theta^T\sigma_A^2\theta\right)^{\frac{1}{2}}\right), \text{ where} \\
\psi(t) &= \int_{\mathbb{R}} \exp(itx)\,F(dx) \text{ with} \\
F(x) &= 1 - F(-x), \quad 0 = \int_{\mathbb{R}} xF(dx) \text{ and } \int_{\mathbb{R}} x^2F(dx) = 1,
\end{aligned}$$

for all portfolios $\theta \in \mathbb{R}^J$ (see also Appendix 2.11.1 for the calculation implying this condition). Thus, applied to vNM–utility functions arbitrary $\phi$ and $F$ satisfying these properties admit mean/variance utility functions.

**Example 14** *Agnew's example is a $\psi$ which mixes Gaussian characteristic functions,* $\exp\left(\frac{-\lambda t^2}{2}\right)$*, by a variance scaling parameter $\lambda \geq 0$, which is itself exponentially distributed: $dG(\lambda) = \exp(-\lambda)\,d\lambda$. This results in the characteristic function:*

$$\begin{aligned}
\psi(t) &= \int_{\mathbb{R}_+} \exp\left(\frac{-\lambda t^2}{2}\right)\exp(-\lambda)\,d\lambda \\
&= -\left(\frac{t^2}{2} + 1\right)^{-1}.
\end{aligned}$$

*Then a single asset is distributed as $x_j\sqrt{\lambda}$, where the $x_j$, $j \in \mathbf{J}$ are indepen-*

*dently standard normal. The density function $f(z)$ of asset pay–offs reads*

$$\begin{aligned} f(z) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \exp\left(-\frac{z^2}{x^2}\right) dx \\ &= \frac{1}{\sqrt{2}} \exp\left(-\sqrt{2}\,|z|\right). \end{aligned}$$

*Thus, assets are two–sided exponentially distributed.*

But Agnew's condition provides a deeper insight than his example. From the definition of characteristic functions the equation:

$$\forall \theta \in \mathbb{R}^{J-1} : \int_{\mathbb{R}^J} e^{iA\theta} P^A(dA) = e^{i\mu_A\theta} \int_{\mathbb{R}} e^{i\left(\theta^T \sigma_A^2 \theta\right)^{\frac{1}{2}} y} F(dy)$$

follows. Note that the riskless asset has a degenerated distribution and needs not to be considered thereby. By normalizing assets instead of portfolio returns the equation is equivalent to

$$\int_{\mathbb{R}^{J-1}} e^{i(A-\mu_A)\left(\sigma_A^2\right)^{-\frac{1}{2}}\theta} P^A(dA) = \int_{\mathbb{R}} e^{i\|\theta\| y} F(dy).$$

The distribution $\hat{P}^A$ of normalized asset returns is defined by

$$\begin{aligned} \hat{P}^A(B) &:= P^A\left(\hat{B}\right) \ \forall B \in \mathbb{B}(\mathbb{R})^{J-1}, \text{ with} \\ \hat{B} &= \left\{ \left( x^T \left(\sigma_A^2\right)^{\frac{1}{2}} + \mu_A \right) \Big| \, x \in B \right\}. \end{aligned}$$

This definition simplifies the equation to

$$\int_{\mathbb{R}^{J-1}} e^{iA\theta} \hat{P}^A(dA) = \int_{\mathbb{R}} e^{i\|\theta\| y} F(dy).$$

Integrating the left hand side over the orthogonal space $\langle\theta\rangle^{\perp}$ and replacing on the right hand side $y$ by $\|\theta\|\,\lambda$ it transforms to

$$\int_{\mathbb{R}} e^{i\|\theta\|^2\lambda} \hat{P}^A(\theta d\lambda) = \int_{\mathbb{R}} e^{i\|\theta\|^2\lambda} F(\|\theta\|\, d\lambda),$$

with $F(\|\theta\|\, d\lambda) := d_\lambda F(\|\theta\|\,\lambda)$. Since the equation shall hold for arbitrary

$\theta \in \mathbb{R}^J$, this condition is equivalent to the equality of the differentials:

$$\forall \theta \in \mathbb{R}^{J-1} : \hat{P}^A (\theta d\lambda) = F (\|\theta\| d\lambda)$$

for a symmetric, univariate distribution $F : \mathbb{R} \to [0,1]$ with mean zero and unit variance. [This follows from Theorem 22.2. p. 286 in (Billingsley, 1995)]. By taking the $j-th$ unit vector for $\theta$ all normalized assets are $F$–distributed. The differential of the marginal distribution of $\hat{P}^A$ depends in any direction only on the norm of $\theta$. Thus, the probability mass is concentrically distributed about the origin. Since probability measures are assumed to be $\mathcal{F}$–finite, a continuous distribution $\hat{P}^A$ admits the density $\hat{p}^A$ and $F$ the density $f$. In this case the density $\hat{p}^A$ depends only on the norm of $\|\theta\|$, because

$$\begin{aligned} \hat{P}^A (\theta d\lambda) &= d_\lambda \hat{P}^A (\theta\lambda) = \hat{p}^A (\theta\lambda) \, d\lambda \\ &= F (\|\theta\| d\lambda) = f (\|\theta\lambda\|) \, d\lambda. \end{aligned}$$

The interpretation is as follows: The variance criterion demands that portfolios of normalized asset returns have the same distribution of pay–offs, whenever their norms are equal, since the variance of their pay–offs are equal. That is to say, for any portfolio $\theta$ one can choose a simple portfolio consisting of $\|\theta\|$ units of a single, normalized asset whose distribution is $F$. If that were not the case, one could choose a von–Neumann–Morgenstern utility function to uncover the distinctions in the distributions.
Therefore, the joint density of asset pay–offs is determined by the norm of pay–offs only. Thus, one arrives at the following conclusion:

**Lemma 15** *Suppose (BOND) holds. Then expected utility implies mean/ variance utility if and only if normalized pay–offs:*

$$(A - \mu_A) \left(\sigma_A^2\right)^{-\frac{1}{2}},$$

*of linear independent risky assets in $A$ admit a $(J-1)$ –dimensional distribution $\hat{P}^A$, which*

*i. is concentric about the origin, i.e. the differential $d_\lambda \hat{P}^A (Y\lambda)$ depends* *only on the norm* $\|Y\lambda\|$ *for all* $Y \in \mathbb{R}^{J-1}, \lambda \in \mathbb{R}$.

*ii. admits the identity as its covariance matrix.*

For instance the symmetric generalized hyperbolic distribution, which has been introduced to economics (and overall) by (Barndoff-Nielsen, 1978), satisfies these conditions. It is indeed a mixture of multivariate normal distributions. Since the full density function is a lengthy expression of six parameters, only a certain example of the standard symmetric case is provided here:

$$f_{\alpha,\gamma}(x) = \frac{\alpha^J}{(2\pi)^{(J-2)/2}\, 2\alpha\delta K_{J/2}(\delta\alpha)} e^{-\alpha\sqrt{\delta^2+\|x\|^2}},$$

for some real parameters $\alpha > 0, \delta > 0$, which determine variance and kurtosis, and the modified Bessel function $K_{J/2}$ of third kind with index $J/2$. The marginal distributions of normalized assets may be discrete as well, but only if assets are strongly dependent, for example if the probability is uniformly distributed on some circles about the origin.

Also (Chamberlain, 1983) used Equation (Concentric) to characterize the distributions implying the mean/variance criterion. Since the condition must hold for arbitrary portfolios it is equivalent to:

$$yS \overset{d}{\sim} y \text{ with } y = (A - \mu_A)\left(\sigma_A^2\right)^{-\frac{1}{2}} \qquad \text{(Spherical)}$$

for all unitary matrices $S \in \mathbb{R}^J \otimes \mathbb{R}^J$ with $S^T S = id$. A random vector $y$ satisfying this condition is called spherically distributed about the origin, whereas $A$ is called spherically generated. Thus, Lemma (15) just spells out the restriction on the distribution imposed by spherically distributed normalized assets. This implicitly assumes either a (BOND) or zero mean for $A$, since $A$ is corrected by its mean. But Chamberlain also shows that Condition (Spherical) is necessary and sufficient if '$\overset{d}{\sim}$' is considered as conditional with respect to the Quasi–Bond. Moreover, he gives a representation for all random pay–offs more general than Equation (Concentric):

$$\forall x \in X : x \overset{d}{\sim} \frac{\mu_x}{\mu_{\mathbf{1}_X}} \mathbf{1}_X + \sqrt{\frac{\sigma_x^2 - \mu_x \left(1 - \mu_{\mathbf{1}_X}\right)}{Var\left(vu_1\right)}} vu_1,$$

where the vector $u = (u_1, \dots, u_{J-1})^T$ is uniformly distributed on the unit sphere in $\mathbb{R}^{J-1}$ conditionally on $v$ and $\mathbf{1}_X$, and with $v$ and $\mathbf{1}_X$ having finite variance. The first expression on the right hand side is ignored if $\mu_{\mathbf{1}_X} = 0$ implying $\mu_x = 0$ (this will be shown in the next subsection). This result extends to an infinite dimensional $X$. In this case $u_1$ converges to a random variable whose distribution is standard normal conditionally on $v$ and $\mathbf{1}_X$.
(SPANNING) has been assumed so far and is now discarded. Agents cannot influence the part $w^i_{X^\perp}$ of their endowments not spanned by $A$. Thus, if normalized asset returns $y$ :

$$y = \left(A - \frac{1}{\mu_{\mathbf{1}_X}} \mu_A \mathbf{1}_X\right) \left(\sigma^2_{A\backslash 1}\right)^{-\frac{1}{2}},$$

are spherically distributed conditional on $\mathbf{1}_X$ *and* $w^i_{X^\perp}$ for all $i \in \mathbf{I}$ portfolios with equal variance and mean are still perceived as equally good by consumers.
If expected utility shall hold, it is often maintained that the CAPM is only applicable to Gaussian returns, for instance in [(Copeland & Weston, 1979) Ch. 4 G. on p. 85]. But this section clearly characterizes a wider range of admissable distribution functions; nevertheless, this range is rather limited.

### 2.3.3 Indifference Curves and Risk Aversion

This subsection is structured in five parts. In the first the general shape of $\mu/\sigma$–indifference curves are described. The measure of absolute risk aversion is motivated in the second part. Then decreasing absolute risk aversion is discussed, which is generally considered as the realistic behaviour of investors. The next issue is the limiting slope of indifference curves, since it determines whether traders have a point of satiation. In the last part the plausible case of constant relative risk aversion is discussed.

**Shape of Indifference Curves**

Consider the indifference curves of mean/variance preferences in the familiar $\mu/\sigma$–diagram, in which $\sigma$ is associated with the horizontal and $\mu$ with the vertical axis. Because of monotonicity in mean, variance aversion, and quasi–concavity the indifference curves are strictly upward sloping and strictly convex graphs in the $\mu/\sigma$–diagram. The indifference curves start off at the vertical axis and form the boundary of the upper contour set, which is located in the upper left area between the curve and the vertical axis. This allows for a further restriction of the $RER$.

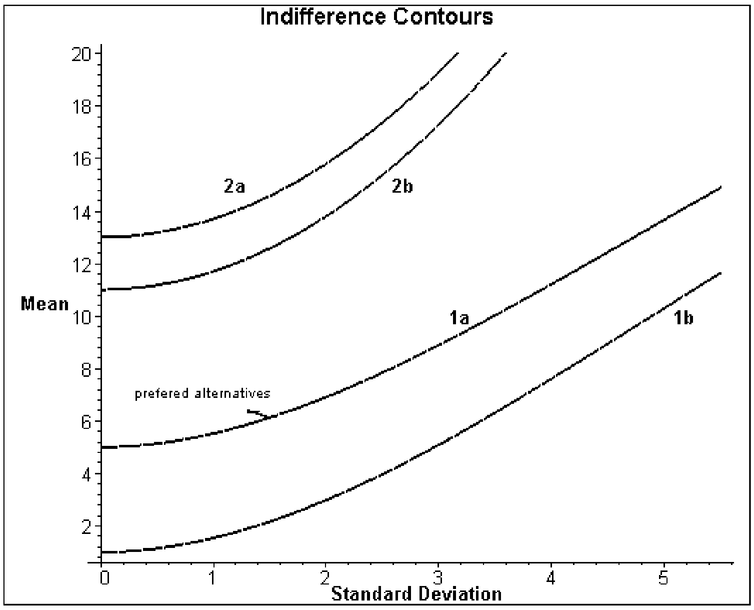


Figure 2: $\mu/\sigma-$Indifference Surfaces

**Lemma 16 (Bounded Consumption)** *In the $RER$ each trader's consumption of mean and standard deviation is bounded.*

**Proof.** Mean is bounded from below by the intercept at which agent $i$'s indifference curve through $(\mu_{w^i}, \sigma_{w^i})$ starts off the vertical axis in the $\mu/\sigma$–diagram. Thus, mean $\mu^i$ is also bounded from infinity, since individuals' $\mu^i$ are bounded from below and sum up to $\mu_{\bar{w}} < \infty$ in $RER$. Obviously, variance is non–negative. Moreover, variance is bounded from above, because the highest acceptable amount for agent $i$ is determined by the indifference curve through $(\mu_{w^i}, \sigma_{w^i})$ at the largest possible mean $\mu_{\bar{w}}$. ■

Quasi–concavity of $\hat{v}$, which is equivalent to the convexity of the upper contour sets (that is to say convexity of preferences), ensures uniqueness of a trader's optimization problem provided that a solution exists and the budget set is weakly convex (see Lemma (46) on p. 63). (Feldstein, 1969) showed Tobin's claim that concavity of $\hat{v}$ follows from a concave vNM–utility function, whenever expected utility implies mean/variance utility. But quasi–concavity in mean and standard deviation is weaker than concavity as well as quasi–concavity in mean and *variance*, since variance is a convex function of standard deviation.

**Lemma 17 (Quasi–Concavity in standard deviation)** *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a monotone increasing function and $v : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ be a monotone decreasing function in the second argument. Consider the conditions:*

*(i) $v$ is strictly quasi–concave in $(\mu, f(\sigma))$.*

*(ii) $v$ is strictly quasi–concave in $(\mu, \sigma)$.*

*Then (i) implies (ii) if $f$ is weakly convex and (ii) implies (i) if $f$ is weakly concave.*

**Proof.** If $v(\hat{\mu}, f(\hat{\sigma})) \geq v(\mu, f(\sigma))$ holds for some distinct $(\mu, \sigma), (\hat{\mu}, \hat{\sigma}) \in \mathbb{R} \times \mathbb{R}_+$ then strict quasi–concavity demands

$$\forall \alpha \in (0,1) : v(\mu, f(\sigma)) < v(\alpha\hat{\mu} + (1-\alpha)\mu, \alpha f(\hat{\sigma}) + (1-\alpha) f(\sigma)).$$

Weak convexity of $f$ in conjunction with variance aversion implies

$$\ldots \leq v(\alpha\hat{\mu} + (1-\alpha)\mu, f(\alpha\hat{\sigma} + (1-\alpha)\sigma)).$$

The second claim can be shown equivalently in the opposite direction. ■

**Measuring Absolute Risk Aversion**

The next aspect drawn to attention is the measure of absolute risk aversion. In the mean/variance world a consumer is a risk averter if he/she always prefers from two consumption plans with equal mean the portfolio with less variance. Since utility functions are strictly decreasing in variance, agents are risk averters in the CAPM. For vNM–utility functions the Arrow–Pratt measure $-\frac{u_{xx}}{u_x}$ is broadly accepted as a risk measure. For the CAPM the coefficient of absolute risk aversion plays this role, which is the tangent slope of the indifference curves in the $\mu/\sigma$–diagram. Since the slope is the marginal rate of substitution of standard deviation for mean, it is the natural choice of a local measure of risk aversion. That is to say, the slope is approximately equal to the amount of additional units in mean (i.e. the riskless asset) for a small additional unit of standard deviation to keep an agent's utility level constant.

**Definition 18 (Risk Aversion)** *Agent i's coefficient of absolute risk aversion $\hat{S}^i$ at $(\mu, \sigma)$ is defined as the slope of his/her indifference curve in the $\mu/\sigma$–diagram at point $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_{++}$ :*

$$\hat{S}^i(\mu, \sigma) := -\frac{\hat{v}_2}{\hat{v}_1} = -\frac{\sigma}{r^i}, \quad \textit{where} \;\; r^i(\mu, \sigma) := \frac{v_1(\mu, \sigma^2)}{2v_2(\mu, \sigma^2)}.$$

Half the marginal rate of substitution of mean for variance, $r^i$, appears in many formulas, why it is used as an abbreviation too. If agents' utility functions satisfy (QUADU) respectively (LINU) their marginal rate of substitution is

$$\begin{aligned} r^i &= \mu_{x^i} - \frac{1}{\varphi^i} && \text{if (QUADU) and} \\ r^i &= -\frac{1}{\varphi^i} && \text{if (LINU), respectively.} \end{aligned}$$

Since $\hat{S}^i$ is regarded to be a *measure* of risk aversion, it should suit for comparisons between the risk aversion of different utility functions. In terms of preference relations agents' risk aversion is comparable in the following way:

**Definition 19** *Investor* 1 *is said to be more risk avers than investor* 2 *if* $\forall x, y \in L, \sigma_{x,y} = 0$ :

$$x \preccurlyeq_1 x + y \quad \textit{implies} \quad x \preccurlyeq_2 x + y,$$

*provided that their mean/variance preferences are* $\preccurlyeq_i, i \in \{1, 2\}$.

This definition states that whenever an agent accepts more variance by $y$, which must be accompanied by more mean due to variance aversion, a less risk avers investor must accept $y$ as well.
Another concept of risk aversion is based on risk premia. If an investor is indifferent between a risky and a riskless consumption plan, the mean of the riskless portfolio is called the safety equivalent. The absolute difference between the means of both consumption streams is the risk premium for taking an additional risk. The generalization of risk premia is as follows.

**Definition 20** *Let* $x, \varepsilon \in L$ *with* $E(\varepsilon | x) = 0$ *and* $Var(\varepsilon | x) > 0$. *Consider a preference relation* $\succsim$ *on* $L$. *The 'compensating risk premium'* $\mu^c_{x,\varepsilon}$ *is defined by*

$$x \sim x + \varepsilon + \mu^c_{x,\varepsilon} \mathbf{1},$$

*whereas the 'equivalent risk premium'* $\mu^c_{x,\varepsilon}$ *is defined by*

$$x - \mu^e_{x,\varepsilon} \mathbf{1} \sim x + \varepsilon.$$

Thus, a risk averter has always positive risk premia, which make him indifferent for taking more risk. Both the risk premia as well as the coefficient of absolute risk aversion cover the notion of higher risk aversion.

**Theorem 21** *Consider two agents with* $\mu/\sigma^2$*–preferences* $\preccurlyeq_1$, $\preccurlyeq_2$ *and the corresponding utility functions* $v^1$, $v^2$. *Then*

1. *Investor* 1 *is more risk avers than investor* 2.

2. *Investor* 1*'s compensating risk premium* $\mu^c_{x,\varepsilon}$ *is never smaller than second investor's.*

3. *Investor* 1*'s equivalent risk premium* $\mu^e_{x,\varepsilon}$ *is never smaller than second investor's.*

4. *First investor's coefficient of absolute risk aversion is never smaller than second investor's.*

5. *For equal endowments investor* 1 *demands never less of the riskless asset than investor* 2*.*

*are equivalent assertions.*

The theorem is borrowed from [(Lajeri & Nielsen, 1994), Proposition 2] and [ (Löffler, 1996a), Ch. 2.4], where the reader is referred to for more detailed results and their proofs. This theorem underpins the reasonability of $\hat{S}^i$ as a risk measure. (Lajeri & Nielsen, 1994) derived the relationship between de–/non–/increasing absolute risk aversion in mean and the equivalent as well as the compensating risk premium. Moreover, they show that risk aversion formulated for vNM–utility carries over to mean/variance utility, if induced preferences coincide on $L$. As their last issue they introduce the concept of 'prudence' or 'precautionary savings' for the CAPM if the utility function is extended for consumption in the investment period.[16]
Quadratic utility functions are criticized for having an increasing absolute risk aversion in mean, i.e. $\hat{S}^i_\mu(\mu, \sigma) > 0$. Thus, an agent with such a utility function is the less willing to bear more risk the more mean he/she consumes. Since mean is a desired good his/her risk aversion increases also in wealth. This is not realistic. Generally it is considered as the natural behaviour that agents are more willing to undertake riskier investments the richer they are. Thus, the more relevant case is non–increasing absolute risk aversion: $\hat{S}^i_\mu(\mu, \sigma) \leq 0$.

[16] Prudence means that agents wish to insure against bad states in the future by purchasing the riskless asset the more the riskier the environment is. Since consumption in the first period is not considered here, this result is only mentioned in passing.

[(Löffler, 1996a), p. 38] has shown that mean/variance preferences with constant absolute risk aversion in mean have a utility representation which is additive separable in mean and variance and linear in mean. Figure (2) at the beginning of this subsection shows two indifference surfaces each for constant ('1a' and '1b') as well as for decreasing absolute risk aversion ('2a' and '2b').

A very interesting axiomatics of $\mu/\sigma^2$–preferences, which is based on decreasing absolute risk aversion, is due to (Epstein, 1985). Epstein's definitions of decreasing absolute risk aversion are stronger in comparison to the definitions proposed by (Pratt, 1964) and (Arrow, 1974), (Ross, 1981), and (Machina, 1982b). His definition relies on the *compensating* risk premium. The distinction between both definitions of risk premia is critical for Epstein's results. Epstein applies a very general definition of "decreasing" absolute risk aversion he calls R–DARA.

**Definition 22** *Let $R$ be a partial pre–order on the space of probability distributions $F_x(\cdot)$ of returns $x \in L$. A preference relation satisfies R–DARA* $:\Leftrightarrow$

$$\mu^c_{x,\varepsilon} \geq \mu^c_{y,\varepsilon} \quad \textit{if} \quad (F_x, F_y) \in R.$$

Under some technical regularity conditions Epstein proves that R–DARA implies $\mu/\sigma^2$–preferences.[17] One of those technical assumptions is incompatible with the independence axiom and thereby with expected utility theory. Constant as well as decreasing absolute risk aversion with respect to the *equivalent* risk premium are shown to imply $\mu/\sigma^2$–linear utility functions. A particular choice for $R$ is that $(F_x, F_y) \in R$ whenever $F_y$ second order stochastically dominates $F_x$. Actually, this choice implies $\hat{S}^i_\mu(\mu, \sigma) \leq 0$, see Theorem 3 in (Epstein, 1985).[18] A particular weakening of this ordering in-

[17](Epstein, 1985) makes use of the "local Bernoulli utility function", which was shown by (Machina, 1982a) to be a local representation for a Fréchet differentiable utility function.

[18]Even without the critical technical conditions the 'first order stochastically dominates'–ordering for $R$ is only compatible with risk neutrality under the expected utility postulate, see (Machina, 1982b).

duces quasi–concavity of the utility function in $(\mu, \sigma^2)$ and $\hat{S}^i_\sigma(\mu, \sigma) \geq 0$, see Theorem 5 ibidem.

**Absolute Risk Aversion in the Limit**

The limiting slope of indifference curves for large mean and standard deviation is important to detect whether a satiating portfolio exists or not (see Section 2.4 for the implications on the pricing functional).

**Definition 23** *Agent i's limiting slope* $s^i : \mathbb{R} \to \mathbb{R}_{++} \cup \{\infty\}$ *is defined by*

$$s^i(\mu_0) = \lim_{\lambda \to \infty} \hat{S}^i(\mu_0 + \lambda, \sigma(\lambda)),$$

*where* $\sigma(\lambda)$ *is the unique solution to* $\hat{v}(\mu_0 + \lambda, \sigma(\lambda)) = \hat{v}(\mu_0, 0)$. *Moreover, let the minimal limiting slope for agent i in RER be*

$$\hat{s}^i = s^i\left(\mu\left(w^i\right)\right),$$

*where* $\mu(w^i)$ *is the unique solution to* $\hat{v}(\mu(w^i), 0) = U(w^i)$.

The limiting slope describes the behaviour of the indifference curves in the far upper right area of the $\mu/\sigma$–diagram. Since mean/variance preferences shall satisfy continuity the absolute risk aversion cannot become infinite for finite mean/variance combinations.[19] If the limiting slope is finite, indifference curves converge from above to a ray for large mean/variance combinations. The limiting slope is non–decreasing in mean, otherwise indifference curves would either cross or would not be convex. This also implies that for non–increasing absolute risk aversion the limiting slope is independent of mean. Figure (2) at the beginning of this subsection shows two indifference surfaces, '2a' and '2b', having decreasing absolute risk aversion and thus a constant limiting slope. From convex analysis the same applies to concave utility functions, see (Nielsen, 1987). Nielsen has shown how the vNM–utility function determines the minimal limiting slope for normally distributed returns.

[19] If the slope is infinite for finite mean/variance combinations there is a pole locus, right from there preferences have to be lexicographic by completeness. That contradicts continuity, but continuity can be relaxed for allocations outside *RER*.

**Relative Risk Aversion**

Relative risk aversion is defined as the expenditure for the riskless asset as a share on total wealth, that is $\mu^i q(\mathbf{1})/q(w_X^i)$ for a pricing functional $q : L \to \mathbb{R}$ and if (BOND) is assumed. A trader's relative risk aversion decreases in wealth if this share decreases in first period income $q(w_X^i)$. Thus, with increasing wealth he/she is willing to bear *relatively* more risk. Empirically, constant relative risk aversion was found to be very plausible.[20] In expected utility theory in–/decreasing/constant relative risk aversion is equivalent to an in–/decreasing/constant coefficient of relative risk aversion, which is defined as $-x\frac{u''(x)}{u'(x)}$. Only the Bernoulli utility functions

$$(1-\alpha)\,x^{1-\alpha}, 0 < \alpha \neq 1 \text{ and } \ln x$$

are compatible with constant relative risk aversion, monotonicity, quasi–concavity and differentiability. Those choices are not applicable to negative returns (non–positive returns for 'ln' and $\alpha > 1$). Since expected utility induces $\mu/\sigma$–utility only if the distribution of returns is centered about the origin, constant relative risk aversion of $\mu/\sigma$–utility functions cannot be founded by expected utility theory, simply because these Bernoulli utility functions are not defined for negative values. Fortunately, $\mu/\sigma$–utility functions need not be derived from expected utility. (Löffler, 1995) shows that non–increasing relative risk aversion in the level of wealth is only compatible with degenerated $\mu/\sigma$–utility functions leading always to optimal portfolios which are riskless.[21] Plausible constant relative risk aversion in conjunction with Löffler's result severely challenges the CAPM. To overcome this 'paradox' (Löffler, 1995) suggested to incorporate wealth – à la Patinkin – in the

[20]See (Friend & Blume, 1975) and the references (Dalal & Arshanapalli, 1993) and (Levy, 1994) cited from (Löffler, 1995).

[21]In his proofs of two important lemmata (Löffler, 1995) claimed that non–increasing relative risk aversion implies $\sigma(\hat{w}_2) \geq \frac{\hat{w}_2}{\hat{w}_1}\sigma(\hat{w}_1)$ for wealth levels $\hat{w}_2 > \hat{w}_1$ and with $\sigma(\hat{w})$ being the optimal standard deviation for monetary wealth level $\hat{w}$ (see Appendix ibidem). This differs from the notion of non–increasing relative risk aversion given in the text. The latter definition requires $\alpha(\hat{w}_2) \geq \alpha(\hat{w}_1)$ for $\hat{w}_2 > \hat{w}_1$ and with $\alpha(\hat{w})$ being the optimal budget share spend on the risky asset $x$. This implies at best $\sigma(\hat{w}_2) = \hat{w}_2\alpha(\hat{w}_2)\sigma_x \geq \hat{w}_1\alpha(\hat{w}_1)\sigma_x = \sigma(\hat{w}_1)$, which is weaker than Löffler's condition for relative risk aversion.

$\mu/\sigma$–utility function. Löffler's result is in contrast to (Epstein, 1985), who established the equivalence between absolute and relative risk aversion, if constant absolute risk aversion is prevalent (see Theorem 6 ibidem). At the end of Section 2.5 this result is extended with a different proof. It is indeed the case that in the CAPM absolute is tantamount to relative risk aversion. Nevertheless, the CAPM is inconsistent with decreasing absolute and constant relative risk aversion thereby, which was empirically found to be very plausible.

## 2.4 Viability of the Price Functional

In this section the possible satiation of investors in the Quasi–Bond and CAPM–related arbitrage are considered. Satiation is troublesome, because it could cause the non–existence of an equilibrium and could make arbitrary price functionals viable. (Nielsen, 1987) pointed out:

> "A LITTLE–RECOGNIZED FEATURE of the mean–variance portfolio–selection model is that induced preferences for asset holdings are not necessarily monotone; more of an asset (or portfolio) is not necessarily better, even if the asset (or portfolio) has positive expected return. ..."

Obviously, more of mean could be outweighed by an increased variance. Indeed the last part could be replaced by: 'even if the risky asset (or portfolio) has just positive pay–offs.' And still more pay–off in every state of the world could be outweighed by a higher variance.
This observation has two implications: Mean/variance preferences are not necessarily monotone in tomorrow's pay–offs and satiation in asset demand is likely to occur. Both are related issues, because non–monotonicity also causes satiation. However, whereas non–monotonicity is a serious problem of coherentness, satiation is – beside of being not representative – 'only' quite troublesome making the calculation of equilibria more difficult. An equilibrium might even not exist, so that one has to switch to the definition of quasi–

equilibria, which are somewhat artificial.[22] This problem is not specific to the CAPM. Also in general equilibrium models with incomplete markets satiation can occur, even if traders have concave, monotone, expected utility functions [see for instance (Mas–Colell, 1992) and (Polemarchakis & Siconolfi, 1993)]. But in contrast to the CAPM it suffices to assume the existence of an asset, which has no negative pay–offs and positive pay–offs with positive probability, since in GEI–models utility functions are monotone in state pay–offs. In the CAPM investors are more likely to be satiated in assets whose ratio of mean per standard deviation is low. Satiation and non–monotone preferences interact in an unpleasant manner, since non–monotonicity enlarges the set of assets, in which consumers could be satiated.

To distinguish between the two related problems, the satiation in state pay–offs is named here as 'the problem of *non–monotonic* utility functions'. And utility functions are said to be monotonic in the CAPM, if traders are not satiated in state pay–offs for allocations in $RER$, even though their utility functions might not be monotone everywhere in $L$. This problem is the main issue of Section 2.9.

If the Quasi–Bond does not satiate an investor it must be infinitely desirable. Thereby the Quasi–Bond could be a free lunch if it has a non–positive price. But even with a positive price for the Quasi–Bond there is a consumption plan which maximizes the ratio between mean and standard deviation subject to the constraint of involving no investments in period zero. Thus, it is a sufficient condition for ruling out arbitrage if for all consumers this particular asset is not infinitely desirable for all allocations in $RER$.

Both satiation and arbitrage must be ruled out for a price system to be regarded as viable in an equilibrium.

**Satiation**

The natural way is to look at satiation first. Satiation appears if traders cannot find utility enhancing opportunities in the marketed subspace. In

[22] In quasi–equilibria aggregated consumption must be not larger than aggregated endowments in contrast to the equality of both demanded in the definition of CAPM–equilibria.

case of (BOND) the riskless asset always improves utility, whereas without a riskless asset the Quasi–Bond is traders' first choice prior to a possible point of satiation. But there might exist portfolios traders are never satiated in. The following definitions distinguish between the two possibilities.

**Definition 24 (Unlimited Improvement)** *The consumption plan $y \in X$ is called an* 'unlimited improvement for agent $i$' *if there is a $x^i \in RER$ such that for all positive real numbers $\alpha_2 > \alpha_1$ :*

$$x^i + \alpha_2 y \succ_i x^i + \alpha_1 y.$$

**Definition 25 (Satiation)** *A consumption plan $y \in X$ is said to be* 'agent $i$'s satiation portfolio' *if for all $x \in X$ :*

$$y + w^i_{X^\perp} \succsim_i x + w^i_{X^\perp}$$

*holds. Investor $i$ is locally not satiated / satiated / over satiated in asset $y \in X$ at $x \in L$ if*

$$\left. \frac{\partial}{\partial \lambda} U^i \left( x + \lambda y \right) \right|_{\lambda=0} \gtreqless 0.$$

Both a satiation portfolio as well as an unlimited improvement can exist for a consumer. This happens if the utility level of an infinite consumption of an unlimited improvement is still less than the utility obtained from the satiation portfolio.

**Lemma 26 (Unlimited Improvements)** *It is a sufficient condition for $y \in L$ to be an unlimited improvement for agent $i$ that*

$$\mu_y > 0 \quad and \quad \left( \mu_y \geq \sigma_y \hat{s}^i \quad or \quad \sigma_y = 0 \right).$$

*The condition is also necessary if the limiting slope $s^i(\mu)$ is constant in $\mu$.*

**Proof.** The bond is obviously an unlimited improvement. And it is the only one if $\hat{s}^i = \infty$. Thus, only the case $\sigma_y > 0$ and $\hat{s}^i < \infty$ deserves attention.

The derivative of the utility function $U(x+\alpha y)$ for $\alpha$, i.e. in direction of an unlimited improvement $y \in X$, must be positive for a sequence of alphas converging to infinity:

$$\begin{aligned} v_1\mu_y + 2v_2\left(\alpha\sigma_y^2 + \sigma_{xy}\right) &> 0 \Leftrightarrow \\ \hat{S}^i \frac{\alpha\sigma_y^2 + \sigma_{xy}}{\sigma_y\sqrt{\alpha^2\sigma_y^2 + 2\alpha\sigma_{xy} + \sigma_x^2}} &< \frac{\mu_y}{\sigma_y}. \end{aligned}$$

The fraction on the left hand side converges to one from below with increasing $\alpha$, whereas $\hat{S}^i$ converges to $s^i(\mu)$ for large $\mu$. Only if $s^i(\mu) \geq \hat{s}^i$ the improvement is relevant in $RER$, otherwise being worse than endowments. Thus $\hat{s}^i\sigma_y \leq \mu_y$ is sufficient. If $s^i(\mu)$ is constant the parabola implied by adding $\alpha y$ to $x$ would cross indifference curves always from below in the $\mu/\sigma$–diagram because of its steeper slope whenever the condition $\mu_y \geq \sigma_y\hat{s}^i$ holds. Thereby for large $\alpha$ more of $y$ is always utility enhancing, what implies the necessity of this condition. ■

An asset may only be infinitely desirable by agent $i$ if its ratio of mean per standard deviation is larger than $i$'s minimal limiting slope; otherwise its indifference curves definitely cross from below the rays with slope equal to asset's $\frac{\mu_y}{\sigma_y}$–ratio starting at any point in $RER$.

The riskless asset is obviously an unlimited improvement. However, the ratio of standard deviation per mean is minimized in $X$ by the Quasi–Bond. Thus, if the Quasi–Bond is not an unlimited improvement no other consumption plan could be, which justifies the name 'Quasi–Bond'. The following lemma and its corollary neatly characterize the properties of the Quasi–Bond (note that $M_A$ is the matrix of second moments of assets in $A$).

**Lemma 27 (Projection of 1 on X)** *The orthogonal projection* $\mathbf{1}_X$ *of* $\mathbf{1}$ *on* $X$ *is:*

$$\mathbf{1}_X = AM_A^{-1}\mu_A.$$

*The mean of the Quasi–Bond satisfies*

$$0 \leq \mu_{\mathbf{1}_X} = \mu_A^T M_A^{-1} \mu_A \leq 1.$$

*The inequality on one of either side holds only with equality either if $\mu_A = 0$ or if $X$ satisfies (BOND), respectively. The Quasi–Bond has variance*

$$\mu_{\mathbf{1}_X} \left(1 - \mu_{\mathbf{1}_X}\right).$$

**Proof.** The orthogonal projection $\mathbf{1}_X$ of $\mathbf{1}$ on $X$ is the solution of the following minimization problem: $\min_{\mathbf{1}_X \in X} ||\mathbf{1} - \mathbf{1}_X||_P$. Minimizing $||\mathbf{1} - \mathbf{1}_X||_P = ||\mathbf{1} - A\theta_{\mathbf{1}_X}||_P = \mathbf{1} - 2\mu_A^T \theta_{\mathbf{1}_X} + \theta_{\mathbf{1}_X}^T M \theta_{\mathbf{1}_X}$, results in $\theta_{\mathbf{1}_X} = M_A^{-1}\mu_A$ and thus $\mathbf{1}_X = AM^{-1}\mu_A$. Note that if (BOND) the first column and the first row in $M_A$ are equal to $\mu_A$. This implies $\mathbf{1}_X = \mathbf{1}$ if (BOND) and $\mathbf{1}_X = 0$ if $\mu_A = 0$, respectively.
Observe that the covariance matrix $\sigma_A^2 = M_A - \mu_A \mu_A^T$ is positive (semi–) definite, which implies $M_A \succ (\succeq)\ \mu_A \mu_A^T \succeq 0$ so that $\zeta(\mu_A \mu_A^T M_A^{-1}) \leq 1$ holds, where $\zeta$ denotes the spectral norm.[23] Since $\text{rank}[\mu_A \mu_A^T M_A^{-1}] \leq 1$ at best one eigenvalue could be non–negative: $0 \leq \zeta(\mu_A \mu_A^T M_A^{-1}) = \text{trace}(\mu_A \mu_A^T M_A^{-1}) = \mu_A^T M_A^{-1} \mu_A = \mu_{\mathbf{1}_X} \leq 1$.
Since $M_A$ is positive definite $\mu_{\mathbf{1}_X}$ is only zero if $\mu_A = 0$. If (BOND) does not hold $\sigma_A^2$ is positive definite, which implies $M_A \succ \mu_A \mu_A^T \succeq 0$ so that $\zeta(\mu_A \mu_A^T M_A^{-1}) = \mu_{\mathbf{1}_X} < 1$ holds.
Finally, $E\left(\mathbf{1}_X^2\right) - \mu_{\mathbf{1}_X}^2 = \mu_{\mathbf{1}_X}\left(1 - \mu_{\mathbf{1}_X}\right)$ yields the variance. ■

**Corollary 28 (Quasi–Bond)** *The Quasi–Bond is the asset with the smallest standard deviation per mean. Any other asset out of $X$ with positive / zero / negative mean has positive / zero / negative covariance with it. The Quasi–Bond is the only asset with this property.*

**Proof.** For fixed mean the minimization problem in the preceding lemma shows the minimal variance of $\mathbf{1}_X$. Now the covariance with $z \in X$ reads:

[23] Only in this proof the ordering $\prec$ on positive semi–definite matrices is applied. See Section 7.7 on p. 469ff and Theorem 7.7.3 on p. 471 in (Horn & Johnson, 1985) about the definition of the ordering and for the result concerning the spectral radius, respectively.

$\sigma_{z,\mathbf{1}_X} = E(z\mathbf{1}_X) - \mu_z\mu_{\mathbf{1}_X} = \mu_z(1-\mu_{\mathbf{1}_X})$.
For any other asset $y \in X : y \nparallel \mathbf{1}_X, \mu_y > 0$ there are numbers $\mu_z$ and $\alpha_z$ such that the covariance between $y$ and the asset $z := \mu_z \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} + \alpha_z \left(y - \mu_y \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right)$ has a different sign than $\mu_z$. ■

Note that assets orthogonal to $\mathbf{1}_X$ have zero mean. Hence, no other asset improves the Quasi–Bond in variance **and** mean. The corollary shows that a satiation portfolio can only consist of the Quasi–Bond. Also local satiation makes sense only with regard to the Quasi–Bond and can be expressed in terms of $r^i$.

**Lemma 29** *Agent $i$ is locally not satiated / satiated / over satiated in the Quasi–Bond at $x^i$ if*

$$\left(r^i - \mu_{x^i}\right)\mu_{\mathbf{1}_X} + \mu_{x^i_X} \lesseqgtr 0.$$

**Proof.** A trader is marginally not satiated at $x^i$ if the derivative of his/her utility in the direction of $\mathbf{1}_X$ at $x^i$ is positive:

$$v_1^i \mu_{\mathbf{1}_X} + 2Cov\left(\mathbf{1}_X, x^i\right) v_2^i > 0.$$

With $Cov\left(\mathbf{1}_X, x^i\right) = \mu_{x^i_X} - \mu_{\mathbf{1}_X}\mu_{x^i}$ the assertion follows. ■

Since the analysis of financial markets is artificial if there is overall satiation in the economy the following assumption is imposed.

**(NonSat)** For all allocations in $RER$ consumers are not over satiated while at least one agent is not satiated in the Quasi–Bond.

Only if the Quasi–Bond does not satiate an agent, i.e. if it is an unlimited improvement to them due to $\sqrt{\mu_{\mathbf{1}_X}} \geq \sqrt{1-\mu_{\mathbf{1}_X}}\hat{s}^i$, there probably exist also other, less utility enhancing, but nevertheless unlimited improvements (just by continuity). Whenever $s^i(\mu)$ depends on $\mu$ it is not straightforward to characterize the set of unlimited improvements, because those depend on the starting portfolio.[24]

[24] An unlimited improvement may well exist for a particular starting portfolio while it never reaches the utility level of the satiating portfolio.

**Viability of the Price Functional**

Following it is intended to define the viability of the price functional for the CAPM. To formalize this argument the price correspondence has to be defined first:

**Definition 30 (Price Functional)** $q : X \twoheadrightarrow \mathbb{R}$ *is the correspondence, which assigns to every traded consumption bundle a set of prices.*

In the CAPM arbitrage shall occur whenever a consumption stream, which is infinitely desirable, has a non–positive price. This is weaker than in the GEI–model, because unlimited improvements can also incorporate disliked components, namely positive variance. From the GEI point of view only the bond is a possible arbitrage portfolio in the CAPM. Nevertheless, in terms of utility, a free unlimited improvement is still a free lunch.[25]

**Definition 31 (CAPM–Arbitrage)** *In a CAPM–economy the price system q permits arbitrage if a trader can purchase an unlimited improvement for a non–positive price.*

Viability of a price system means that it is not a priori incompatible with an equilibrium of an economy with (AGENTS). But for agents satiated by endowments every price system is viable. Therefore viability is defined for a particular choice of (AGENTS), namely not saturated consumers.

**Definition 32 (CAPM–Viability)** *In a CAPM–economy with (NonSat) a price system is called viable if it does not permit arbitrage and if the Quasi–Bond has a positive price.*

[25] An unlimited improvement with a non–positive price offers an unlimited arbitrage opportunity. However contrary to unlimited arbitrage, local arbitrage is defined as a would–be opportunity which cannot be exploited because of trading constrains. In the General Equilibrium Theory of Incomplete Markets with trading constraints there is a very clean characterization of the implications local and no–unlimited arbitrage have on the price functional, see (Elsinger & Summer, 1998). However, trading constraints are not considered here.

The viability condition is a minimal requirement for a price system to be regarded as a possible candidate for sustaining an equilibrium.[26] The implication that the no–arbitrage condition has on the price system can be considerably tightened. This will be explored at the end of this section. Viability is stronger than no–arbitrage and will mainly be considered from now on. The definition suffices to show the linearity of the price functional.

**Lemma 33 (Linear pricing functional)** *Suppose at least one asset has a non–zero mean and (NonSat) holds. The pricing correspondence $q$ is viable only if $q$ is a continuous linear functional on $X$ with the representation:*

$$\forall x \in X : q(x) = \pi \bullet x$$

*for a unique $\pi \in X$ with positive mean.*

Due to its importance in the subsequent analysis $\pi$ deserves its own definition.

**Definition 34** *$\pi$ is called the **pricing asset**.*

**Proof.** Note that $q$ must be uniquely determined. If not, the Quasi–Bond could be purchased at zero costs, contradicting viability. Thus, the price of every consumption stream is the price of the replicating portfolio:

$$\forall x \in X : q(x) = q\theta_x,$$

which is a continuous, linear functional on $X$, since $X$ is closed with respect to the relative topology of $L$.
Since $P$ is $\mathcal{F}$–finite there is – up to a set of $P$–measure zero – a unique $\pi \in X$ by Riesz' Representation Theorem [see for instance (Billingsley, 1995) on p. 244] such that $q(x) = \pi \bullet x$ for all $x \in X$. By viability the price of the non–zero Quasi–Bond is positive: $\pi \bullet \mathbf{1}_X = \mu_\pi > 0$. ■

[26] In the GEI a price system is called viable if it can support an equilibrium for an economy with rational investors. It is not intended to give a similar definition of viability for the CAPM. The only conclusion would be a positive price for the Quasi–Bond because arbitrage can always be ruled out by a particular choice of (AGENTS), see the respective result below.

The price functional determines when agents are saturated in the Quasi–Bond:

**Lemma 35** *Suppose the Quasi–Bond satiates trader $i$ at some amount of $\alpha^{i*}$. Then he/she is able to purchase his/her satiating portfolio $x^i(\alpha^{i*}) = \alpha^{i*}\mathbf{1}_X + w^i_{X^\perp}$ if*

$$\left(\mu_{w^i_{X^\perp}} - r^i \circ x^i\left(\alpha^{i*}\right)\right)\mu_\pi \leq q(w^i_X)(1-\mu_{\mathbf{1}_X}), \qquad \text{(CondSat)}$$

*where $\alpha^{i*}$ has to be the fixed point of:*

$$\left(1-\mu_{\mathbf{1}_X}\right)\alpha^{i*} = \mu_{w^i_{X^\perp}} - r^i \circ x^i\left(\alpha^{i*}\right).$$

**Proof.** The point of satiation is the solution to

$$\begin{aligned}
\max_{\alpha\in\mathbb{R}} v^i(\alpha\mu_{\mathbf{1}_X} + \mu_{w^i_{X^\perp}}, Var(\alpha\mathbf{1}_X + w^i_{X^\perp})) &\Leftrightarrow \\
v^i_1\mu_{\mathbf{1}_X} + 2\alpha\sigma^2_{\mathbf{1}_X}v^i_2 + 2Cov\left(\mathbf{1}_X, w^i_{X^\perp}\right)v^i_2 &= \\
v^i_1\mu_{\mathbf{1}_X} + 2\alpha(1-\mu_{\mathbf{1}_X})\mu_{\mathbf{1}_X}v^i_2 - 2\mu_{\mathbf{1}_X}\mu_{w^i_{X^\perp}}v^i_2 &= 0,
\end{aligned}$$

which results in the solution for $\alpha^{i*}$. Pricing the satiation portfolio yields the inequality $q\left(\alpha^{i*}\mathbf{1}_X\right) = \alpha^{i*}\mu_\pi \leq q(w^i_X)$. ■

This condition of satiation is endogenous since it involves the pricing asset $\pi$ being actually part of an equilibrium. Without further knowledge about $\pi$ it is not obvious how to rule out satiation just by an *exogenous* and *explicit* condition. Note that (NonSat) is an exogenous but *abstract* condition on (AGENTS). With a solution for the equilibrium pricing asset the condition for non–satiation will be refined in Section 2.6.

It is now intended to derive the unique, freely available asset with the largest ratio of mean per standard deviation. For that the next definition in conjunction with the following lemma simplifies the formulas to be calculated afterwards. Both generalize the *centered* second (cross–)moment (or pseudo inner product), (co–)variance, onto a marketed subspace without a riskless asset, since the mean–correction by the bond is not feasible.

**Definition 36** *The X–covariance of $z, y \in X$ is defined by*

$$Cov_X(z, y) := \mu_{\mathbf{1}_X}(z \bullet y) - \mu_z \mu_y.$$

*Likewise, the X–variance of $z \in X$ is $Var_X(z) := Cov_X(z, z)$ and the X–correlation of $z, y \in X$ is $Cov_X(z, y) / \sqrt{Var_X(z) Var_X(y)}$.*

**Lemma 37 (Non–negative X–variance)** *The X–variance $Var_X(z)$ is non–negative for all $z \in X$. It is zero only if $z \parallel \mathbf{1}_X$ or $\mu_{\mathbf{1}_X} = 0$. $Cov_X(\mathbf{1}_X, z)$ is zero for all $z \in X$. Moreover, if $\mu_{\mathbf{1}_X} \neq 0$ the following equation holds for all $z \in X$ :*

$$Var\left(z - \mu_z \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right) = \frac{1}{\mu_{\mathbf{1}_X}} Var_X(z).$$

**Proof.** By $Cov_X(\mathbf{1}_X, z) = \mu_{\mathbf{1}_X}(\mathbf{1}_X \bullet z) - \mu_{\mathbf{1}_X}\mu_z = 0$ the last but one assertion holds. If $\mu_{\mathbf{1}_X} = 0$ all assets have zero mean which implies zero $X$–variance. Suppose, $\mu_{\mathbf{1}_X} \neq 0$. Then

$$\begin{aligned} Var\left(z - \mu_z \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right) &= E\left(z^2\right) - 2\frac{\mu_z^2}{\mu_{\mathbf{1}_X}} + \frac{\mu_z^2}{\mu_{\mathbf{1}_X}^2}\mu_{\mathbf{1}_X} \\ &= \frac{1}{\mu_{\mathbf{1}_X}} Var_X(z) \\ &\geq 0, \end{aligned}$$

is only zero if $z \parallel \mathbf{1}_X$. ■

Even if the Quasi–Bond has a positive price, free unlimited improvements could exist. To characterize the best of those, the asset with the highest ratio of mean per standard deviation which has additionally a zero price is determined by the following lemma.

**Lemma 38** *Given that the price of the Quasi–Bond is positive and $\pi \nparallel \mathbf{1}_X$,*

$$\chi(\pi) = \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} - \frac{\mu_\pi}{Var_X(\pi)}\left(\pi - \mu_\pi \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right)$$

*is the asset with the smallest ratio of standard deviation per mean which is* *purchasable at zero costs and has a mean of one. This ratio is*

$$\sqrt{\frac{q(\pi)}{Var_X(\pi)} - 1}.$$

**Proof.** Consumption streams in $\langle \pi, \mathbf{1}_X \rangle^{\perp} \cap X$ have a zero mean and a zero price. They just add additional variance and are therefore not desirable to achieve minimal standard deviation per mean. Thus, desirable consumption plans can only have the following representation

$$y = \alpha_y \left( \pi - \mu_\pi \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} \right) + \mu_y \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}$$

for a positive mean $\mu_y$ and a real $\alpha_y$. The Quasi–Bond has the highest ratio of mean per standard deviation but a positive price. The first asset on the right hand side has zero mean and a positive price. It is used to make $y$ having a zero price at minimal variance:

$$\begin{aligned} q(y) &= \alpha_y \frac{Var_X(\pi)}{\mu_{\mathbf{1}_X}} + \mu_y \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} = 0 \Leftrightarrow \\ \alpha_y &= -\mu_y \frac{\mu_\pi}{Var_X(\pi)}. \end{aligned}$$

This yields the asset given in the Lemma. Then $y$'s variance is given by

$$\sigma_y^2 = \mu_y^2 \left[ \frac{q(\pi)}{Var_X(\pi)} - 1 \right].$$

■

Since the asset $\chi(\pi)$ is important in agents' portfolio choice and in detecting satiation it deserves its own definition:

**Definition 39** *The consumption stream $\chi(\pi)$ defined in Lemma 38 is called the **'zero–cost efficient asset'**.*

If there are free lunches in the market the zero–cost efficient asset would be the most desirable one. Thus for ruling out arbitrage opportunities it is sufficient that $\chi(\pi)$ is not an unlimited improvement. The following corollary provides a sharper condition for the viability of $\pi$. The condition should hold even if there is a riskless asset.

**Corollary 40** *In a CAPM–economy with (NonSat) and where $\pi \nparallel \mathbf{1}_X \neq 0$ holds a sufficient condition for the viability of the pricing asset $\pi$ is*

$$\mu_\pi > 0 \quad and \quad \left(\frac{q(\pi)}{Var_X(\pi)} - 1\right)^{-\frac{1}{2}} \leq \min_{i \in \mathbf{I}} \hat{s}^i. \qquad \text{(Viability)}$$

*Both conditions are also necessary if the limiting slopes are independent of mean, i.e. if $s^i(\mu) \equiv \hat{s}^i$ for all $i \in \mathbf{I}$.*

**Proof.** $\mu_\pi > 0$ rules out arbitrage and free satiation in the Quasi–Bond. The second condition rules out arbitrage in $\chi(\pi)$ as well as arbitrage in other unlimited improvements, since those are never efficient at zero costs like $\chi(\pi)$ is. Whereas the first condition is always necessary, the latter one is necessary if the limiting slopes do not depend on agents' starting portfolios. ■

Viability will be reconsidered when in Section 2.6 it is solved for the equilibrium properties. Therewith one is able to tighten the conditions for local satiation and arbitrage moreover.

## 2.5 Portfolio Selection

In this section traders' optimal portfolio choice is derived by proceeding in the following way. First, the Capital Market Line is calculated. Then it is shown that the first order conditions for traders' choice problem are also sufficient for a unique optimum. The Mutual Fund Theorem is then a straight forward implication. The corollary establishing the equivalence between absolute and relative risk aversion concludes this section.

Some calculations are investigated to determine the Capital Market Line in terms of minimal standard deviation for a level of mean $\mu$ to be attained and a given budget $\bar{b} = q(w_X^i)$. Since an investor can achieve the wanted mean without spending the whole budget the indicator of satiation is defined first:

**Definition 41** *Agent i's indicator function of satiation is denoted by*

$$\Upsilon^i \equiv \Upsilon\left(\mu^i, q\left(w^i_X\right)\right) := \begin{cases} 1 & \text{if } \mu^i \mu_\pi > q\left(w^i_X\right) \mu_{\mathbf{1}_X} \quad \text{and} \quad \pi \nparallel \mathbf{1}_X \\ 0 & \text{otherwise.} \end{cases}$$

Traders' decision problem sounds like '*picking their favorite pair out of the efficient mean/variance combinations they can afford within their budget'.* The Capital Market line characterizes this efficient set.

**Lemma 42 (Capital Market Line)** *If a $z \in X$ exists satisfying*

$$\mu_z = \mu \quad \text{and} \quad q(z) \leq \bar{b}$$

*the optimization problem:*

$$\min_{x \in X} Var(x + y) \quad \text{with respect to}$$
$$\mu_x = \mu \quad \text{and} \quad q(x) \leq \bar{b}, \ \text{for some} \quad y \in X^\perp$$

*has a unique solution $\hat{x}\left(\mu, \bar{b}\right)$ in $\langle \mathbf{1}_X, \pi \rangle$. The standard deviation of the solution is a weakly convex, continuously differentiable function of the parameters $\mu$ and $\bar{b}$. If $\pi \nparallel \mathbf{1}_X \neq \mathbf{0}$ the solution sounds*

$$\hat{x}\left(\mu, \bar{b}\right) = \begin{cases} \frac{\mu}{\mu_{\mathbf{1}_X}} \mathbf{1}_X & \text{if } \Upsilon\left(\mu, \bar{b}\right) = 0, \\ \frac{\bar{b}}{\mu_\pi} \mathbf{1}_X + \left(\frac{\bar{b}\mu_{\mathbf{1}_X}}{\mu_\pi} - \mu\right) \chi(\pi) & \text{if } \Upsilon\left(\mu, \bar{b}\right) = 1. \end{cases}$$

*The minimum variance is given by*

$$\hat{\sigma}^2\left(\mu, \bar{b}, y\right) = \Upsilon\left(\mu, \bar{b}\right) \frac{\left(\bar{b} - \mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right)^2 \mu_{\mathbf{1}_X}}{Var_X(\pi)} + \mu^2 \frac{1 - \mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu\mu_y + \sigma_y^2.$$

**Proof.** See Appendix 2.11.2, also for some degenerated cases not mentioned here and non–existence of solutions. ■

Note that $\sigma_y$ and $\mu_y$ must satisfy the constraint

$$\sigma_y^2 \geq \frac{\mu_{\mathbf{1}_X}}{1-\mu_{\mathbf{1}_X}}\mu_y^2 \overset{\mathbf{1}_X \to \mathbf{1}}{\longrightarrow} 0.$$

The Capital Market Line consists only of efficient $\mu/\sigma-$combinations:

**Definition 43** *The* Capital Market Line *for investor i given their budget* $q\left(w_X^i\right)$ *and their non–spanned endowments* $w_{X^\perp}^i$ *is the upper branch of the graph*

$$\left(\hat{\sigma}\left(\mu, q\left(w_X^i\right), w_{X^\perp}^i\right), \mu + \mu_{w_{X^\perp}^i}\right)$$

*for* $\mu \geq \frac{q\left(w_X^i\right)}{\mu_\pi}$ *if (BOND) holds and for* $\mu \geq \frac{\mu_{\mathbf{1}_X}}{1-\mu_{\mathbf{1}_X}}\mu_{w_{X^\perp}^i}$ *otherwise.*

Standard deviation, $\hat{\sigma}\left(\mu, \bar{b}, y\right)$, is a weakly convex, continuously differentiable 'parabola' in $\left(\mu, \bar{b}\right)$. It is constant in income if $\Upsilon\left(\mu, \bar{b}\right) = 0$ and decreasing if $\Upsilon\left(\mu, \bar{b}\right) = 1$. Its cross–derivative in $\left(\mu, \bar{b}\right)$ is negative for $\Upsilon\left(\mu, \bar{b}\right) = 1$, and its derivative with respect to $\mu$ is positive at the efficient, upper part of the parabola. Second derivatives do not exist only at $\mu = \bar{b}\mu_{\mathbf{1}_X} / \mu_\pi$. As long as the budget is not fully exhausted, i.e. if $\Upsilon\left(\mu, \bar{b}\right) = 0$ , it is optimal to buy only the best feasible asset $\frac{\mu}{\mu_{\mathbf{1}_X}}\mathbf{1}_X$. But if the investor is running out of money the gap: $\bar{b}\mu_{\mathbf{1}_X} / \mu_\pi - \mu > 0$, has to be 'financed' by purchasing exactly this amount of the zero–cost efficient asset $\chi(\pi)$. This is the most convenient way to finance the gap in the demand for mean, namely at the lowest ratio of standard deviation per mean and at zero additional costs.
If $\pi \parallel \mathbf{1}_X$ the zero–cost efficient asset is zero, $\chi(\pi) = \mathbf{0}$, which implies that only the Quasi–Bond and the non–spanned endowments are consumed. Another degenerated case is $\mathbf{1}_X = \mathbf{0}$, for which it is optimal to purchase the pricing asset only to balance the budget. Both cases are not considered in the sequel, but are mentioned in Appendix 2.11.2.
The Efficient Frontier does not involve the riskless asset. Thus, if there is no riskless asset, the Capital Market Line and the Efficient Frontier coincide. If (BOND) holds, the lemma describes only the Capital Market Line.

To recover from Lemma 42 the Efficient Frontier with (BOND), one has to interpret the Quasi–Bond as the projection of the Bond onto the marketed subspace *without* the Bond, see Figure (4) of the following example.

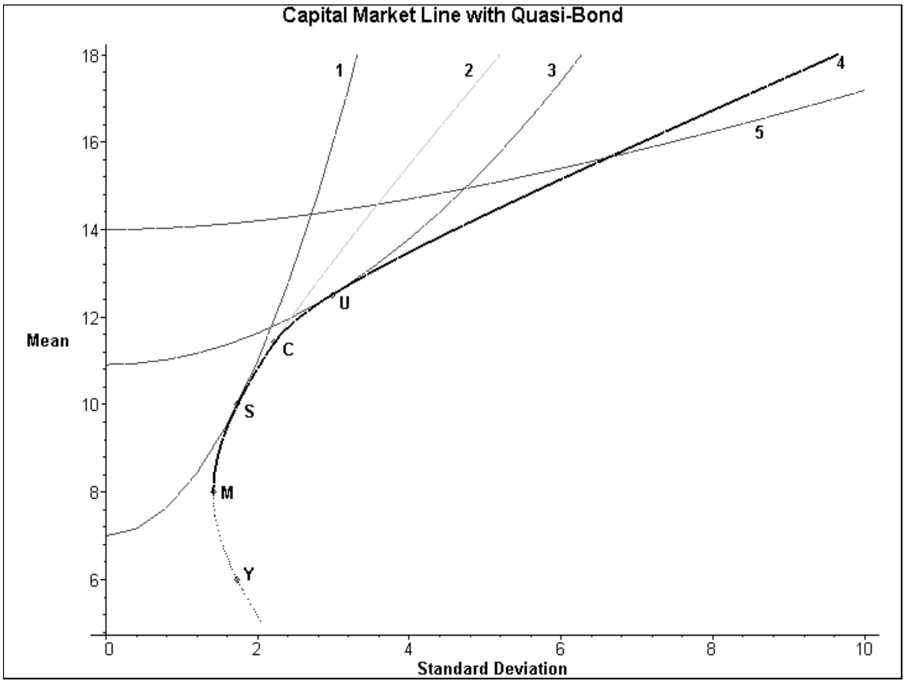


Figure 3: Capital Market Line with Quasi–Bond

**Example 44** *Consider the Figure (3). Curve '4' is the Capital Market Line as well as the Efficient Frontier, since a riskless asset does not exist. It starts in 'Y', where no money is spent at all. At this point an agent consumes only his/her non–spanned endowments. The dotted lower part of the curve represents just inefficient portfolios. The inefficient part ends at the minimum variance portfolio, denoted by point 'M'. Since non–spanned endowments with positive mean have negative correlation with the Quasi–Bond the investor reaches point 'M' by buying an amount of* $\frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}}\mu_{w^i_{X^\perp}}$ *of the Quasi–Bond. At point 'C' the whole budget is spent for the Quasi–Bond. Between 'M' and 'C' only the Quasi–Bond is purchased but the budget is not fully exploited.*

*Thus, for a preference relation represented by curve '1' an agent would be satiated with a portfolio yielding a* $\mu/\sigma-$*combination represented by point 'S'. By assuming Non–Sat, this situation cannot happen if at the same time markets clear. If a trader's budget was infinite he/she could choose mean/ variance combinations on curve '2', which then has the highest* $\mu/\sigma-$*ratio of achievable portfolios, i.e. containing only the Quasi–Bond beside non–spanned endowments. Increasing mean at point 'C' demands for purchasing the second best asset,* $\chi(\pi)$, *which has the highest achievable* $\mu/\sigma-$*ratio at zero cost. Thus, from point 'C' onwards on the Efficient Frontier the budget is exhausted and* $\chi(\pi)$ *worsens the* $\mu/\sigma-$*ratio of the portfolio. This causes the kink in the curvature of the Capital Market Line at point 'C'. A consumer with indifference curve '3' would be best of with a portfolio yielding the* $\mu/\sigma-$ *combination of point 'U', whereas an individual with indifference curve '5' has no optimal choice at all, since his/her limiting slope is less than the* $\mu/\sigma-$ *ratio of asset* $\chi(\pi)$.

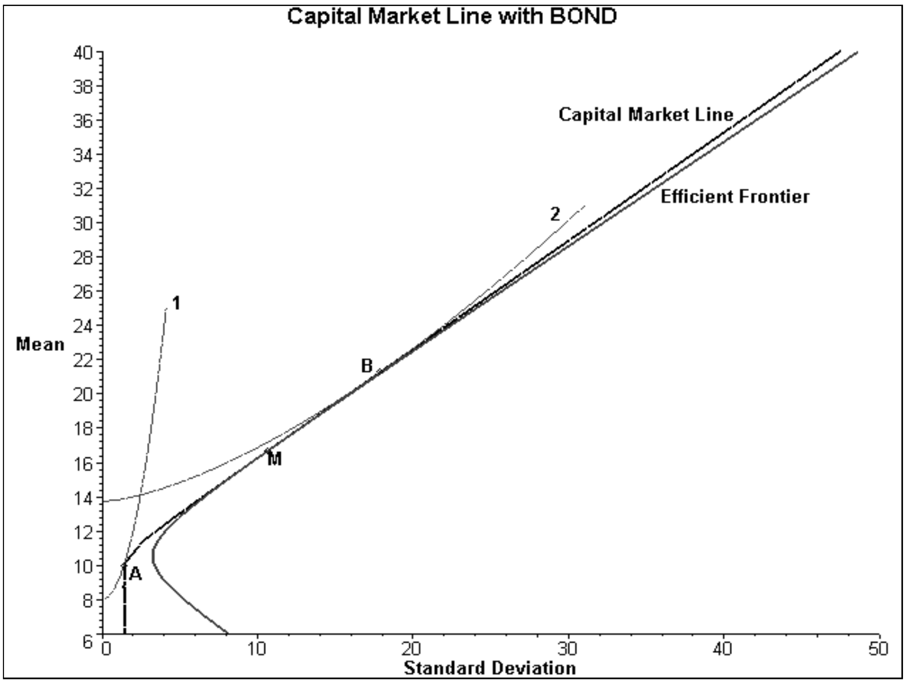


*Figure 4: Capital Market Line with Bond*

*If a riskless asset is introduced the situation is quite different, see Figure (4). The Efficient Frontier and the Capital Market Line do not coincide. Moreover, the Efficient Frontier exists only separately if the first asset is really a bond and other assets admit a non–zero mean. Then it is defined as the set of variance–efficient $\mu/\sigma-$combinations relative to the starting portfolio 'M' without incorporating the bond and any additional costs. Since the Capital Market Line is always more efficient than the 'Efficient Frontier' the latter does not play that important role. If non–spanned endowments have positive variance, like in the case considered here, the Capital Market Line is strictly concave. In the upper right corner of the graph it converges to a ray whose slope is equal to the $\mu/\sigma-$ratio of $\chi(\pi)$. At point 'A' the complete budget is spent on purchasing the bond. An agent with indifference curve '1' would stick at this point as an optimal choice. After that point agents – being not as variance–averse as demonstrated by agent '1' – demand the zero–cost efficient asset $\chi(\pi)$, which simplifies to $\mathbf{1} - \frac{\mu_\pi}{Var(\pi)}(\pi - \mu_\pi \mathbf{1})$ in case of (BOND). An agent with indifference curve '2' would demand the $\mu/\sigma-$combination represented by point 'B'.*

A slight modification of Condition (CondSat) given in Lemma 35, which rules out satiation, has also a geometric interpretation, which is even more general. Consider the point where the whole budget is spent on purchasing the Quasi–Bond and the indifference curve through that point (corresponding to point 'C' in Figure (3) and point 'A' in Figure (4)). If the slope of the indifference curve is smaller than the slope of the Capital Market Line at this point (taking the derivative from the right in case of (BOND)) $\mu/\sigma-$combinations right from that point are preferable. Thus, satiation in the Quasi–Bond as well as satisfaction with a riskless portfolio are ruled out.

**Corollary 45** *Given that $\pi \nparallel \mathbf{1}_X \neq 0$ and Condition (Viability) hold the* *following condition:*

$$\left(\mu_{w^i_{X^\perp}} - r^i\right)\mu_\pi > q(w^i_X)(1-\mu_{\mathbf{1}_X}) \quad at \tag{NonSat}$$

$$\left(\frac{q(w^i_X)}{\mu_\pi}\mu_{\mathbf{1}_X} + \mu_{w^i_{X^\perp}}, \hat{\sigma}^2\left(\frac{q(w^i_X)}{\mu_\pi}\mu_{\mathbf{1}_X}, q(w^i_X), w^i_{X^\perp}\right)\right)$$

*ensures that investor $i$ is neither satiated in the Quasi–Bond nor satisfied with a riskless portfolio in $X$.*

**Proof.** Non–satiation in the Quasi–Bond and dissatisfaction with a riskless portfolio is equivalent to a Capital Market Line having a larger slope than the indifference curve at the point where the entire budget is spent on purchasing the Quasi–Bond. This particular $\mu/\sigma$–pair is given in the second row of the Condition (NonSat). The slope of the Capital Market Line is equal to:

$$-\frac{2\sigma}{-\sigma^2_\mu} = \frac{2\sigma}{2\hat{\mu}\frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu_{w^i_{X^\perp}}} = \frac{2\sigma}{2q(w^i_X)\frac{1-\mu_{\mathbf{1}_X}}{\mu_\pi} - 2\mu_{w^i_{X^\perp}}}.$$

And the marginal rate of substitution of standard deviation for mean sounds:

$$-\frac{2\sigma v_2}{v_1} = -\frac{\sigma}{r^i}.$$

Thus, the inequality between both slopes transforms to the condition stated in the corollary. Involved derivatives are taken from the right in case of (BOND). ■

The Capital Market Line, which yields traders' consumption of standard deviation as a function of their desired mean, their budget and their non–spanned endowments, encourages to re–formulate traders' decision problem as follows:

$$\max_{\mu\in\mathbb{R}} \hat{v}^i\left(\mu + \mu_{w^i_{X^\perp}}, \hat{\sigma}\left(\mu, q(w^i_X), w^i_{X^\perp}\right)\right) \tag{$P^*$}$$

By quasi–concavity of $\hat{v}$ and by weak convexity of the Capital Market Line the first order condition of a trader's decision problem $(P^*)$ is also sufficient for obtaining an unique optimum in most circumstances. Attention has to be payed to unlimited improvements. Even if the Quasi–Bond has a positive price the price system might permit arbitrage opportunities preventing the existence of a solution. Moreover, a trader could be extremely risk averse, that is to say their coefficient of absolute risk aversion is larger than the slope of the Capital Market Line (Condition (NonSat) is violated). Then a riskless portfolio is most favourable to them.

**Lemma 46 (Sufficient First Order Conditions)** *Trader i's decision problem defined in Assumption (AGENTS) has a unique solution if the Condition (Viability) holds. The first order condition (FOC):*

$$\hat{v}^i_\mu + \hat{v}^i_\sigma \hat{\sigma}_\mu = 0, \tag{FOC}$$

*of the simplified choice Problem ($P^*$) is sufficient for determining the solution, if $\pi \nparallel \mathbf{1}_X \neq \mathbf{0}$ and Condition (NonSat) in case of (BOND) is satisfied.*

**Proof.** The Capital Market Line defines the minimal standard deviation $\hat{\sigma}\left(\mu, q(w^i_X), w^i_{X^\perp}\right)$ as a function of mean $\mu$. Thereby, the optimization problem can be simplified to the equivalent, one dimensional Problem ($P^*$). Since $\hat{\sigma}$ is convex in $\mu$ (weakly if (BOND) and $\sigma_y = 0$ hold) and the preference relation $\precsim_i$ is strictly convex, continuous and complete at most one solution exists. Existence is guaranteed since $\pi$ is viable by Condition (Viability) excluding arbitrage opportunities. Then Condition (FOC) is also sufficient for a differentiable $v^i$ [see Theorem 1 in (Mukherji, 1989)], if the solution is not located at the boarder of the choice set. Boarder solutions are ruled out by $\pi \nparallel \mathbf{1}_X \neq 0$ and Condition (NonSat) in case of (BOND), see Lemma 42 and Corollary 45. ■

A consumer buys the Quasi–Bond as long as they like its mean more than they dislike its variance. If they like more mean than they are able to finance, they have to purchase the zero–cost efficient asset, while their position of variance is becoming worse. At one point the additional variance of a bit more of the zero–cost efficient asset possibly off–sets its additional mean, which determines the optimal choice. It could well be that they are satiated in the Quasi–Bond without spending all of their budget. Their optimal demand for the Quasi–Bond as well as the zero–cost efficient asset is determined by the first order condition (FOC), which – spelled out – yields the so–called Theorem of Mutual Fund Separation in the CAPM.

**Corollary 47 (Mutual Fund)** *Let Condition (Viability) and* $\pi \not\parallel \mathbf{1}_X \neq \mathbf{0}$ *as well as Condition (NonSat) in case of (BOND) be satisfied. Then a trader's optimal choice of mean* $\mu^i$ *solves the equation*

$$\mu^i = \left( \frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} + \Upsilon^i \frac{\mu_\pi^2}{\mu_{\mathbf{1}_X} Var_X(\pi)} \right)^{-1} \left( \mu_{w^i_{X^\perp}} + \Upsilon^i \frac{\mu_\pi q\left(w^i_X\right)}{Var_X(\pi)} - r^i \right),$$

*with a consumption of variance determined by the Capital Market Line. A trader's consumption plan lies in the span* $\langle \mathbf{1}_X, \pi \rangle$ *and is given by Lemma 42.*[27] *Individual demand is positively homogenous of degree zero in* $\mu_\pi$.

**Proof.** See Appendix 2.11.3. ■

The preceding Corollary 47 places one in the position to prove the one–to–one correspondence between absolute and relative risk aversion. This claim was made at the end of Section 2.3.3 contradicting the result of (Löffler, 1995).

**Corollary 48** *Suppose (BOND),* $q(\mathbf{1}) = 1$, *Condition (NonSat) and Condition (Viability) hold. Then in–/decreasing/constant absolute and in–/decreasing/constant relative risk aversion are equivalent notions, respectively.*

**Proof.** Let $i$'s budget be denoted by $b := q\left(w^i_X\right)$. The first order condition (FOC), which is applicable because the promises of Lemma 46 are satisfied, sounds

$$\mu - b - Var(\pi)\, r^i = 0,$$

by using Corollary 47 and (BOND). It induces a consumption of mean $\mu$ and

[27] *If* $\pi \parallel \mathbf{1}_X$ *investors only consume* $\mathbf{1}_X$ *and if* $\mathbf{1}_X = \mathbf{0}$ *they consume* $\pi$ *to balance their budget whenever* $q\left(w^i_X\right) < 0$.

of variance $\hat{\sigma}^2 \equiv \hat{\sigma}^2\left(\mu, b, w^i_{X^\perp}\right)$. Because of

$$\frac{\partial}{\partial b}\left(\frac{\mu}{b}\right) = \frac{\frac{\partial}{\partial b}\mu - 1}{b^2}$$

in–/decreasing/constant relative risk aversion means

$$\frac{\partial}{\partial b}\mu \gtreqless 1.$$

The first order condition yields the implicit derivative of mean for budget:

$$\begin{aligned}\frac{\partial}{\partial b}\mu &= -\frac{-1 - r^i_{\sigma^2}\hat{\sigma}^2_b Var(\pi)}{1 - \left(r^i_{\sigma^2}\hat{\sigma}^2_\mu + r^i_\mu\right) Var(\pi)} \\ &= \frac{1 + 2r^i_{\sigma^2}(b-\mu)}{1 + 2r^i_{\sigma^2}(b-\mu) - r^i_\mu Var(\pi)}.\end{aligned}$$

The second equation follows from Lemma 42 in conjunction with (BOND). Since the first order condition (FOC) is also sufficient by Lemma 46 the nominator of this fraction has to be negative whenever Condition (FOC) holds, which is actually the second order condition. Then it is equivalent to write for in–/decreasing/constant *relative* risk aversion

$$r^i_\mu \lesseqgtr 0.$$

By the definition of the coefficient of absolute risk aversion these three cases are equivalent to

$$\frac{\hat{\sigma}}{\left(r^i\right)^2} r^i_\mu = \hat{S}^i_\mu\left(\mu^i, \sigma^i\right) \lesseqgtr 0,$$

which corresponds to in–/decreasing/constant *absolute* risk aversion. ■

This pre–work is an important step towards the determination of CAPM–equilibria, which is the main issue of the next subsection.

## 2.6 The CAPM–Equilibrium

The equilibrium properties are derived by solving the market clearing condition with respect to consumption plans and endowments (see $i$. of Definition 5 on page 23) for the pricing asset. Since (SPANNING) and (BOND) is not assumed the marginal rate of substitution of mean for variance appears in a slightly different shape. Therefore the following definition is introduced:

**Definition 49** *The* 'adjusted marginal rate of substitution' $R^i$ *is defined by*

$$R^i = \mu_{\mathbf{1}_X}\left(r^i - \mu_{w^i}\right) + \mu_{w^i_X}.$$

Since all aggregated quantities are abbreviated by a bar on top the identity

$$\bar{R} = \mu_{\mathbf{1}_X}\left(\bar{r} - \mu_{\bar{w}}\right) + \mu_{\bar{w}_X}$$

holds. The following proposition establishes the properties of CAPM–equilibria.

**Proposition 50 (CAPM–(EQU))** *Suppose that an asset with non–zero mean exists and $\bar{w}_X \neq 0$ holds. CAPM–Equilibria exhibit the following properties.*

1. *Traders are not over satiated in the Quasi–Bond nor are they satisfied with a riskless portfolio in case of (BOND). This implies the condition*

   $$\bar{R} \leq 0$$

   *and for all $i \in \mathbf{I}$ the condition*

   $$Cov_X\left(\bar{w}_X, w^i_X\right)\left(1 - \mu_{\mathbf{1}_X}\right) \geq -R^i \bar{R}.$$

   *Both inequalities hold strictly, i.e. with '$<$' and '$>$' respectively, if an investor $i$ is not satiated in the Quasi–Bond.*
   *This is assumed for the following properties (and ensured by Condition (NonSat) for at least one $i \in \mathbf{I}$).*

2. *The pricing asset separates into the Quasi–Bond and the market asset:*

$$\pi = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\mathbf{1}_X + \frac{\mu_\pi}{\bar{R}}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right).$$

*It has a price*

$$q(\pi) = \frac{\mu_\pi^2}{\mu_{\mathbf{1}_X}}\left[1 + \frac{Var_X(\bar{w}_X)}{\bar{R}^2}\right],$$

*and the price of the market asset shows to be*

$$q(\bar{w}_X) = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\left(\mu_{\bar{w}_X} + \frac{Var_X(\bar{w}_X)}{\bar{R}}\right).$$

3. *If $\bar{w}_X \nparallel \mathbf{1}_X$ the Beta–pricing formula reads:*

$$q(x) = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\mu_x + \frac{Cov_X(\bar{w}_X, x)}{Var_X(\bar{w}_X)}\left(q(\bar{w}_X) - \mu_{\bar{w}_X}\frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right).$$

4. *Investors' demand for mean in $X$ sounds:*

$$\mu^{*i} = \mu_{w_X^i} + \frac{\bar{R}Cov_X(\bar{w}_X, w_X^i) - R^i Var_X(\bar{w}_X)}{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)}.$$

*Their equilibrium consumption plan in $X$ satisfies the Tobin Separation in $\bar{w}_X$ and $\mathbf{1}_X$ :*

$$x_X^{*i} = \frac{\bar{R}R^i + Cov_X(\bar{w}_X, w_X^i)\left(1 - \mu_{\mathbf{1}_X}\right)}{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) + \frac{\mu^i}{\mu_{\mathbf{1}_X}}\mathbf{1}_X.$$

5. *Consuming more mean than endowed with, incorporates a higher consumption of $X$–standard deviation than the standard deviation of endowments projected onto the market asset. If $\bar{w}_X \nparallel \mathbf{1}_X$ the ratio of both excess demands is equal across investors:*

$$\left(\mu^i - \mu_{w_X^i}\right)\frac{-\bar{R}}{\sqrt{Var_X(\bar{w}_X)}} = \sqrt{Var_X(x_X^{*i})} - \frac{Cov_X(\bar{w}_X, w_X^i)}{\sqrt{Var_X(\bar{w}_X)}}.$$

6. *Equations involving the pricing asset are homogeneous in $\pi$ of degree zero.*

**Proof.** See Appendix 2.11.4. ■

Without loss of generality the mean of the pricing asset $\pi$ can be normalized due to equilibrium Property 6 and the property of homogeneity of individual demand established in Corollary 47 respectively. It is natural to choose:

$$\mu_\pi = q(\mathbf{1}_X) \equiv \mu_{\mathbf{1}_X},$$

which simplifies almost every expression in the CAPM. Moreover, it allows to interpret the pricing asset as the projection of a Radon–Nikodym derivative between the true and a pricing probability measure[28] on $X$ (or a pricing kernel, if preferences appear to violate monotonicity in an equilibrium, see also Section 2.9).

A desirable property of equilibria is Pareto–efficiency. Since markets are not necessarily complete, one can only expect equilibria to be constrained Pareto–efficient, because some Pareto enhancing exchanges lie in $X^\perp$. Moreover, non–monotonicity of utility functions can destroy general Pareto–efficiency, which calls for a narrower definition than in the GEI–world.

**Definition 51** *A CAPM–allocation $x^{\mathbf{I}}$ is said to be constrained Pareto–efficient if it is in RER and no allocation in RER is Pareto–superior.*

**Corollary 52** *Under the same prerequisites of Proposition 50 every equilibrium allocation is constrained Pareto–efficient.*

(NonSat) ensures traders' local non–satiation in the Quasi–Bond in any equilibrium, which induces constrained Pareto–efficiency. A proof of constrained Pareto–efficiency for the CAPM with (BOND) can be found in

[28] Some authors choose $\mu_\pi = 1$, so that $\pi$ itself might be interpreted as a Radon–Nikodym derivative. Unfortunately, for this choice a strictly positive extension of $\pi$ to $X^\perp$ would not induce a probability measure. However, if consumption in the first period is considered, the normalization $q(\mathbf{1}_X) = \frac{\mu_{\mathbf{1}_X}}{1+Y}$ with $Y$ being the market yield is more appropriate.

[(Magill & Quinzii, 1996), (iv) of Theorem 17.3 on p. 181] and with non–homogenous expectations in [(Güth *et al.*, 2000), Proposition 1].
Constraint Pareto–efficiency generally holds for financial market GEI–economies, i.e. with only one physical good in all states of nature. With multiple goods this is not the case generically in endowments, utility functions and asset structures, because income effects in the subsequent good markets in each state are not efficiently taken into account by the preceding asset markets, see Ch. 4 in (Hens, 1998). But (Geanakoplos & Shubik, 1990) have shown that for (QUADU) the equilibrium allocation is Pareto–optimal, if (SPANNING) holds, a 'special' riskless asset exists for one trader, and agents show to be on the monotone part of their quadratic utility function.[29]

**Remark 3** *Note that in RER traders are forced to consume their entire portfolio pay–offs. Thereby, free–disposal is ruled out in RER. In the CAPM where free–disposal is allowed it is not always optimal to consume the entire portfolio pay–offs since preferences are not necessarily monotone in state pay–offs. In this case a financial market equilibrium might exist, but not an exchange equilibrium. Thus, for mean/variance utility functions monotonicity is required if the GEI–definition of constrained Pareto–efficiency is considered, where only the trade of consumption streams is restricted to the (incomplete) marketed subspace, but free–disposal is permitted.*

(Hara, 1997) considered the question, how equilibria change when new assets are introduced. Hara has shown for the CAPM that a sequence of Pareto–improving asset innovations always exists. Moreover, she investigated the price changes of the market portfolio due to asset innovations in a CAPM with (BOND) and (SPANNING). Suppose an introduction of assets from $X^{\perp}$, which are also orthogonal to the market portfolio and the riskless asset but not to everyone's endowments ((BOND) and (SPANNING) is not critical for this line of reasoning). These assets have a zero price and have

[29] Their result holds also for other special vNM–utility functions, for which the derivative of the Walrasian demand function with respect to income is independent of income. Thereby the income effect leading to Pareto–inefficiency is ruled out.

no direct income effects. But they are used by agents to hedge each other's risks, thus the consumption of variance strictly shrinks with those assets.[30] The price changes depend on how $\bar{R}$ reacts on an overall decreasing variance, see (Hara, 1998). However, in general the changes of equilibria due to innovations not orthogonal to the market portfolio and the riskless asset are unpredictable.

It remains a problem that the equilibrium equations are just implicit solutions in mean. This endogeneity will pass on to almost every statement, but it is common in the CAPM literature. Nevertheless, the properties of equilibria established in Proposition 50 further restrict the $RER$. For instance only pricing assets in $\langle \bar{w}_X, \mathbf{1}_X \rangle$ have to be considered in $RER$. This allows to re–formulate the conditions for no–arbitrage and non–satiation in a sharper way.

**Corollary 53** *Possible equilibrium pricing assets:*

$$\pi\left(\bar{R}\right) = \mathbf{1}_X + \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right),$$

*rule out satiation in the Quasi–Bond by:*

$$R^i\bar{R} + Cov_X\left(\bar{w}_X, w_X^i\right)\left(1 - \mu_{\mathbf{1}_X}\right) \geq 0 \quad and\ \bar{R} < 0,$$

*as well as arbitrage opportunities by:*

$$\bar{R}^2 \geq Var_X\left(\bar{w}_X\right)\left[\frac{\mu_{\mathbf{1}_X}}{\left(\min_{i\in\mathbf{I}} \hat{s}^i\right)^2} - \left(1 - \mu_{\mathbf{1}_X}\right)\right],$$

*for possible equilibrium $R^{\mathbf{I}} \in RER$.*

**Proof.** The first claim is already mentioned in Proposition 50. The second assertion follows from Condition (Viability) on page 56 by replacing the pricing asset with $\pi\left(\bar{R}\right)$. ■

In case of $\mathbf{1}_X \neq \mathbf{1}$ it may well be that both conditions contradict for some $R^i > 0$. Then $\left|\bar{R}\right|$ should be small to satisfy the first, but large to guarantee

[30] Those assets may not be Pareto–improving because of indirect price effects.

the second condition. If (BOND) is assumed the first condition reduces to $r^i \bar{r} > 0$, which is already satisfied by definition. But note that this condition also rules out riskless portfolios in any equilibrium.
The next example exploits the fact that in case of (QUADU) the equilibrium is explicitly solvable. For (QUADU) it is unnecessary to check the no–arbitrage condition since for large mean satiation always occurs, because in this region utility is no longer monotone in mean. Neither is it necessary to consider arbitrage in the case of (LINU) because the limiting slopes are infinite. Moreover, the condition for satiation is trivial in the case of (LINU). The following example considers the verification of the condition of satiation (NonSat) if (AGENTS) satisfy (QUADU):

**Example 54 (QUADU and Satiation)** *Suppose (AGENTS) satisfy (QUADU). Then any equilibrium requires that* $\bar{R} = \mu_{\mathbf{1}_X} \sum_{i \in \mathbf{I}} \frac{1}{\varphi^i} - \mu_{\bar{w}_X}$ *and* $R^i = \mu_{\mathbf{1}_X} \left( \mu^{*i} - \frac{1}{\varphi^i} - \mu_{w^i} \right) - \mu_{w^i_X}$ *hold. Equilibrium Property (4), which yields the consumption in mean, transfers to:*

$$\left( \mu^{*i} - \mu_{w^i_X} \right) \left( \bar{R}^2 + Var_X \left( \bar{w}_X \right) \left( 1 - \mu_{\mathbf{1}_X} \right) \right) = \bar{R} Cov_X \left( \bar{w}_X, w^i_X \right) - R^i Var_X \left( \bar{w}_X \right).$$

*By replacing* $\mu^{*i}$ *with the help of the definition of* $R^i$ *this equation solves for* $R^i$ *implicitly in* $\bar{R}$ *:*

$$\begin{aligned} R^i &= \frac{1}{\bar{R}^2 + Var_X \left( \bar{w}_X \right)} \Big\langle \mu_{\mathbf{1}_X} \bar{R} Cov_X \left( \bar{w}_X, w^i_X \right) + \\ &\quad \left[ \mu_{w^i_X} - \mu_{\mathbf{1}_X} \left( \frac{1}{\varphi^i} + \mu_{w^i_{X^\perp}} \right) \right] \left[ \bar{R}^2 + Var_X \left( \bar{w}_X \right) \left( 1 - \mu_{\mathbf{1}_X} \right) \right] \Big\rangle . \end{aligned}$$

*Now* $R^i$ *and therewith* $\mu^{*i}$ *are exogenous constants since* $\bar{R}$ *is. The products* $R^i \bar{R}$ *have to be applied to the condition of non–satiation. If the condition is satisfied for all investors a unique equilibrium is established.*

The calculations and formulas without (BOND) are very messy. For the matter of completeness the result for a CAPM–equilibrium with (BOND) is given next:

**Corollary 55 (Equilibrium with (BOND))** *Suppose (BOND) and* $Var(\bar{w}_X) > 0$ *hold and that* $\mu_\pi$ *is normalized to one. Then CAPM–equilibria exhibit the following properties:*

1. *The pricing asset separates into the Bond and the market asset:*

$$\pi = \mathbf{1} + \frac{1}{\bar{r}}\left(\bar{w}_X - \mu_{\bar{w}_X}\mathbf{1}\right).$$

*It has a price of*

$$q(\pi) = 1 + \frac{Var(\bar{w}_X)}{\bar{r}^2},$$

*and the price of the market asset sounds*

$$q(\bar{w}_X) = \mu_{\bar{w}} + \frac{Var(\bar{w}_X)}{\bar{r}}.$$

2. *The Beta–pricing formula reads:*

$$q(x) = \mu_x + \frac{Cov(\bar{w}_X, x)}{Var(\bar{w}_X)}\left(q(\bar{w}_X) - \mu_{\bar{w}}\right).$$

3. *Trader i's consumption in X satisfies the Tobin Separation:*

$$x_X^i = \frac{r^i}{\bar{r}}\left(\bar{w}_X - \mu_{\bar{w}_X}\mathbf{1}\right) + \mu^i\mathbf{1},$$

*while his/her consumption of mean and variance in X sounds:*

$$\begin{aligned} \mu^i &= \mu_{w^i} + \frac{1}{\bar{r}}Cov(\bar{w}_X, w_X^i) - \frac{r^i}{\bar{r}^2}\sigma_{\bar{w}_X}^2 \quad \textit{and} \\ \sigma^i &= \frac{r^i}{\bar{r}}\sigma_{\bar{w}_X}, \quad \textit{respectively.} \end{aligned}$$

4. *Consuming more mean than endowed with, incorporates a higher consumption of standard deviation than the standard deviation of endowments projected onto the market asset. The ratio of both excess demands*

*is equal across investors:*

$$\left(\mu^i - \mu_{w^i}\right) \frac{-\bar{r}}{\sigma_{\bar{w}_X}} = \sigma^i - \frac{Cov(\bar{w}_X, w^i_X)}{\sigma_{\bar{w}_X}}.$$

**Proof.** Agents are never satiated with a riskless portfolio since $R^i = r^i < 0$ and $r^i \bar{r} > 0$. The properties follow immediately from Proposition 50. ■

## 2.7 The Mean/Variance–Economy

A consumer's demand shows to have the 'nice' characteristic in that it satisfies the Tobin Separation. Since every trader is only interested in mean and variance the two chosen assets, $\mathbf{1}$ and $\bar{w} - \mu_{\bar{w}}\mathbf{1}$, can exclusively be associated with the two goods, mean and variance, respectively. This leads to the presumption that a corresponding economy exists, in which traders just optimize between mean and variance as if those were perceived as real goods. Then the market clearing condition shall hold also with respect to those goods. This idea goes back to (Dana, 1999). It considerably simplifies investigations about the aggregated demand in mean, since investors' portfolio choice is reduced to a 2–dimensional problem without considering the complexity of the original consumption space $L$. This bears the risk, that problems of satiation and monotonicity are neglected.

Neither (BOND) nor (SPANNING) are necessary for the re–formulation of the CAPM–economy Dana proposed. The mean/variance–economy is defined as follows:

**Definition 56 (Mean/Variance–Economy and its Equilibrium)** *A tuple* $\left\{X,\ (\mathbb{R} \times \mathbb{R}_+),\ (\hat{v}^i, w^i, \rho^i)^{\mathbf{I}}\right\}$ *is said to be a mean/variance–economy,* *if – provided that* $\mu_{\mathbf{1}_X} > 0$ *– agents* $i \in \mathbf{I}$ *solve the decision problem:*

$$\max_{(\mu^i,\sigma^i)\in\mathbb{R}\times\mathbb{R}_+} \hat{v}^i\left(\mu^i + \mu_{w^i_{X^\perp}}, \tilde{\sigma}\left(\mu^i,\sigma^i,w^i_{X^\perp}\right)\right)$$

$$s.t. \quad \mu^i - p\sigma^i \leq \mu_{w^i_X} - p\rho^i\sqrt{Var_X(w^i_X)} \text{ and}$$

$$\tilde{\sigma}\left(\mu^i,\sigma^i,w^i_{X^\perp}\right) = \sqrt{\frac{1}{\mu_{\mathbf{1}_X}}\left(\sigma^i\right)^2 + \left(\mu^i\right)^2\frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu^i\mu_{w^i_{X^\perp}} + \sigma^2_{w^i_{X^\perp}}},$$

*with a negative price* $-p \in \mathbb{R}_{--}$ *for standard deviation and an individual weighting factor* $\rho^i$ *for standard deviation of spanned endowments.*
*In this economy an equilibrium* $\left\{p^*,(\mu^*,\sigma^*)^{\mathbf{I}}\right\}$ *is such that investors solve their decision problem and markets for mean and standard deviation clear:*

$$\bar{\mu}^* = \mu_{\bar{w}} \text{ and } \bar{\sigma} = \sqrt{Var_X(\bar{w}_X)}.$$

To establish a one–to–one correspondence between the two notions of an economy, traders' demand in mean and standard deviation must behave in the same way.

**Lemma 57 (Demand Equivalence)** *The demand–allocations with respect to mean and variance of the CAPM–economy and the mean/variance–Economy correspond one–to–one if*

1. *the utility functions coincide*[31]*:*

$$\hat{v}^i(\mu,\sigma) = v^i(\mu,\sigma^2) \quad on \quad \mathbb{R}\times\mathbb{R}_+,$$

2. *the weighting factors of the standard deviations of spanned endowments are equal to the X–correlations between* $-\pi$ *and* $w^i_X$ *:*

$$\rho^i \equiv \rho_X\left(-\pi, w^i_X\right), \quad and$$

[31] *Positive affine transformations would be sufficient.*

*3. the pricing asset and the (relative) price of standard deviation are related by*

$$p = \frac{\sqrt{Var_X(\pi)}}{\mu_{\mathbf{1}_X}} \quad \textit{with} \quad \mu_\pi = \mu_{\mathbf{1}_X}.$$

**Proof.** If two investors receive in both economies the same standard deviation as a function of mean they would choose the same mean, provided their utility functions coincide in both economies. In order to show the if–conjecture both investors' choice sets have to be identical for the same endowments $w^i$.

In the mean/variance–economy agent $i$'s demand for standard deviation $\sigma^i$ is determined by his/her budget constraint for an optimal level of mean $\mu^i$:

$$\sigma^i = \frac{\mu_{w_X^i} - p\rho^i \sqrt{Var_X(w_X^i)} - \mu^i}{-p}.$$

By using the conditions for $p$ and $\rho^i$ this equation transforms further to

$$\begin{aligned}
\sigma^i &= -\frac{1}{p}\left(\mu_{w_X^i} - \frac{\sqrt{Var_X(\pi)}}{\mu_{\mathbf{1}_X}} \frac{Cov_X(-\pi, w_X^i)}{\sqrt{Var_X(\pi)\, Var_X(w_X^i)}} \sqrt{Var_X(w_X^i)} - \mu^i\right) \\
&= -\frac{1}{p}\left(\mu_{w_X^i} - \frac{\mu_{\mathbf{1}_X}(-\pi \bullet w_X^i) + \mu_{\mathbf{1}_X}\mu_{w_X^i}}{\mu_{\mathbf{1}_X}} - \mu^i\right) \\
&= -\left(q\left(w_X^i\right) - \mu^i\right) \frac{\mu_{\mathbf{1}_X}}{\sqrt{Var_X(\pi)}}.
\end{aligned}$$

With this result the consumption of standard deviation $\tilde{\sigma}\left(\mu^i, \sigma^i, w_{X^\perp}^i\right)$ is identical to the optimal standard deviation $\hat{\sigma}\left(\mu^i, q\left(w_X^i\right), w_{X^\perp}^i\right)$ derived in Lemma 42 (Capital Market Line). Thus, both investors face the same choice sets. Because utility functions are identical the optimal demand for mean and standard deviation is identical as well in both formulations. ■

It remains to show that equilibrium allocations in mean and variance coincide in both economies.

**Corollary 58** *The equilibrium allocations in mean and variance coincide in* *the CAPM–economy and the mean/variance–Economy.*

**Proof.** Under the assumptions of the previous lemma the demand for mean and standard deviation is equivalent in both economies. That in CAPM–equilibria markets for mean and standard deviation in $X$ are cleared remains to be shown. While in equilibrium markets for mean clear by the fourth property of Proposition 50 the market clearing condition for standard deviation is implied by the fifth property of this proposition in conjunction with the market clearing for mean.[32] ■

Some authors [e.g. (Bottazzi *et al.*, 1998)] use Dana's mean/variance–Economy as a starting point for analyzing existence and uniqueness in the CAPM. The framework of the CAPM–economy **(ENV)** is not abandoned here, because monotonicity can only be considered in this setting and the market demand has been already reduced to a one dimensional problem by Corollary 47.

## 2.8 Existence and Multiplicity of Equilibria

In this section some important results are reviewed regarding existence and uniqueness (or multiplicity) of CAPM–equilibria. Whenever the properties of CAPM–equilibria are applied, existence ensures that one is not guided by an artefact, while uniqueness is a prerequisite of comparative studies.

The set of CAPM–equilibria shall be denoted by $T_E \subset RER$. The main existence result about CAPM–equilibria is due to (Nielsen, 1990b):

**Theorem 59 (Existence)** *In a CAPM–economy with (SPANNING) and (NonSat) $T_E \neq \emptyset$ holds.*

Nielsen's most general contributions about existence are (Nielsen, 1989) and (Nielsen, 1990b). In the first article he applies the ideas first introduced in (Hart, 1974), who proved a rather general existence result for financial market economies. But Hart's assumptions, most notably vNM–utility functions

[32] This establishes Walras' Law for mean and standard deviation.

with priors depending on prices, bounded consumption, and non–negative asset pay–offs, are not valid in the CAPM. Moreover, Hart pointed out that existence may fail, if choice sets are unbounded, which is the case in the CAPM with admissable unlimited shot selling and unbounded support of asset returns. In a more general financial market model as the CAPM (Nielsen, 1989) has shown that (NonSat)[33] is one of the sufficient conditions[34] guaranteeing existence. In (Nielsen, 1990b) he applied this result to the CAPM. Unfortunately, agents are endowed with assets which implies (SPANNING). But this assumption does not seem to be essential to his result, since the non–spanned endowments only determine a 'new' mean/variance utility function exhibiting important properties of the old one. In (Nielsen, 1990b) Nielsen gives sufficient conditions for (NonSat) to be satisfied:

1. Existence of a riskless asset.
2. (BOND) and $\hat{s}^i = \infty$ for all $i \in \mathbf{I}$, if investors do not agree on expected returns.
3. The Quasi–Bond is an unbounded improvement for any trader.
4. Agents are not satiated in the Quasi–Bond while consuming the entire market asset, and they are better off with their endowments than having nothing.

(Nielsen, 1990a) considered two special cases, (LINU) and $1_X \parallel \bar{w}_X$. In both cases a unique equilibrium exists, if everybody's endowments do not suffice to purchase their satiating portfolio.
(Dana, 1999) assumes concave utility functions in mean and variance[35], a riskless asset and (SPANNING). Her very useful contribution was to show that the equilibrium correspondence can be reduced to that of a two goods

[33] Nielsen calls it 'Non–satiation at Pareto attainable portfolios'.
[34] 'Positive semi–independence of directions of improvement' is a necessary assumption, but which is not critical for the CAPM.
[35] This is a much stronger requirement than quasi–concavity in mean and standard deviation as was pointed out above in Lemma 17 on page 38.

economy in mean and standard deviation only (see Section 2.7). Thereby, Dana's proof is analytically and applies the most basic version of a standard fixed point argument. She shows that the individual, continuous net supply of standard deviation is negative for small and positive for large relative prices of standard deviation which implies an equilibrium price in between.

In a CAPM with concave utility functions in mean and variance and with (SPANNING) (Allingham, 1991) investigated existence of equilibria with positive prices for assets having a positive mean. He gave two examples of non–existence, of which one is due to satiation and the other is due to the violation of the price boundary. His conclusion is intuitive: Existence with positive asset prices is guaranteed, if agents' risk aversion is sufficiently small. (Laitenberger, 1997) extended Allingham's result within Dana's mean/variance–economy. In a CAPM–economy, in which free disposal in first period income is prohibited, he derived sufficient conditions for existence – more general but similar to Allingham's – with a positive as well as a negative price for the Quasi–Bond.[36]

The properties of an equilibrium are approachable to comparative investigations and are especially of practical value, if investors can be assured of a unique equilibrium. But this is generally not the case. The possibility of multiple equilibria in the CAPM has been illustrated by (Nielsen, 1988). Nielsen constructed two CAPM–economies with two equally endowed investors showing to have two different equilibria each. In his first example investors have decreasing absolute risk aversion. And in the second example each investor consumes the same amount of standard deviation in both equilibria. The latter case may arise if one investor has a lower and the other has a higher absolute risk aversion from one equilibrium allocation to the other. Before turning to the more refined results on uniqueness, one should see a direct consequence of the equilibrium properties.

**Corollary 60 (Uniqueness)** *Suppose that $T_E \neq \emptyset$.*

· *Any equilibrium in $T_E$ is sustained by a unique pricing asset.*

[36]Existence with the Quasi–Bond having a negative price is due to over–satiation and prohibited free disposal, which is a suspicious situation.

- *Equilibria with the same pricing asset are identical.*

- *Equilibria coincide if an investor's ratio of non–zero access demands in mean and standard deviation (due to Property 5 of Proposition (50)) stays the same.*

- *Equilibria coincide if all investors consume the same standard deviation in $X$ across equilibria and if their risk aversion is decreasing in mean (due to (Nielsen, 1988)).*

- *$T_E$ is a singleton, if the sum of the MRSs between mean and variance $2\bar{r}$ is equal across equilibria.*

- *The equilibrium is explicitly solvable, if each individual's MRS $2r^i$ is linear in mean and constant in standard deviation.*

**Proof.** The Proposition (50) states all equilibrium quantities as parameterized formulas in $\bar{R}$ and $R^i$. The formula for the pricing asset yields the first assertion, since in a single equilibrium $\bar{R}$ is uniquely determined. The second claim follows from the unique solution of each trader's decision problem, see Lemma 46. The third condition determines an equal $\bar{R}$ and therewith an equal pricing asset across equilibria. The same argument holds for the fifth and the sixth assertion. The formula establishing the consumption of mean is linear in mean and thereby explicitly solvable, if the last prerequisite is assumed. For a proof of the fourth assertion see Proposition 2 in (Nielsen, 1988). ■

(Nielsen, 1988) claimed that a pricing asset may support distinct equilibrium allocation. This is not possible due to Lemma 46. From a GEI point of view multiplicity is not surprising because of Debreu's, Mantel's and Sonnenschein's results on the structure of market excess demand [for an overview consult (Shafer & Sonnenschein, 1982) and (Hens, 1998)]. Their results were successfully transferred to the CAPM by (Bottazzi *et al.*, 1998):

**Theorem 61 (Structure of Market Demand)** *In a CAPM with a riskless asset for any market portfolio with positive mean and variance and any* *finite set $\mathbb{P} \subset X$ of in pairs different and normalized pricing assets $\pi \in \mathbb{P}$, any function $\bar{x}(\pi) : \mathbb{P} \to X$ with nonzero variance that satisfies Walras' Law and the Tobin Separation Property, i.e. $\bar{x}(\pi) \in \langle \mathbf{1}, \bar{w}_X \rangle$, is the aggregate demand of two CAPM–investors.*

The somewhat sobering conclusion is, that one could find economies with any finite number of equilibria and with varying pricing assets across equilibria. This negative result is soothed by some contributions investigating the uniqueness of equilibria. (Dana, 1999) proved uniqueness under the assumptions (BOND) and (SPANNING), if investor have additive separable, quasi–concave utility functions in mean and variance and the utility functions satisfy $-\mu \frac{v_{11}}{v_1} < 1$.[37] These assumptions lead to monotone individual excess demand functions in mean, which implies uniqueness. But these assumptions are only compatible with increasing absolute risk aversion, which is criticized to be unrealistic. Within Dana's framework of a mean/variance–economy (Hens *et al.*, 2000) contribute a corresponding result which allows also for non–increasing absolute risk aversion – generally considered as the more realistic behaviour. The following theorem is a slight modification of their result, which goes through with the proof of the original version.

**Theorem 62 (Uniqueness)** *In a CAPM–economy with (BOND), (SPANNING) and $\bar{w} \nparallel \mathbf{1}$, $T_E$ is singleton if $\forall (\mu, \sigma)^I \in RER$ every consumer $i \in \mathbf{I}$ has non–decreasing or non–increasing absolute risk aversion in mean whenever $(\mu^i, \sigma^i) < \left(\mu_{w^i}, \frac{Cov(\bar{w}, w^i)}{\sigma_{\bar{w}}}\right)$ or $(\mu^i, \sigma^i) \geq \left(\mu_{w^i}, \frac{Cov(\bar{w}, w^i)}{\sigma_{\bar{w}}}\right)$, respectively.*

However, (Hens *et al.*, 2000) also give a counter–example showing that the result is not robust in markets without a riskless asset. They construct an economy with two constantly absolute risk averse investors and without a riskless asset, which has at least two equilibria. To all appearance uniqueness in the general setup without (SPANNING) and (BOND) is up to now not nicely characterized.

[37] This condition would imply gross substitution in an exchange economy, which in turn leads to uniqueness, but which is not easy to interpret in the CAPM. For vNM–utility functions it would mean a relative risk aversion less than one. (Dana, 1999) pointed out that the so called Mitjushin–Polterovitch–condition is not transferable to the CAPM.

## 2.9 Monotonicity and Positive State Prices

A very unpleasant weakness of the CAPM is the possibility of negative state prices. This is caused by mean/variance utility functions, which are not necessarily monotone in the consumption of state pay–offs: With more consumption in a particular event it might be possible to increase variance in such a huge (and bad) manner that the increase in mean is outweighed. Consequently, investors would try to avoid more consumption in those events or, in other words, they value an asset paying positive amounts in those events negatively. The risk measure, the variance of pay–offs, is the source of such a 'negative' behaviour of the pricing functional, which counteracts any economic intuition, especially if one has in mind the definition of arbitrage and the Fundamental Theorem of Asset Pricing (FTP) from the General Equilibrium Theory of Incomplete Markets (GEI). An arbitrage opportunity exists in a GEI–economy, if there is a traded consumption plan with a non–positive price, which offers non–negative returns and positive returns with positive probability[38], formally:

**Definition 63 (GEI-arbitrage)** *A price system $q$ permits GEI–arbitrage if and only if there is a $x \in L_{++} \cap X$ with $q(x) \leq 0$. A pricing asset $\pi$ permits GEI–arbitrage if and only if the price system $q : X \rightarrow \mathbb{R}$ defined by $q(x) = \pi \bullet x$ does.*

The FTP states that arbitrage for the pricing asset $\pi$ is ruled out if and only if an extension $\eta \in X^{\perp}$ exists such that $\pi + \eta$ is a Radon–Nikodym derivative of an equivalent probability measure $Q \sim P$ with respect to $P$. In this case all positive pay–offs in essential events $z \in L_{++}$ are valued positively by $(\pi + \eta) \bullet z$ and the extended pricing functional coincides with $q$ on $X$ since $(\pi + \eta) \bullet x = q(x)$ for $x \in X$.

**Theorem 64 (Fundamental Theorem of Asset Pricing)** *Markets do not offer GEI–arbitrage opportunities for the pricing asset $\pi \in X$ if and*

[38] Consumption today is not considered in the investment based CAPM.

*only if there is a* $\eta \in X^{\perp}$ *such that* $Q$ *defined by* $dQ = (\pi + \eta)\, dP$ *is an* $P$*–equivalent probability measure on* $(\Omega, \mathcal{F})$.

For finite $\mathcal{F}$ the proof can be found in [(Magill & Quinzii, 1996) in §9]. The infinite dimensional case is covered by the results of (Clark, 1993), which are just combined by the following proof:

**Proof.** The existence of a pricing asset $\pi$ has already been established in Lemma 33, which implies a continuous, linear pricing functional on $X$. Since $X$ is a finite dimensional subspace of $L$ 'approximate arbitrage' does not apply (with regard to $X^{\perp}$), see Clark's Theorem 6. Moreover, the probability space, the probability measure and the consumption space satisfy the promises of his Theorem 7. By this theorem a continuous, strictly positive, and linear extension of the pricing functional to the entire $L$ exists if and only if $\pi$ does not offer arbitrage opportunities in $X$. The decomposition of the pricing extension into $\pi$ and $\eta$ follows from Riesz' representation theorem. ■

The presumption of the existence of $\pi$ is justified for the CAPM by Lemma 33, if an asset with non–zero mean exists.[39] $\pi + \eta$ is called the pricing density, $dQ$ the state price functional (particularly in case of a finite state space), and $Q$ the equivalent pricing measure.[40] Since $\pi$ could be negative on essential events, it does not necessarily define a density on its own, see Footnote 42 for an example.

Surprisingly, in the CAPM–literature the problem of non–monotonicity is known, see (Nielsen, 1987) and (Nielsen, 1992). But the question in concern has been the existence of positive asset prices instead of state prices in (Nielsen, 1992), (Hiroshi & Hiroshi, 1995), and (Allingham, 1991). The following example demonstrates that positive asset prices do not matter for GEI–arbitrage, however the positiveness of state prices really does:

[39] For the GEI–model the case is different. $X$ might be such that $X \cap L_{++}$ is empty. Then every price functional is obviously arbitrage–free in the GEI–world, which also could be non–linear and discontinuous. Provided that $X \cap L_{++} \neq \emptyset$ a price functional is arbitrage–free if and only if it is strictly positive and linear by Clark's Theorem 2 and continuous by Theorem 3, since $X$ is a closed vector space, see (Clark, 1993).

[40] In intertemporal arbitrage pricing models $Q$ is known as a martingale measure.

**Example 65 (Negative State Prices)** *Suppose the asset structure consists only of two assets: the (BOND) and the market asset $\bar{w}$. The sum of all endowments shall be a non–negative random variable with positive variance. Furthermore, let (LINU) hold, so that $\bar{r} = -\sum_i \frac{1}{\varphi^i}$. Hence, the price of the market asset reads:*

$$q(\bar{w}) = \mu(\bar{w}) + \frac{1}{\bar{r}}\sigma_{\bar{w}}^2.$$

*For negative state prices it is sufficient have:*

$$q(\bar{w}) < 0 \Leftrightarrow E(\bar{w}^2) > \mu_{\bar{w}}^2(1 - \bar{r}).$$

*To achieve this requirement $\bar{w}$ should have in some states of nature a very high pay–off with a very low probability to overweight the second to the first moment. The inequality is fulfilled in the following economy:*

$$\begin{aligned}
\Omega &= \{\omega_1, \omega_2, \omega_3\}, \\
P(\{\omega_1\}, \{\omega_2\}, \{\omega_3\}) &= (0.01, 0.95, 0.04), \\
\bar{w} &= (100, 1, 0), \\
(\varphi^1, \varphi^2) &= (1/18, 1/30), \\
\bar{r} &= -\left(\frac{1}{1/18} + \frac{1}{1/30}\right) = -48, \\
\mu_{\bar{w}} &= 1.95, \\
E(\bar{w}^2) &= 100.95.
\end{aligned}$$

*The pricing asset and the price of the market asset turn out to be:*

$$\begin{aligned}
\pi &= \frac{1}{\bar{r}}[\bar{w} - \mu(\bar{w})\mathbf{1}] + \mathbf{1} \\
&= -\frac{1}{48}[(100, 1, 0) - 1.95 \cdot (1, 1, 1)] + (1, 1, 1) \\
&\approx (-1.0427, 1.0198, 1.0406) \quad \textit{and} \\
q(\bar{w}) &= E_Q(\bar{w}) = E(\pi\bar{w}) \\
&\approx 100 \cdot (-1.0427) \cdot 0.01 + 1 \cdot 1.0198 \cdot 0.95 + 0 \cdot 1.0406 \cdot 0.04 \\
&= -0.07389 < 0.
\end{aligned}$$

*In this example the non–negative market asset has positive pay–offs with positive probability. Nevertheless one would get money for buying this asset. A third asset, which would complete the market, does not change this result, since the bond and the market asset are already spanned. Imagine an additional Arrow–security, which pays only one unit in the first state of the world. Its price would be negative here contradicting the no–arbitrage condition in GEI–models.*

The question of positive security prices was raised, because one observes almost everywhere in financial markets positive prices. Hence negative asset prices in the CAPM would be an artefact suggesting a refutation of the model. One must oppose this view, because of at last two reasons: First of all, asset prices are a matter of normalization. Equilibria stay the same if different asset structures are chosen while the marketed subspace remains unchanged. For example, if there is an consumption plan with positive mean and positive price one is able to impose this property on all assets, i.e. by adding a large amount of this consumption plan onto all assets. Second, in real financial markets, assets with negative pay–offs involve the risk of bankruptcy, i.e. the buyer's promise to pay positive amounts in the future may not be met. The default risk implies costs of monitoring or demands for collateral, which is not explicitly modelled either in the CAPM nor in the GEI.[41] These frictions

[41] In the GEI–framework research on bankruptcy has been started recently by (Geanakoplos *et al.*, 1996) and (Dubey *et al.*, 2000).

impose a direction upon the evolution of securities in favour of developing assets with just positive pay–offs, where the short side is kept by a large institution with enough collateral, a broad diversification of its assets and low costs of monitoring enforced by law. Then just by arbitrage these assets have positive prices. Much less often one observes assets with mixed, positive and negative, pay–offs. Futures belong to that kind of securities. In these markets one has usually restricted participation, rules to put in collateral or margin payments, etc.

Obviously, if for every spanned consumption stream with positive mean a positive price is demanded, GEI–arbitrage is ruled out. But this would be an unnecessarily strong condition.

Nevertheless, sufficient conditions for asset prices to be positive are briefly summarized, see (Nielsen, 1992), where the first two statements are cited from:

**Proposition 66 (Positive Asset Prices)** *Suppose one normalizes assets to have non–negative mean. Then the following conditions are sufficient for the price $p(A_j)$ of asset $A_j \in A$ to be positive in any equilibrium with (BOND):*

1. *For all $x^{\mathbf{I}} \in RER$ there is an investor $i \in \mathbf{I}$ such that*

$$\mu_{A_j} > \frac{\sigma_{A_j,x^i}}{-r^i \circ x^i}.$$

2. *For all $x^{\mathbf{I}} \in RER$ there is an investor $i \in \mathbf{I}$ such that*

$$\frac{\mu_{A_j}}{\sigma_{A_j}} > \hat{S}^i \circ x^i.$$

3. *For all $x^{\mathbf{I}} \in RER$ it holds that*

$$\frac{\mu_{A_j}}{\sigma A_j} > \frac{\sigma_{\bar{w}_X}}{-\bar{r} \circ x^{\mathbf{I}}}.$$

*From the first to the third the assertions are becoming consecutively stronger.*

**Proof.** [Sketch of the proof] The first expression results from the marginal utility in the direction of the particular asset. The second involves the upper bound $\sigma_{x,y} \leq \sigma_x \sigma_y$. And the third condition uses the fact, that $\hat{S}^i \geq \sigma_\pi = -\frac{\sigma_{\bar{w}X}}{\bar{r}}$ in equilibrium (" = " if (SPANNING) holds), see the fifth property in Corollary 55. For the first two conditions one needs that marginal utilities point in the same direction in equilibrium. See also (Nielsen, 1992). ■

The first condition is also necessary if the equilibrium allocation is concerned. It involves a covariance term, which is omitted in the second, whereas the third makes use of an aggregated level instead of individual values. Those conditions do not rule out arbitrage opportunities as the next example points out.

**Example 67 (Arbitrage with positive asset prices)** *Suppose one changes in the last example the set of risky assets to be*

$$A_2 = (+70, +5, -80) \quad \textit{and} \quad A_3 = (-30, +4, -80),$$

*so that* $A_2 - A_3 = \bar{w}$ *with* $\mu_{\bar{w}}/\sigma_{\bar{w}} \approx 0.198$. *Agents shall share the entire endowments symmetrically, while anything else remains unchanged. Their equilibrium demand is* $x^{*1} \approx (+37.49, +0.366, -0.009)$, *and* $x^{*2} \approx (+62.51, +0.634, +0.009)$, *respectively. Some computations show*

$$\begin{aligned} \frac{\sigma_{A_1,x^{*i}}}{-r^i \circ x^{*i}} &\approx 1.466, \quad \mu_{A_2} = 2.25, \quad \mu_{A_2}/\sigma_{A_2} \approx 0.125, \quad \textit{and} \\ \frac{\sigma_{A_1,x^{*i}}}{-r^i \circ x^{*i}} &\approx -.558, \quad \mu_{A_2} = 0.30, \quad \mu_{A_3}/\sigma_{A_3} \approx 0.018. \end{aligned}$$

*Since* $\hat{S}^i \approx 0.205$ *the first and not the second of above conditions ensures positive asset prices:*

$$p(A_2) \approx 0.784 \quad \textit{and} \quad p(A_3) \approx 0.858.$$

*Although* traded *assets have positive prices arbitrage opportunities are still possible, since the market asset has a negative price by the first of above* *conditions:*

$$\frac{\sigma_{\bar{w},x^{*i}}}{-r^i \circ x^{*i}} \approx 2.024 > 1.95 = \mu_{\bar{w}}.$$

Though an equivalent pricing measure has to exist to prevent arbitrage opportunities, one also may achieve an arbitrage–free pricing asset in the GEI–model which is negative on essential events if markets are incomplete. Even if the Bond is spanned the projection $\pi$ of a positive pricing density $\pi + \eta$ onto the market span might not stay positive.[42] The following theorem describes the restrictions the no–arbitrage condition of the GEI–model imposes on $\pi$ for the CAPM.

**Theorem 68 (No–Arbitrage and Monotonicity)** *The following assertions hold for every CAPM–(EQU) in which at least one consumer is not satiated:*

***(NoArb)*** *The price system is free of GEI–arbitrage, if for all $\bar{r} \in RER$ there is a $\eta \in X^{\perp}$ such that*

$$0 > (\bar{r} - \mu_{\bar{w}})\,\mathbf{1} + \bar{w} + \eta \qquad P - a.s.$$

***(PosPi)*** *The pricing asset assigns a strictly positive value to every essential event if $\forall \bar{r} \in RER:$*

$$0 > (\bar{r} - \mu_{\bar{w}})\,\mathbf{1}_X + \bar{w}_X \qquad P - a.s.$$

***(MPref)*** *The equilibrium allocation makes agents not satiated in tomor-*

[42] Example: The following figures describe a market, a pricing density and individuals' excess demand $z^i$, which may well constitute an equilibrium for a suitable choice of utility functions and endowments. $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $P = (1/2, 3/8, 1/8)$, $x = (-7/8, 1, 1/2)$, $y = (2/3, 22/9, -10)$, $A = (y, \mathbf{1})$, $\mathbf{1} \perp x \perp y \perp \mathbf{1}$, $9\frac{5}{9}(\pi + \xi) = x + y + 9\frac{5}{9}\mathbf{1} > \mathbf{0}, 9\frac{5}{9}\pi = y + 9\frac{5}{9}\mathbf{1} \ngtr \mathbf{0}$, $i \in \{1, 2\}$, $z^1 = 1.29y - 1.72\mathbf{1}$, $z^2 = -z^1$, $z^i \perp \pi$.

*row's consumption, if* $\forall i \in \mathbf{I}, R^I \in RER$ *:*

$$\left[\bar{R}R^i + Cov_X\left(\bar{w}_X, w_X^i\right)\left(1-\mu_{\mathbf{1}_X}\right)\right]\left(\bar{w}_X + \left(\bar{r}-\mu_{\bar{w}}\right)\mathbf{1}_X\right) +$$
$$\left[\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)\right]\ \left(w_{X^\perp}^i + \left(r^i - \mu_{w^i}\right)\mathbf{1}_{X^\perp}\right) -$$
$$\left[\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) - R^i Var_X\left(\bar{w}_X\right)\right]\mathbf{1}_{X^\perp} < 0 \qquad P-a.s.$$

*The three conditions are also necessary if those are only applied to the equilibrium allocations.*

**Proof.** See Appendix 2.11.5. ■

It will be referred to the equilibrium properties stated in the theorem also as conditions, e.g. Condition (MPref) ensures monotonicity. (PosPi) implies (NoArb) because $\eta$ can absorb the difference. Note that $\bar{R}R^i + Cov_X\left(\bar{w}_X, w_X^i\right)$ $\left(1-\mu_{\mathbf{1}_X}\right)$ is positive and $\bar{R}$ is negative since in an equilibrium investors could not be satiated in the Quasi–Bond. Thus, if (SPANNING) and (BOND) is assumed (MPref) is equivalent to (PosPi). In that case sufficiently small absolute risk aversion, i.e. large $|\bar{r}|$, guarantees the monotonicity of preferences. Since the theorem is rather comprehensive it needs further explanation and will be illuminated by three examples.

The most important question is the GEI–viability of CAPM–equilibria, i.e. whether equilibrium prices are free of GEI–arbitrage. The first condition guarantees this property, which differs for incomplete, $X \subsetneq L$, from complete markets, $X = L$. The aggregate risk parameter is changing from $\bar{r}_X$ to $\bar{r}_L$ when markets are completed. Moreover, the orthogonal pricing asset $\eta$ must be zero for complete markets. In special cases where $\bar{r}_X \leq \bar{r}_L$ holds the no–GEI–arbitrage condition for incomplete markets can be checked via complete markets. The next example shows that the opposite way, i.e. no–GEI–arbitrage in complete shown by no–GEI–arbitrage in incomplete markets, is generally not possible.

**Example 69 (Arbitrage by completing the market)** *Suppose one changes the sum of initial endowments of Example 65 to be* $\bar{w} = (100, 1, -1)$*. Then the pricing asset changes to* $\pi \approx (-1.0435, 1.019, 1.061)$*. If the* *security structure consists of the riskless and the market asset only, arbitrage is not possible though the first state price is negative. To see this, add* $(.8081, -.4296, 10) \in X^{\perp}$ *approximately two times to the pricing asset to get a positive pricing density. Thus, there is a positive pricing density ruling out GEI–arbitrage by Condition (NoArb). With a third, non–redundant asset the null space of the original asset matrix becomes empty, but the pricing asset stays the same. In this case GEI–arbitrage is possible again since the Arrow security paying one unit in state one is traded at a negative price.*

No–GEI–arbitrage demands positive prices for spanned positive pay–offs. But it does not imply monotonicity. At investors' optimal demand $\hat{x}$ their utility differential with respect to consumption $\hat{x}(\omega)$ in states $\omega$, which is

$$\xi^i(\omega) := \frac{d_{\hat{x}(\omega)} U^i(\hat{x})}{dP(\omega)},$$

projected on $X$ has to be a positive multiple of the pricing asset:

$$\xi^i_X = \lambda\pi, \ \lambda \in \mathbb{R}_{++}.$$

Thus, by no–GEI–arbitrage agents' preferences must be monotone in *all traded* positive pay–offs, but not necessarily monotone in *all* positive pay–offs in equilibrium. In other words, monotonicity in $X^{\perp}$ can be destroyed by a suitable choice of the utility functions and $w^i_{X^{\perp}}$ although monotonicity in $X$ is assured by Condition (NoArb). The third condition (MPref) ensures general monotonicity with respect to all positive pay–offs, so that $\xi^i > 0$ $P - a.s.$ holds, because both are equivalent inequalities. Condition (PosPi) is not that important itself. At best it simplifies the validation of the first condition. A positive pricing asset does actually not imply monotonicity, since it determines only $\xi^i_X$ and not $\xi^i$.

It should be emphasized that all conditions are endogenous properties, because $\bar{r}$ generally depends on the particular allocation. Only if $r^i$ is linear in mean and standard deviation, $\bar{r}$ adds up to an exogenous constant in equilibrium. This holds for (QUADU), for which the following example considers the Condition (MPref) more accurately.

**Example 70 (Positive State Prices with QUADU)** *Let the utility functions satisfy (QUADU), so that $\bar{r} = \mu_{\bar{w}} - \sum_{i\in\mathbf{I}} \frac{1}{\varphi^i}$ in equilibrium. In the case of complete markets the Condition (NoArb) transforms to:*

$$\sum_{i\in\mathbf{I}} \frac{1}{\varphi^i} > \bar{w} \qquad P-a.s. \tag{Q–NoArb}$$

*Observe that $\bar{r}$ does not change between equilibria for different asset structures. Thus, this condition is also sufficient to guaranty no–GEI–arbitrage in any incomplete market structure. If additionally (SPANNING) and (BOND) is assumed, this condition is also sufficient for monotonicity and a positive pricing asset, even in incomplete markets.*

In the literature two other sufficient conditions are known to ensure monotonicity in a CAPM–equilibrium with (QUADU). The first one is well known [see for instance (Magill & Quinzii, 1996) p. 180 last sentence of the last paragraph but one or (Geanakoplos & Shubik, 1990)]:

$$\forall i : \frac{1}{\varphi^i} > \bar{w} \qquad P-a.s. \tag{MQ}$$

It is sufficient to guarantee monotonicity whenever consumption and endowments are non–negative. The second condition is due to (Pilgrim, 1998):

$$\forall i : \frac{1}{\varphi^i} > w^i \qquad P-a.s., \tag{P}$$

which presupposes (BOND) and (SPANNING). The first condition implies the second if individual endowments are relatively small compared with total endowments, e.g. if $w^i \geq 0$ or $\bar{w} \geq w^i$. If one assumes (BOND) and (SPANNING) both inequalities are stronger than the one given in the example. This is because Condition (Q–NoArb) is equal to the aggregation of the last inequality and because one has $\sum_{i\in\mathbf{I}} \frac{1}{\varphi^i} > \frac{1}{\varphi^i}$. If positivity of consumption is required the Condition (MQ) is indeed sufficient for monotonicity in equilibrium, but it is unfortunately very strong. *Without* positivity constraint for the first and (SPANNING) for the second condition there are examples such that preferences violate monotonicity although the inequality (MQ) or (P) is satisfied. Thus, even though Condition (MPref) is not as simple as the conditions (Q–NoArb), (MQ) and (P), it is generally the weakest possible, provided that the non–negativity constraint is not binding if imposed.[43] This property of Condition (MPref) is demonstrated in the next example by three different economies.

**Example 71 (Restricted Constraints)** *Consider two consumers* $i \in \{1, 2\}$ *having (QUADU) with risk parameters* $\varphi^1 = 1/11$ *and* $\varphi^2 = 10/101$. *The environment and the asset structure is shown in the following table:*

| $\Omega$ | $P$ | $\mathbf{1}$ | $A_2$ | $A_3$ |
|---|---|---|---|---|
| 1 | 1/2 | 1 | −7/8 | 2/3 |
| 2 | 3/8 | 1 | 1 | 22/9 |
| 3 | 1/8 | 1 | 1/2 | −10 |

*Only the riskless asset and the first risky asset* $A_2$ *shall belong to the marketed subspace. Note that all assets are in pairs orthogonal.*
*In the second and third table endowments, the pricing asset, trader* 1*'s net supply* $z_X$, *the equilibrium allocation (all rounded figures) and the sign of the monotonicity condition (MPref) of Theorem (68), denoted by* $M^i$, *are given for two different economies. In both cases arbitrage opportunities do not exist, since the pricing asset is positive by Condition (PosPi), i.e. the market portfolio,* $\bar{w}_X$, *is smaller than the sum of reciprocal risk parameters,* $\bar{r} - \mu_{\bar{w}} = 1/\varphi^1 + 1/\varphi^2 = 21.1$.
*In the first equilibrium (see the following table) both consumers have reciprocal risk parameters larger than each entry of aggregated endowments, i.e. (MQ) holds. Endowments are strictly positive so that Condition (P) is valid as well.*

| $w^1$ | $w^2$ | $\bar{w}_X$ | $\bar{w}_{X^\perp}$ | $\pi$ | $z_X$ | $x^1$ | $x^2$ | $M^1$ | $M^2$ |
|---|---|---|---|---|---|---|---|---|---|
| $0.89\bar{3}$ | $1.79\bar{3}$ | 2.8 | $-0.11\bar{3}$ | 1.17 | 1.39 | 2.28 | 0.40 | − | − |
| $9.44\bar{2}$ | $0.44\bar{2}$ | 10.3 | $-0.41\bar{5}$ | 0.81 | −2.20 | 7.24 | 2.64 | − | − |
| 0.100 | 9.900 | 8.3 | 1.700 | 0.91 | −1.24 | −1.14 | 11.14 | − | + |

[43] For a binding positivity constraint the following inequality ensures monotonicity: $x^{*i} - \mu_{x^{*i}}\mathbf{1} + r^{*i}\mathbf{1} < 0$ $P - a.s.$, see the proof of Theorem (68).

*The second agent has much endowments in the third state, much less in the first and the second state. His/her marginal utility at $w^1$ is highest for the first, most likely state and lowest for the third, least likely state. Agent 2 prefers more consumption in state 2 than in state 3. Since asset 3 is not traded, agent 2 is not able to share his/her 'disliked' endowments in the third state with agent 1. But agent 2 is willing to accept more of state three and less of state one consumption to get more consumption in state two. The first investor consumes a negative amount in state three, which makes consumer 2 violating monotonicity in that state, i.e. $M_3^2 > 0$.*
*In the second equilibrium (see next table) the individual risk parameters are larger than individual endowments, so that (P) holds but not (MQ). Now consumption is strictly positive but non–monotonicity is detected by (MPref) and not by (P).*

| $w^1$ | $w^2$ | $\bar{w}_X$ | $\bar{w}_{X^\perp}$ | $\pi$ | $z_X$ | $x^1$ | $x^2$ | $M^1$ | $M^2$ |
|---|---|---|---|---|---|---|---|---|---|
| $0.17\bar{3}$ | 2.540 | 3.39 | $-0.68$ | 1.16 | 1.39 | 1.56 | 1.15 | $-$ | $-$ |
| $6.80\bar{2}$ | $1.31\bar{3}$ | 10.61 | $-1.30$ | 0.82 | $-2.20$ | 4.60 | 3.51 | $-$ | $-$ |
| 10.90 | 9.900 | 8.69 | 5.31 | 0.91 | $-1.24$ | 9.66 | 11.14 | $-$ | $+$ |

*If investor 2's endowments are reduced to 8 in the third state preferences are monotone at the equilibrium allocation, i.e. $M_3^2$ becomes a minus. However, this is not detected by Condition (MQ), since reciprocal risk parameters are then still smaller than aggregated endowments.*

Altogether it is not surprising that state prices may be negative in a CAPM–equilibrium. Agents perceive just two *aggregated* goods, mean and variance, on which their utility functions depend positively and negatively, respectively. Thereby it is always possible to manipulate the economy in such a way that non–monotonicity appears in all equilibria for arbitrary mean/variance utility functions.[44]

[44] Keep $\rho(\mathbb{E}) = P(\mathbb{E})\, x^i(\mathbb{E}) > 0$ constant for some $\mathbb{E} \in \mathcal{F}$, while $x^i(\mathbb{E})$ is increased. Mean and variance are independently effected by $\rho(\mathbb{E})$ and $x^i(\mathbb{E})$. Hence one could choose both values in such a way that trader $i$ dislikes to get more of $x^i(\mathbb{E})$. This can always be achieved by a suitable choice of (ENV), (MARKET) and endowments.

It is obviously not a realistic situation when agents wish to discard money in particular states of the consumption period. But exactly this is the case in Example 71 investor 2 wishes to do in state three. On the other hand, a fundamental axiom of orthodox economic theory is to demand every agent to choose the best alternative regarding his/her preferences and capabilities. To overcome this empirical paradox *within* the CAPM[45] but without assumptions like proposed in Theorem 68 two modifications of the decision problem are briefly discussed. The first allows for free disposal and the second one changes preferences to be lexicographic. Afterwards, alternative risk measures outside the CAPM are considered, which describe the choice problem consistent and more thoroughly.

### 2.9.1 CAPM with Free Disposal

Suppose agents differentiate between decision utility (ex ante) and received utility (ex post) just because of simplification, idleness or ignorance. Since received utility does not influence the allocation, because it is irrelevant for the decision, it is not explicitly modelled. Due to their decision utility traders would dispose money into the wastebasket in states of the world in which they are satiated. But as these particular states occur traders recycle the 'saved' money from the bin. Thus, when Assumption (AGENTS) shall allow free disposal agent $i$'s decision problem has to be re–defined in the following way:

$$\max_{\left(y^i,\theta^i\right)\in\mathbb{B}^i(q),x\in L} U^i(x) \ \text{ s.t. } \ x\leq y,$$

whereas agent $i$ actually consumes $y^i = A\theta^i + w^i_{X^\perp}$ tomorrow. The definition of CAPM–equilibria remains unchanged but with regard to $\bar{y}$ and $\bar{\theta}$.

**Lemma 72 (Zero–priced Bundles)** *Suppose $(x^i, y^i) \in L\times\mathbb{B}^i(q)$ is trader $i$'s optimal choice with respect to prices $q$. Then $q(z^i) = 0$ with $z^i = y^i - x^i_X > x^i_{X^\perp}$ holds if $i$ is not satiated in the Quasi–Bond.*

[45] That is to say without replacing variance as the risk measure.

**Proof.** If the price of $z^i$ were positive (negative) $i$ would sell (purchase) a small amount of the 'useless' but also 'harmless' $z^i$ to buy the Quasi–Bond. ■

**Remark 4** *Note that for agent $i$ there is neither a reason to give $z^i$ away, because it could be useful in the future, nor to get any $z^j$ from another trader $j$ for free, because $j$ has no reason to give it away. Thereby a weakly monotone utility function is constructed.*

Since strictly positive pay–offs in the marketed subspace may still be valued with a zero price free disposal does not solve, only mitigates the problem of non–monotonicity. Moreover, for a rational individual the differentiation between decision and received utility must be criticized to be not time consistent in a frictionless world.

### 2.9.2 Lexicographic Preferences

A second suggestion to solve the problem of non–monotonicity proposed to the author has been lexicographic preference relations within the CAPM. Individuals shall prefer a consumption plan to another if it is larger or if – provided that none of both is larger – it yields a higher mean/variance utility:

$$\begin{gathered} x \succsim y :\Leftrightarrow \\ \langle x \geq y \ P - a.s. \text{ OR} \\ \left[P\left(x<y\right) \cdot P\left(x>y\right) > 0 \text{ AND } U(x) \geq U(y)\right]\rangle . \end{gathered}$$

Unfortunately, the lexicographic preference relation is neither continuous nor transitive[46]. The latter, indispensable requirement can be repaired by eliminating all $\geq$–dominated consumption plans from the budget set first, but discontinuity remains and could cause non–existence of equilibria:

[46] It is possible to choose $U$ and $P$ such that: $x \succ y \succ z \succ x$ with $x = (1,2), y = (100,1), z = (2,1000)$. The idea is that variance is increasing from $x$ over $y$ to $z$. The increase and the variance–aversion has to be strong enough.

**Example 73** *Suppose agents eliminate all $\geq$–dominated consumption plans from their budget set first and then choose out of the remaining consumption plans due to their mean/variance utility function. Imagine a complete asset market with Arrow–securities for a finite $\Omega$. Let traders have in one particular state, with a low but positive probability, relatively large endowments. Hence, when agents choose due to the mean/variance criterion all of them shall wish to get rid of that particular Arrow–security which pays in the state where all of them are satiated. Thus, an equilibrium settles down only if the price of that particular Arrow–security is negative (zero, if free disposal is allowed). But for a non–positive price the choice set explodes in that particular Arrow–security, so that the satiation point in that particular Arrow–security is no longer contained in the choice set. This contradicts the existence of an equilibrium.*

### 2.9.3 Conclusions

To all appearances non–positive equilibrium state prices are an immanent problem of mean/variance utility functions. Only the three conditions pointed out in Theorem (68) guaranty arbitrage–free prices, a positive pricing asset, and the monotonicity of consumers' preferences in an equilibrium. These endogenous restrictions ensure the reasonability of CAPM–equilibria just from a General Equilibrium point of view, which is not really satisfactory.

Nevertheless, the CAPM's simplicity and illuminating power led to its broad usage in practice. An important field of future research in applied finance will be the development of a comparably simple risk measure without incurring the problem of monotonicity. The concluding chapter throws some light onto alternative risk measures which are currently in the discussion. Since a further equilibrium analysis of these alternatives would be far beyond the scope of this chapter, the discussion remains mainly on the surface. Moreover, the following investigation concentrates more on the practical aspects of portfolio choice for large financial institutions, where particularly mathematical choice problems are applied.

## 2.10 Coherent Measures of Risk in Portfolio Choice – An Outlook

As it has been pointed out in Section 2.9 the non–monotonicity of CAPM–utility functions is in fact caused by a major deficiency of the risk–measure, the variance of returns: Variance is a symmetric risk measure, i.e. irrespectively of the sign it penalizes quadratically the pure distance of returns from mean.[47] Thus, variance displays no sensitivity with regard to a negative skewness of returns. This allows to find for any $\mu/\sigma$–utility function a probability space, endowments, and assets causing non–monotonicity on certain subspaces of the market: A possible gain in mean is more than outweighed by an increase in variance for certain consumption plans. Nevertheless variance is still the most popular criterion for portfolio selection in applied finance. Even if this negative property has been recognized, the biggest advantage of variance is *simplicity*, which implies *intuitive* results and makes it thereby easily *communicable* and *usable* for investors.

The intention of this outlook is to sketch an alternative to variance within a broader context of asset/liability–management in banks. Since variance and the CAPM are sometimes regarded as old–fashioned in finance the following discussion will refer to more topical issues of risk management. The general story of this section is summarized by the question: What is a *"good" risk measure* for *institutional investors* in *practice*? Beside variance the so–called Value–at–Risk plays a similar important role in portfolio optimization. Actually, Value–at–Risk satisfies a different purpose than variance. Whereas variance shall represent investors' attitude towards risk, Value–at–Risk should mainly secure stake–holders' interests, since investors' bankruptcy cannot be

[47] To avoid the penalization of gains one could think of replacing variance by semi–variance: $\sigma_-^2 := E(\min[x-\mu,0]^2)$, in a $\mu/\sigma_-^2$–utility function. At first, actually, Markowitz preferred semi–variance as he pointed out in his speech at the Nobel Prize awarding 1990 (see (Reichling, 1999)). While this choice repairs monotonicity it is not capable to mirror risk aversion with respect to gains. Since $\sigma_-^2$ is only weakly convex, mean preserving spreads of $x$ where $x > \mu$ would not affect the utility, i.e. it implies risk neutrality for $x > \mu$. This could cause multiple solutions of the portfolio problem. Thus, semi–variance is either not 'perfectly' applicable to the portfolio choice problem.

ruled out. Sometimes this important difference is not appropriately honored, especially if Value–at–Risk is unfoundedly considered as a risk measure for investors. Thus, dealing only with one of both risk measures the picture of real portfolio optimization would become incomplete. Variance and Value–at–Risk have in common that both are inconsistent with monotonicity. Therefore a further step will be made by introducing two promising alternative risk measures, which go beyond variance and Value–at–Risk: *Worst Conditional Expectation* (abbreviated WCE), proposed by (Artzner *et al.*, 1998) to replace Value–at–Risk, and *Weighted Value–at–Risk* (abbreviated WVaR), suggested by (Aspandiiarov *et al.*, 1998) instead of variance as well as (!) Value–at–Risk. These risk measures satisfy certain consistent and convincing axioms, which classify them as *coherent*. Monotonicity belongs to the axioms, which excludes variance and Value–at–Risk from the class of coherent risk measures. Weighted Value–at–Risk is a flexibly parameterized risk measure even though it cannot cover all preferences towards risk. The WVaR can easily be integrated in an utility function. The Worst Conditional Expectation is indeed a generalization of the Lower Partial First Moment.

The grounds of General Equilibrium Theory are abandoned here in favour of more heuristic arguments. This is because an equilibrium model with explicitly coherent risk measures and the possibility of investors' bankruptcy is a complete research project by itself.[48] However, it will be shown that the distinct purposes of both new risk measures are sometimes misinterpreted. Moreover, it will become clear that although a public intervention in the risk management of banks is economically justified Value–at–Risk alone is not the right tool for it. For that reason the more realistic situation is considered in which an investor's bankruptcy is taken into account as a possible event. Thus, the bond is no longer riskless, but its pay–offs are endogenously determined by investors' leveraged portfolio strategy. The partial equilibrium between the investor (the borrower) and the lender determines the risk adjusted interest rate on foreign capital.

[48] A full General Equilibrium model with bankruptcy and incomplete markets has been developed by (Dubey *et al.*, 2000).

The reader should notice that WCE and WVaR – though theoretically founded – are chosen with the view of illuminating a real world portfolio problem and not because these are regarded as final and unique recommendations. From a decision–theoretical point of view the question for a coherent measure of risk has no definite answer. The set of axioms which are discussed in this section are normative in that they postulate a consistent behaviour under risk gained from introspection. In the literature different foundations of choice under risk and uncertainty exist.[49] The normative axiomatics of subjective expected utility theory, see (Anscombe & Aumann, 1963), has been generalized to Choquet Expected Utility, see (Schmeidler, 1989) and (Chateauneuf, 1994).[50] This new approach avoids Allais' and Ellsberg's decision paradoxes and covers also decision problems under uncertainty. Beside those normative approaches there is a positive descriptive theory – called Prospect Theory – which investigates by experiments agents' real investment behaviour under uncertainty, see (Tversky & Kahneman, 1992). Its tenor is that people value gains by a concave and losses by a convex, but steeper utility function, while attaching proportionally more weight to events with low probability. (Wakker & Tversky, 1993) founded the empirical Prospect Theory by an axiomatics, which is based on Choquet Expected Utility Theory. While from Evolutionary Finance there is strong support for the logarithmic utility function, since it maximizes the expected growth rate of wealth. (Blume & Easley, 1992) have shown that investors who deviate from log–utility die out in the long run measured by their share on total wealth.

This section is organized as follows. In the first subsection the concept of Value–at–Risk with its intention, definition, practical implementation, and properties is described. Non–monotonicity of variance and Value–at–Risk is the reason for introducing the coherent risk measures WCE and WVaR in

[49] The modern distinction between risk and uncertainty is as follows: Under uncertainty the decision maker is unable to form a unique probability prior about possible events, i.e. they consider more than one possible probability–scenario. Unique objective as well as unique subjective probability priors are subsumed to decision under risk.

[50] Essentially, the axiom of independence of prospects has been weakened to co–monotonic independence.

the subsection thereafter. Then the portfolio choice problem is consistently re–formulated with the help of WCE and WVaR. Finally the conclusions summarize the results of this section.

### 2.10.1 Value–at–Risk

Before the coherent risk measures are considered a digression about the so called Value–at–Risk (abbreviated VaR) is appropriate. It has similar disadvantages as variance, but has been introduced to applied finance for a different purpose than variance. Knowing VaR is necessary to understand the need for those mentioned and further new developments regarding the measurement of risk.

**Excursion about Value–at–Risk:**

At a quick glance the following paragraphs explain, why Value–at–Risk became a very popular risk measure, how it is defined analytically, how it is used by regulators to enforce a risk management in banks, and why it should be used carefully in portfolio selection.

*Historical background:* In the past decades some major crisis of banks and corporates distressed financial markets. A recent example has been BARING's Bank Inc. and the Bayrische Hypotheken Bank AG. In the sector of corporates Metallgesellschaft AG was one famous example and recently Holzmann AG got into trouble. In all cases the risk management was inappropriate with regard to the business undertaken by the firms. At BARING's the risk of criminal betray was almost ignored, Metallgesellschaft failed because of a hedging strategy which incorporated high liquidity risk and the Bayrische Hypotheken Bank underestimated the risk of investments in real estates in eastern Germany after the unification. Holzmann had almost no risk management for their real estate development projects. BARING's, Metallgesellschaft and Holzmann caused economic costs of restructuring, the latter two also forced creditors to renounce parts of their loans. Whereas in case of the Bayrische Hypotheken Bank it seems to outsiders that the hidden losses were partially transferred onto foreign shareholders' shoulders in a merger with the Bayrische Vereinsbank AG.

In perfect markets there is no need for legislative interventions in the risk management of public firms to improve the efficiency of markets. Actually, markets are imperfect. Asymmetric information, high transaction costs and bounded rationality are inevitably leading to incomplete contracts between owners respectively stake holders and a company's management. Incomplete contracts are mirrored in the difficulty to enforce contractors' interest, which are *implicitly* expressed in the contracts, by a first best incentive scheme accompanied by a monitoring of agents' actions. For example a change in a company's risk profile might increase the probability of bankruptcy with the consequence that the risk premium of senior debt does not compensate the risk of default anymore.[51] A single creditor is hardly in the position to impose a risk management in the credit contract to monitor and ensure a certain risk profile of the debtor.[52]

Moreover, these market imperfections are accompanied by internal imperfections in the principal–agent relation. A company's responsible agents, i.e. traders and portfolio managers in this context, are generally exposed to adverse incentives. Incentive payments of traders and managers have usually option character. Their bonuses (despite non–financial incentives) depend in an asymmetric way positively on gains or volume and less – if ever – on losses. Since the value of an option increases in taking higher risks (adjoint with higher chances), the incentives are possibly concurrent with owners' and often adverse to stake holders' risk preferences.[53]

[51] A defaultable bond can be duplicated by a riskless bond and a short–put on the company's assets with strike price equal to the amount of outstanding debt [see (Merton, 1976)]. The price of the put–option rises with higher risk, i.e. the price of the outstanding debt shrinks. One should distinguish between changes of the default probability caused by the conventional business risk or caused by a strategical change of a firm's risk/return strategy. The latter source of risk may counteract stakeholders' interest implicitly agreed on and mirrored in the contracted credit spread. Important is not only the probability of default itself but its change in time.

[52] To impose a risk–constraint might be in the interest of owners as well, but it is argued here only in favour of stakeholders, since they are external to the firm and deserve more protection thereby.

[53] Zero risk is generally not in the interest of stakeholders as well because chances – accompanied by risk – increase expected growth and reduce thereby in the long run the

Last but not least high financial losses are usually accompanied by high economic costs, for instance unemployment. Those arguments justify regulators' intention to limit financial risks by imposing capital requirements appropriate to banks' risk profile.[54]
A different approach, which is also followed by regulators, is to increase the transparency of banks' risk profiles and risk management. This enables the market to take the reputation of banks' risk management better into consideration. The disclosure of meaningful risk characteristics enforces a strong incentive for banks to limit the risk of shortfall, because the conditions of re–financing worsens with higher risk. Nevertheless, the market requires in this case at least one commonly accepted, *coherent* risk measure as well.
From major financial crisis of banks and corporates Value–at–Risk gained high popularity, since it suits nicely for measuring the risk of large shortfalls. The Value–at–Risk is the expected amount of capital, which at the worst can be lost at a given confidence level for an assumed probability distribution of portfolio returns. If consequently applied the VaR–concept helps getting aware of and reducing the risk of extreme shortfalls.

*Definition:* Whereas variance displayed its strength in portfolio selection, but was unable to discriminate between positive chances and negative risks, the Value–at–Risk covers especially the risk of shortfall. The $VaR_\alpha^P$ is defined as the expected magnitude of loss the portfolio returns $x$ fall at least below the mean with a probability of $\alpha$–percent. That is to say $VaR_\alpha^P(x)$ is the solution $V$ of the following equation:

$$P(x \leq \mu_x - V) = \alpha,$$

for a continuous probability distributions $P$ and a confidence level $\alpha \in (0,1)$. The $\alpha$–quantile $\mu_x - V$ is called the $\alpha$–Risk Limit of $x$. The interpretation is, that if a portfolio has an expected return of $\mu_x$ then an amount $V$ of economic

probability of default. Moreover, zero risk implies a zero credit spread, which lenders could achieve also by purchasing government bonds.

[54]Corporates are not concerned yet. In case of banks see the publications of the Basle Committee of Banking Supervision (www.bis.ch).

capital, which covers first losses, secures lenders' capital with a probability of $\alpha-$percent given the probability distribution of returns $P$.[55]

For discrete distributions of $x$ the $VaR_{\alpha}^{P}$ is generally not unique. Let $\alpha^{-} \leq \alpha \leq \alpha^{+}$ be the nearest neighbours of $\alpha$ for which there is a solution to the previous equation.[56] Then the $VaR_{\alpha}^{P}(x)$ lies between the last point of support before and the first after the probability distribution of $x$ crosses the $\alpha-$percentile:

$$VaR_{\alpha}^{P}(x) \in \left[VaR_{\alpha^{+}}^{P}(x), VaR_{\alpha^{-}}^{P}(x)\right].$$

For discrete distributions the VaR–concept lags a definition, which guarantees uniqueness and continuity in $\alpha$. The following arbitrary choice is suggested:

$$VaR_{\alpha}^{P}(x) = (1-\beta)\, VaR_{\alpha^{+}}^{P}(x) + \beta VaR_{\alpha^{-}}^{P}(x),$$

where $\beta$ solves

$$\beta f\left(\alpha^{+} - \alpha^{-}\right) = f\left(\alpha^{+} - \alpha\right)$$

for a continuous (or differentiable), positive and monotone function $f$ with $f(0) = 0$. This formula defines a unique and continuous VaR–number as a convex combination of the extreme alternatives.

*Regulators' implementation:* The VaR–concept has successfully been implemented by regulators to supervise the market risk of banks' trading books.[57] This is because the market risk is seen as an important source of shortfall risk and is the most tractable kind of risks.[58] The Basle Committee of Banking

[55] It should be stressed that a "lender" is not necessarily an outsider to the bank, it could be a different unit within the bank, if the portfolio management is a separate profit centre.

[56] If $\alpha$ is not in the range of $P$ let $\alpha^{-} = \alpha^{+}$ hold for only one nearest neighbour of $\alpha$.

[57] This concept does not apply to all banks, but to the trading books of large international banks. Market risk covers the risk of negative price movements of liquid assets. A bank collects in its trading book all assets and liabilities bought for short term speculations on positive price changes.

[58] Generally excepted is the following classification of risks: 1) market risk of the trading book (e.g. negative price movements), 2) market risk of the investment book (e.g. interest

Supervision[59] is planning to extend VaR–similar concepts onto more fields of risk, see (Basle Committee of Banking Supervision, 2001). The broad idea[60], which is only partly realized yet, is to demand from financial institutions an allocation of economic capital to all different sources of risk.[61] The capital allocated display the maximal amount the bank must be willing to risk on that particular category or portfolio with respect to a predefined loss–percentile and time–horizon. The loss–percentile mirrors the risk–profile of a bank. For example a bank with a Standard & Poor's rating of A probably has a different – less rigorous – loss–percentile than an AAA–bank. If a loss in one risk category exceeds the associated economic capital, it goes 'hypothetically bankrupt'. The bank itself should be still alive, since it is very unlikely that the economic capital of all categories are exhausted by a negative shock in only one category or that exhausting shocks in more than one category occur at the same time despite the fact that a bank has often much higher capital reserves than actually allocated. Suppose the bank calculates by a certain theoretical model its VaR for a horizon of one trading day and a 5%–percentile. Then the number of days for which the actual day–to–day losses exceed the associated capital should be about five out of hundred trading days.[62] If significantly more heavy losses occur, the bank is asked by regulators to put a multiple of the model's VaR for this source of risk as economic capital aside.[63] The bank should then refine its theo-

rate risk for illiquid loans), 3) default risk of counterparts (e.g. credit risk), 4) operational risk (e.g. fraud), and 5) operative risks (e.g. delays in discharging settlements).

[59]The Basle Committee of Banking Supervision is an international forum founded by the G10–countries in 1974 to coordinate the national acts about the regulations of banks. In January 2001 the committee proposed a new Capital Accord, which considerably extends the first Capital Accord from 1988.

[60]**Not** described here is any national legislative translation, but only what is considered as a good application of the VaR–concept in an ideal world.

[61]Note that an allocation of economic capital does usually not imply a physical association. In general any association between non–collateralized assets and liabilities is hypothetical and thereby arbitrary. However, the allocation of economic capital might be finer than the classification of risks due to organizational purposes, for instance with regard to separate units, country wide credit portfolios etc.

[62]The test–procedure capturing this property of the estimated VaR is called "backtesting".

[63]Regulators check the consistency of banks' risk models and risk management. They

retical model to make better VaR predictions, since, if economic capital is fully allocated, more risk is costly, marginally in an amount of the internal shadow price of owners' economic capital.[64] Beside the capital requirements the VaR–concept caused a fast development in banks' risk management. A bank must have a theoretical model describing their risk, even if it is only a rough estimate. Thereby the risk management becomes *aware* of sources of risk. For many banks this has been or is still a big challenge.

*Shortcomings:* The major advantage of VaR is – similar to variance – its simplicity. But the VaR has some serious limitations, which exclude it from being a meaningful criterion for measuring risk in portfolio optimization. First of all VaR is *not quasi–convex* (see the second of the following tables) and it is only *weakly monotone* with respect to mean preserving spreads. Moreover, if not defined carefully as above, VaR is *discontinuous* for discrete distributions. All this implies that a portfolio optimization problem might not have a unique solution or solutions behave discontinuously ("instable") with respect to small changes in the underlying distribution of returns. Thereby VaR is only meaningful as an additional constraint in well defined portfolio optimization problems.

Consider two random variables $x$ and $y$ with probability distributions $F_x$ and $F_y$, respectively. Then $y$ is a mean preserving spread of $x$ if

$$\mu_x = \mu_y$$

and

$$\forall u \in \mathbb{R} : \int_{-\infty}^{u} F_x(v)\, dv \leq \int_{-\infty}^{u} F_y(v)\, dv.$$

The last condition means that $F_y$ has more weight in the tails than $F_x$. In other words, $x$ second order stochastically dominates $y$. A meaningful risk

value the success of banks' internal models by *backtesting* based on the empirical distribution of returns for a certain history.

[64]Usually a shadow price between 15% and 30% return per annum before taxes is demanded for banks' economic capital.

measure should detect a mean preserving spread. Variance does and any concave vNM–utility function detects it as well. The latter would attach less utility to $y$ than to $x$, for details see (Copeland & Weston, 1979).
The point is that Value–at–Risk should be sensitive to the risk of shortfall below a certain Risk Limit. But even though the severity of a loss can increase by a mean preserving spread the VaR does not necessarily detect it, because the VaR is insensitive to the conditional distribution of returns on both sides of the Risk Limit when the mean is constant. The following table shows an example:

| $P$ | 80% | 5% | 5% | 5% | 5% | $\mu$ | $VaR^P_{10\%}$ | $WCE^P_{10\%}$ |
|---|---|---|---|---|---|---|---|---|
| $x$ | 3.7500 | $-60$ | $-10$ | 0 | 10 | 0 | 10 | $-7/2$ |
| $y-x$ | 0 | $-10$ | 0 | 0 | 10 | 0 | – | – |
| $y$ | 3.7500 | $-70$ | $-10$ | 0 | 20 | 0 | 10 | $-8/2$ |

$y-x$ and $x$ are comonotone, which makes $y$ riskier. The loss severity increases in the second state from $-60$ to $-70$, but the VaR remains the same. Thus, whenever the Risk Limit, which is in this example equal to the VaR due to $\mu = 0$, does not alter by a mean preserving spread neither the VaR does change. The Worst Conditional Expectation, defined in the sequel, is the lower partial first moment below the Risk Limit (mean below the 10%–quantile). It changes from $-7/2$ for $x$ to $-8/2$ for $y$. Thus, the WCE is superior in measuring the severity of losses, but – as will be shown below – is not capable to measure *risk* below the risk limit.

| $P$ | 80% | 5% | 5% | 5% | 5% | $\mu$ | $VaR^P_{10\%}$ |
|---|---|---|---|---|---|---|---|
| $x$ | 3.7500 | $-60$ | $-10$ | 0 | 10 | 0 | 10 |
| $y$ | 5.6250 | $-10$ | 0 | $-20$ | $-60$ | 0 | 20 |
| $\frac{1}{2}(x+y)$ | 4.6875 | $-35$ | $-5$ | $-10$ | $-25$ | 0 | 25 |
| $\frac{1}{10}(9x+y)$ | 3.9375 | $-55$ | $-9$ | $-2$ | 3 | 0 | 9 |

Table: Neither quasi–concavity nor quasi–convexity holds

A more methodological criticism of VaR is that a capital requirement does not really mirror stake holders' interest as it should be the intention of the supervisor. This issue will be clarified at the end of this section.

**Remark 5** *Sometimes the Value–at–Risk is defined as the Risk Limit itself, for instance in (Aspandiiarov* et al.*, 1998) and (Artzner* et al.*, 1998). But the Risk Limit is no longer interpretable as the capital requirement. Moreover, it is often not workable if used as a constraint, because a solution to the portfolio optimization of mean given a minimal Risk Limit might not exist. For example if the returns $x$ are normally distributed the Risk Limit as well as the mean increase with increasing volume whenever*

$$\mu_x - 1 + \sigma_x \Phi^{-1}(\alpha) > 0,$$

*where $\Phi$ is the cumulative standard normal distribution function (the price of the portfolio could be zero if $q(x) = 1$).[65] The VaR as defined in the text is positively homogenous in volume. Thus, increasing volume implies increasing risk, such that a portfolio optimization problem has always a solution if VaR is constrained.*

### 2.10.2 Coherent Risk Measures versus Variance and VaR

It is not surprising that this dichotomy between two popular risk measures, variance suitable for portfolio selection and VaR applied to shortfall risk, evoke the incentive for researches to find alternatives matching both needs with less inconsistencies and comparable simplicity. But a closer look to the portfolio choice problem reveals that two different risk measures reflect two conflicting interests. On the one hand stakeholders, notably creditors, are interested in a small default probability and – in case of default – a high recovery rate on their exposures (low severity of losses). Their claims are

[65] For spherically distributed returns the so–called safety–first approaches by (Roy, 1952), (Kataoka, 1963), and (Telser, 1956) (see also (Bawa, 1978)) apply the Risk Limit to the $\mu/\sigma$–portfolio selection. But for spherical distributions $\mu/\sigma$–efficient portfolios are also $\mu/\alpha$–efficient since any percentile $\alpha$ is a strictly increasing function of $\sigma$.

in general of fixed interest, so that stakeholders can ignore the distribution of returns above the default–threshold[66]. On the other hand, stockholders are especially interested in the distribution of returns above the default–threshold. In case of default they hardly gain something from the bankrupt's estate, such that the distribution of returns below the default–threshold is of minor interest to them.[67] Since the default–threshold is an endogenous quantity this would not be an equilibrium point of view. Indeed, the conditional distribution of returns below the default–threshold effects the condition of re–financing and has therewith an influence on the overall distribution. Moreover, in a dynamical context the conditional distribution of returns above the default–threshold determines the default–thresholds in the subsequent periods. This mitigates the conflict of interest between creditors and owners, which nevertheless endures. Therefore regulators enhance transparency of risk management to internalize the effects of banks' risk strategy in the market place of re–financing and create incentives for voluntary protections against the risk of big shortfalls thereby.

From the regulators' point of view two risk measures seem reasonable. One risk measure is imposed by the regulator as a constraint to protect creditors' interest, whereas the other risk measure is used by investors in portfolio selection while taking creditors' constraint into consideration. Variance suits better for an optimization criterion whereas VaR helps reducing the risk of large shortfalls. Since both risk measures have their deficiencies the two mentioned alternatives; Weighted Value–at–Risk for portfolio optimization and Worst Conditional Expectation as an optimization constraint, are introduced now.

[66] The amount of total outstanding liabilities is called the default threshold. The distance to default is the amount of economic capital that can be lost before a firm goes bankrupt. It is usually measured in units of standard deviation of firms' asset values.

[67] Whether stockholders can expect any returns in case of default really depends on the law of bankruptcy. If one assumes an economically efficient bankruptcy procedure as proposed by (Aghion *et al.*, 1992) the ownership of residual rights immediately passes to creditors in case of a firm's default on debt. Then it is very unlikely that stockholders might regain anything. In real legal codification the situation is quite different, since some rights remain at the stockholders, which often allows them to compel stakeholders to cede returns.

The most appropriate way is to develope a minimal set of desired consistent axioms a coherent measure of risk should satisfy. From a heuristic point of view the risk measure should be: a single number imposing a natural and complete order on portfolio returns; easily understandable and computable while reflecting one's intuitive understanding of what risk means; and, applicable to all real financial risks. Analytical desired properties are sensitivity to and continuity in changes of the distribution of returns. (Artzner *et al.*, 1998) see in the following axioms the best description of a coherent measure of risk.

**Definition 74** *A measure of risk $\rho : L \to \mathbb{R}$ is said to be* coherent *if it satisfies the following five axioms for any $x, y \in L$ :*

1. *Translation invariance:* $\rho(x + \alpha\mathbf{1}) = \rho(x) - \alpha$ *for* $\alpha \in \mathbb{R}$.

2. *Subadditivity:* $\rho(x + y) \leq \rho(x) + \rho(y)$.

3. *Positive homogeneity:* $\rho(\lambda x) = \lambda\rho(x)$ *for* $\lambda \in \mathbb{R}_+$.

4. *Monotonicity:* $y \geq x$ *implies* $\rho(y) \leq \rho(x)$.

5. *Relevance:* $x \leq 0$ *and* $x \neq 0$ *implies* $\rho(x) > 0$.

*A prospect $x \in L$ is said to be an acceptable risk if and only if $\rho(x) \leq 0$.*

(Artzner *et al.*, 1998) define zero risk as the threshold, which separates acceptable from unacceptable risks. Moreover, the authors derive a correspondence between axioms on acceptance sets and axioms on risk measures.
Translation invariance stresses that more of the riskless asset reduces risk additively. It is intuitive that more of the riskless asset makes more risk bearable, whereas additivity is a simplification. Subadditivity points out the importance of diversification. This axiom means that the sum of risks of two separated portfolios is never smaller than the risk of the joined portfolio. Subadditivity is convincing, since the returns of both portfolios might insure each other. Positive homogeneity is a kind of normalization, which seems to be not necessary, but it emphasizes the simplicity of the risk measure.

Moreover positive homogeneity can be redone in a utility function depending on mean and the risk measure. Monotonicity is the property, which variance is criticized for not to fulfill. It says that strictly higher returns cannot imply higher risk. Relevance demands that purely negative returns must have positive risk, i.e. are not acceptable.
A risk measure satisfies the first four axioms if and only if it admits the following representation (Proposition 4.1 in (Artzner *et al.*, 1998)):

$$\rho(x) = -\inf_{P \in \mathbb{P}} E_P(x),$$

for a set $\mathbb{P}$ of probability distributions on $(\Omega, \mathcal{F})$, which are called "generalized scenarios". Relevance is only satisfied, if the union of supports of probability measures in $\mathbb{P}$ covers all measurable events in $\mathcal{F}$ (Propositions 3.1 in (Artzner *et al.*, 1998)). The authors propose a coherent risk measure they call *Worst Conditional Expectation* (WCE):

$$WCE_\alpha^{\mathbb{P}}(x) = -\inf_{P \in \mathbb{P}} E_P\left(x \middle| x \leq \mu_x - VaR_\alpha^P(x)\right).$$

In the WCE the conditioning event restricts the set of probability scenarios to assign zero probability to returns above the Risk Limit. $-WCE_\alpha^P(x)$ is never larger than the $\alpha-$Risk Limit, which is $\mu_x - VaR_\alpha^P(x)$. That makes $\mu_x + WCE_\alpha^P(x)$ a more conservative risk measure than the Value–at–Risk. Like the Risk Limit the WCE cannot sensibly serve as a capital requirement. For this purpose the sum $\mu_x + WCE_\alpha^{\mathbb{P}}(x)$ would suite better (see next subsection). WCE is advantageous over VaR, since it takes into account the conditional distribution of losses exceeding the Risk Limit.[68] Moreover, by a non–singleton set $\mathbb{P}$ WCE allows also to consider model risk and uncertainty. WCE is not applicable to portfolio optimization, because it neglects the distribution of returns larger than the scenario conditional Risk Limit. Moreover, it is not strictly quasi–convex, which could lead to non–unique solutions in portfolio selection. Indeed, the authors only speak from an acceptance set, the set of portfolios for which $\rho(x) \leq 0$. They leave open the

[68] Actually, WCE is the partial first moment below the Risk Limit.

question, how to rank acceptable portfolios. Therefrom one can conclude that the authors developed a risk measure which is more suitable for regulators' purposes as a *constraint* in portfolio selection.
(Aspandiiarov *et al.*, 1998) propose another set of axioms for a coherent risk measure:

**Definition 75** *A measure of risk* $\hat{\rho} : L \to \mathbb{R}$ *is said to be* coherent *if it satisfies the following four axioms for any* $x, y \in L$ *:*

1. *Risk–Free Condition:* $\hat{\rho}(x + \alpha\mathbf{1}) = \hat{\rho}(x) - \alpha$ *for* $\alpha \in \mathbb{R}$.
2. *Strict Quasi–Convexity:*

$$\hat{\rho}(x) \leq \hat{\rho}(y) \quad \textit{implies} \quad \hat{\rho}(\beta x + (1-\beta) y) \leq \hat{\rho}(y) \quad \forall \beta \in (0,1),$$

*where equality holds if and only if* $x = y$ $P - a.s.$

3. *Strict Continuity:* $\hat{\rho}(x) \to \hat{\rho}(y)$ *as* $x \xrightarrow{d} y$.
4. *Strict monotonicity:*

$$(y - x) \in L_{++} \quad \textit{implies} \quad \hat{\rho}(y) < \hat{\rho}(x).$$

An important distinction between both axiomatics is that the first applies also to uncertainty whereas the second presumes a single subjective probability measure.[69]
The axioms are ordered in this way, because there is a weak relationship between both sets. Note that for $\rho$ subadditivity in conjunction with homogeneity leads to strict convexity, whenever $\rho$ is not additive:

$$\begin{aligned}
\rho(\alpha x + (1-\alpha) y) &\leq \rho(\alpha x) + \rho((1-\alpha) y) \\
&= \alpha\rho(x) + (1-\alpha)\rho(y) \\
&\leq \rho(y) \text{ if } \rho(x) \leq \rho(y) \text{ and} \\
&= \rho(y) \text{ only if } \rho(x) = \rho(y).
\end{aligned}$$

[69] To cover also different scenarios strict quasi–convexity and strict monotonicity should hold for all $P \in \mathbb{P}$.

Subadditivity is on the one hand weaker than strict quasi–convexity, since it allows for additivity. On the other hand, whenever additivity in the first inequality does not hold, subadditivity in conjunction with homogeneity implies convexity, which is stronger than quasi–convexity. But since $\rho$ is applied as an inequality constraint, i.e. $\rho \leq \bar{\rho}$, where $\bar{\rho}$ could be the threshold imposed by the regulator, it preserves weak convexity of the choice set, which is very appropriate to achieve uniqueness with regard to an quasi–concave optimization problem. Strict quasi–convexity should hold for $\hat{\rho}$ to ensure uniqueness, because it is used as a criterion and not as a constraint in portfolio selection. For example, the efficient frontier in terms of risk $\hat{\rho}(x)$ and expected return $\mu_x$ subject to the constraint $\rho \leq \bar{\rho}$ would exist and be unique. The following example illuminates for WCE that subadditivity is indecisive with regard to higher risk when additivity holds.

| $P$ | 80% | 5% | 5% | 5% | 5% | $\mu$ | $VaR^P_{15\%}$ | $WCE^P_{15\%}$ | $WCV^P_{15\%}$ |
|---|---|---|---|---|---|---|---|---|---|
| $x$ | 3.75 | −60 | −10 | 0 | 10 | 0 | 0 | 70/3 | $688,\bar{8}$ |
| $y-x$ | 0 | − 5 | 5 | 0 | 0 | 0 | – | – | – |
| $y$ | 3.75 | −65 | − 5 | 0 | 10 | 0 | 0 | 70/3 | $872,\bar{2}$ |

Again, $y$ is a mean preserving spread of $x$. Since only the payoffs below the Risk Limit change from $x$ to $y$ the WCE stays equal, although $y - x$ adds more risk to $x$ in states two and three. The 'Worst Conditional Variance' in the last column shows it.

Neither of both axiomatics imply strict risk aversion. Risk aversion would demand $\hat{\rho}(y) > \hat{\rho}(x)$ for $y$ being a mean preserving spread of $x$. For the WCE this is only the case if the Risk Limit of $y$ is worse than that of $x$. The weaker condition $\hat{\rho}(x) > \hat{\rho}(0) - \mu_x$, i.e. risk aversion defined by the certainty equivalent for some non–zero $x$, is satisfied for WCE but not necessarily for general $\hat{\rho}$ or $\rho$.

Strict continuity of $\rho$ is indeed implied by homogeneity and monotonicity. Since homogeneity does not hold for $\hat{\rho}$, continuity must be assumed separately. Continuity implies stability of risk minimizing portfolios with respect to small changes in the distribution of asset returns. Strict monotonicity is a desired property for an optimization criterion since it counts for strictly larger returns. Monotonicity over the whole domain of pay–offs is not necessarily required if bankruptcy is concerned. $P$ can be defined as the conditional probability measure of returns above the Risk Limit such that investors do only take into account returns above the default threshold.

The axioms proposed by (Aspandiiarov *et al.*, 1998) seem to be compatible with a wider class of possible risk measures. Accepted, minimal requirements a utility function should satisfy are continuity, monotonicity and quasi–concavity. The only property the risk measure $\hat{\rho}$ adds to these axioms is the risk free condition.[70] However, Aspandiiarov et al. did not derive analytically a class of risk measures from their axioms as it has been done for the first set of axioms. The authors propose an *example*, which they call Weighted Value–at–Risk:

$$WVaR_{a,p}^{P}(x) = \frac{1}{a} \ln \left[ \frac{E_P(\exp(-ax))}{p} \right],$$

where $a > 0$ is the penalizing factor. For a normalizing constant $p$ being equal to one the $WVaR$ satisfies above axioms. Convexity of $\exp(-ax)$ is stronger than the demanded strict quasi–convexity on $L$. The authors show that if $p$ corresponds to the percentile $\alpha$ to which the $VaR_\alpha$ is chosen the Weighted Value–at–Risk is more conservative than the $\alpha$–Risk Limit for all parameters $a$, i.e.:

$$\forall a \in \mathbb{R}_{++} : WVaR_{a,p}^{P}(x) > VaR_{p}^{P}(x) - \mu_x.$$

The $WVaR$ is indeed a simple and flexible measure of risk, although it just shows the well known exponential utility function in a favorable light. Moreover, this special choice is also compatible with Savage's axiomatics induc-

[70] The risk free condition gives grounds for the following conjecture not proven here: A strictly quasi–concave, continuous utility function $\tilde{U}$ induces the coherent risk measure $-\left(\tilde{U}(x-\mu_x)+\mu_x\right)$; and a utility function $\tilde{v}$ applied only to mean and risk can mimic together with a coherent risk measure $\tilde{\rho}$ only utility functions showing to have constant absolute risk aversion.

ing expected utility, since it is minus a monotone transformation of a von–Neumann–Morgenstern utility function.

### 2.10.3 Coherent Risk Measures in Portfolio Choice

Suppose a portfolio manager invests owners' capital $e$ and foreign capital $f$ in $J$ assets with non–negative pay–offs $A\theta$. The return on investment is $A\theta - \hat{r} \cdot f$, where $\hat{r} - 1$ is the risk–adjusted interest rate on foreign capital. The price of the *riskless* bond is normalized to one; therewith time preferences are ignored here. In case of default, which occurs if $A\theta < \hat{r} \cdot f$, creditors receive all remaining pay–offs $A\theta$. The pricing measure $Q$ applies to the valuation of risky pay–offs. Let the probability of default be $\alpha$ and $\hat{\alpha}$ under $P$ and $Q$, respectively. Thus, in case of default creditors can expect a present value at an amount of $-WCE_{\hat{\alpha}}^{Q}(A\theta)$ from the bankrupt's estate.
Owners impose their utility function $U(x)$ as the optimization criterion, whereas creditors demand a credit spread in an amount of $\hat{r} - 1$, which is determined endogenously by the portfolio choice $\theta$. All participants agree on a unique scenario $P$, which is assumed to be a continuous probability distribution of returns. The conditions for a partial equilibrium between creditors and the portfolio manager are as follows:

1. The portfolio manager maximizes owners' utility by choosing the portfolio and the right amount of foreign capital:

   $$\max_{f \in \mathbb{R}_+, \theta \in \mathbb{R}_+^J} U(A\theta - \hat{r}f) \text{ s.t. } q^T\theta = e + f.$$

   Their utility function might be linear in expected returns and risk – measured by WVaR –, where only the returns above the default threshold, i.e. $A\theta \geq \hat{r}f$, count:

   $$U(x) = \beta E_P(x|\, x \geq 0) - \frac{1}{a} \ln E_P(\exp(-ax)|\, x \geq 0)$$
   $$\text{for some } \beta \geq 0 \text{ and } a > 0.$$

   Furthermore, on demand of the regulator the bank's portfolio manager has to comply with the risk–constraint:

$$e \geq VaR_{\alpha}^{P}(A\theta), \tag{CondReg}$$

where $\alpha$ is chosen to be a conservative estimate of the bank's default probability.[71]

2. Creditors are indifferent between the riskless asset and the risky debt lend to investors' fund with respect to the pricing measure $Q$. This determines the fair interest rate $\hat{r}-1$ and the *risk neutral* probability of default $\hat{\alpha}$ :

$$\begin{aligned}\hat{r} &: \; f = (1-\hat{\alpha})\,\hat{r}f - \hat{\alpha}WCE_{\hat{\alpha}}^{Q}(A\theta)\\ \hat{\alpha} &: \; VaR_{\hat{\alpha}}^{Q}(A\theta) = \mu_{A}^{Q}\theta - \hat{r}f.\end{aligned}$$

The last equation is equivalent to $\hat{\alpha} = Q\left(A\theta \geq \hat{r}\cdot f\right).$

In an ideal world the endogenous Condition 2. would be part of a debt agreement, which is usually contracted prior to the investment decision. Regulators can internalize the latter condition for $\hat{\alpha}$ into the portfolio decision of leveraged banks, firms or funds as an additional constraint. Thus, the portfolio manager would be enforced to comply with a second constraint, which – given a contracted yield $\hat{r}$ – sounds

$$-WCE_{\hat{\alpha}}^{Q}(A\theta) \geq \underbrace{\frac{1-(1-\hat{\alpha})\,\hat{r}}{\hat{\alpha}}}_{\geq 0} f. \tag{CondRisk}$$

But this condition is indeed different from the condition imposed by the regulator. That should not be the case, if the banking regulation intends to protect stake holders' interest. Only Condition (CondRisk) is economically founded but not Condition (CondReg). The following line of reasoning will show how the two conditions are related to each other.

[71] For rated banks the rating agencies publish historical default frequencies, which are proper estimates of the default probabilities.

Taking the budget constraint and the pricing rule $q = \mu_A^Q$ into account while expanding Condition (CondRisk) with a zero it transforms to

$$
\begin{aligned}
e \geq\ & VaR_{\alpha}^{P}(A\theta) \\
& \underbrace{+\mu_A^Q\theta + WCE_{\hat{\alpha}}^{Q}(A\theta) - VaR_{\alpha}^{P}(A\theta)}_{\text{is} \lesseqgtr 0\ ?} \\
& \underbrace{+\frac{(1-\hat{\alpha})(\hat{r}-1)}{1-(1-\hat{\alpha})\hat{r}} WCE_{\hat{\alpha}}^{Q}(A\theta)}_{\leq 0}.
\end{aligned}
$$

The first row is equal to Condition (CondReg), which would be more conservative if the remaining expressions of the right hand side were negative together. The Risk Limit (equal to the $\hat{\alpha}-$quantile) is never smaller than the corresponding lower partial first moment. This implies:

$$\mu_A^Q\theta + WCE_{\hat{\alpha}}^{Q}(A\theta) \geq VaR_{\hat{\alpha}}^{Q}(A\theta).$$

The risk neutral default probabilities $\hat{\alpha}$ are empirically much higher[72] than the historical default frequencies $\alpha$. And the expected return under $P$ includes usually a positive risk premium and is therefore higher than the expected return under $Q$. Both presumptions would imply a much more conservative estimate by the VaR with $P$ than with $Q$:

$$VaR_{\alpha}^{P}(A\theta) \gg VaR_{\hat{\alpha}}^{Q}(A\theta).$$

Thus, the sign of the sum of all effects which occur if one moves from Condition (CondRisk) to Condition (CondReg) remains unclarified. It is hard to say which of the two conditions is more conservative than the other. This seems to have only an empirical answer.

To rule out the intervention of regulators, the quantities $\hat{\alpha}, \hat{r}$ and $f$, the procedure how to build the risk neutral probability measure $Q$, and the rules of monitoring the portfolio decisions by auditors have to be specified in the

[72] Those are derived from interest rate spreads and expected recovery rates.

debt contract. Because of transaction costs it might be justified to trust in Condition (CondReg), which to apply is much more convenient.

### 2.10.4 Conclusion

In this section it has been pointed out that two risk measures are necessary to model a portfolio manager's choice problem more realistic. Since investors' bankruptcy is a possible event, creditors' preferences come into play via an additional risk–constraint. Because of high transaction costs of contracting on a certain risk profile, regulators impose capital constraints on leveraged investments. Currently, the most advanced approach is to demand a level of economic capital in an amount of the Value–at–Risk for a certain confidence level. Therefore, it is still common in portfolio management to take variance as the optimization criterion while holding the Value–at–Risk equal to the amount of ventured capital. However, both measures of risk, variance and VaR, are not strictly monotonic. Recently, two alternative risk measures have been suggested: Worst Conditional Expectation and Weighted Value–at–Risk, which satisfy some convincing axioms of coherentness. It has been shown that WCE suits very well for regulators' interest to restrict the risk–level, especially the risk of losses on debt, while WVaR might replace variance as a consistent optimization criterion.
By formulating the partial equilibrium between leveraged investors and creditors it turned out that regulators' demand for economic capital determined by the VaR can only be a rough estimate of the capital necessary to protect stake holders' interest. In a frictionless world the right amount of economic capital would be determined by the WCE rather than by the VaR.

## 2.11 Appendix with Proofs

### 2.11.1 Agnew's condition

The existence of a Bond is assumed, which allows for a correction of assets' mean. But only a set of linear independent, risky assets are considered explicitly, which is denoted in this subsection with $A$ and not with $A_{\backslash\mathbf{1}}$ for the sake of simplicity. Let $P^A : \mathbb{R}^{J-1} \to [0,1]$ be the distribution of asset pay–offs. Then its characteristic function $\phi^A : \mathbb{R}^{J-1} \to \mathbb{C}$ is defined as

$$\phi^A(\theta) := \int_{\mathbb{R}^J} e^{iA\theta} P^A(dA),$$

implying a characteristic function $\phi^{A\theta} : \mathbb{R} \to \mathbb{C}$ for the pay–offs of certain portfolios $\theta$ :

$$\phi^{A\theta}(t) = \int_{\mathbb{R}^J} e^{itA\theta} P^A(dt) \equiv \phi^A(t\theta).$$

If pay–offs are normalized to a mean of zero and a variance of one, i.e.

$$\left(\theta^T \sigma_A^2 \theta\right)^{-\frac{1}{2}} (A - \mu_A)\,\theta,$$

the characteristic function transforms to

$$\phi^{\left(\theta^T\sigma_A^2\theta\right)^{-\frac{1}{2}}(A-\mu_A)\theta}(t) = e^{-it\left(\theta^T\sigma_A^2\theta\right)^{-\frac{1}{2}}\mu_A\theta}\phi^A\left(t\left(\theta^T\sigma_A^2\theta\right)^{-\frac{1}{2}}\theta\right).$$

Let $\psi(t)$ be the characteristic function of a distribution function $F : \mathbb{R} \to [0,1]$ generating a random variable with zero mean and unit variance. Expected utility induces mean/variance utility if and only if for all $t \in \mathbb{R}$ and $\theta \in \mathbb{R}^{J-1}$ yielding $\theta^T \sigma_A^2 \theta > 0$ it holds that

$$\psi(t) = e^{-it\left(\theta^T\sigma_A\theta\right)^{-\frac{1}{2}}\mu_A\theta}\phi^A\left(t\left(\theta^T\sigma_A^2\theta\right)^{-\frac{1}{2}}\theta\right).$$

Thus portfolio pay–offs with equal mean and variance coincide in distribution. By replacing $t\left(\theta^T\sigma_A^2\theta\right)^{-\frac{1}{2}}$ with $\hat{t}$ and $\hat{t}\theta$ with $\hat{\theta}$, respectively, this equation transforms to Agnew's condition (while ignoring the circumflexes):

$$\forall \theta \in \mathbb{R}^{J-1}, \zeta \in \{-1,1\}: \; \phi^{A}(\theta) = e^{i\mu_{A}\theta}\psi\left(\zeta\left(\theta^{T}\sigma_{A}^{2}\theta\right)^{\frac{1}{2}}\right).$$

Since the left hand side does not depend on $\zeta$ the distribution $F$ must ensure that this holds for $\psi$ either. This condition implies symmetry for $F$, i.e. $F(t) = 1 - F(-t)$. The economic reason for this is that altering the sign of a portfolio, which yields a zero mean, does not effect the distribution of portfolio returns.

### 2.11.2 Efficient Frontier

**Proof. of Lemma** 42: Note $Var(x+y) = \sigma_x^2 + 2\sigma_{x,y} + \sigma_y^2$, and $\sigma_{x,y}$ $= M_{x,y} - \mu_x\mu_y = -\mu\mu_y$ so that just $\sigma_x^2$ is minimized by $x$ with subject to the constraints: $\mu = \mu_x$ and $\bar{b} \geq q(x)$. It follows from Corollary 28 and Lemma 38 that whenever $\mu_{\mathbf{1}_X} > 0$ a variance efficient solution must have the representation:

$$\begin{aligned} x &= \gamma\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} + \alpha\chi \\ &= (\alpha+\gamma)\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} - \alpha\frac{\mu_\pi}{Var_X(\pi)}\left(\pi - \mu_\pi\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right). \end{aligned}$$

The coefficients $\alpha$ and $\gamma$ are determined by the constraints:

$$\begin{aligned} \alpha+\gamma &= \mu, \\ (\mu-\alpha)\frac{\mu_\pi}{\mu_{\mathbf{1}_X}} &\leq \bar{b} \text{ and} \\ \alpha &\geq 0. \end{aligned}$$

The solution depends on different cases:

1. Case: $\mu_{\mathbf{1}_X} > 0, \pi \nparallel \mathbf{1}_X$, and $\mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \geq \bar{b}$ :

$$\begin{aligned} (\mu - \alpha) \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} &= \bar{b} \Rightarrow \\ \alpha &= \mu - \bar{b} \frac{\mu_{\mathbf{1}_X}}{\mu_\pi} \text{ and} \\ \gamma &= \bar{b} \frac{\mu_{\mathbf{1}_X}}{\mu_\pi}. \end{aligned}$$

The variance is:

$$\begin{aligned} Var(x+y) &= (\alpha+\gamma)^2 Var\left(\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right) - 2\mu\mu_y + \sigma_y^2 \\ &\quad + \left(\frac{\alpha\mu_\pi}{Var_X(\pi)}\right)^2 Var\left(\pi - \mu_\pi \frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}\right) \\ &= \mu^2 \frac{1 - \mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} + \left(\mu - \bar{b} \frac{\mu_{\mathbf{1}_X}}{\mu_\pi}\right)^2 \frac{\mu_\pi^2}{Var_X(\pi)\, \mu_{\mathbf{1}_X}} - 2\mu\mu_y + \sigma_y^2 \\ &= \frac{\bar{b}^2 \mu_{\mathbf{1}_X} - 2\mu\mu_\pi \bar{b} + \mu^2 q(\pi)}{q(\pi)\mu_{\mathbf{1}_X} - \mu_\pi^2} - \mu^2 - 2\mu\mu_y + \sigma_y^2. \end{aligned}$$

This is the most relevant constellation mainly considered in the text.

2. Case: $\mu_{\mathbf{1}_X} > 0, \pi \parallel \mathbf{1}_X$ and $\mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} > \bar{b}$ :
Thus, the asset $\pi - \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \mathbf{1}_X$ is equal to zero. But then the constraints cannot be fulfilled, since all assets orthogonal to $\mathbf{1}_X$ have a zero price. The best investment, which is allowed to be done, is

$$x = \frac{\bar{b}}{\mu_\pi} \mathbf{1}_X \text{ with } \mu_x = \frac{\bar{b}}{\mu_\pi} \mu_{\mathbf{1}_X} < \mu \text{ and } q(x) = \bar{b},$$

which minimizes the difference between the desired and the feasible mean.

3. Case: $\mu_{\mathbf{1}_X} > 0$ and $\mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \leq \bar{b}$ :
Observe that the almost best solution is possible if $\frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \leq \frac{\bar{b}}{\mu}$ (for in-

stance if $\mu_\pi \leq 0$). Thus

$$\begin{aligned} x &= \frac{\mu}{\mu_{\mathbf{1}_X}} \mathbf{1}_X \quad \text{with} \\ Var(x+y) &= \mu^2 \frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu\mu_y + \sigma_y^2. \end{aligned}$$

4. Case: $\mu_{\mathbf{1}_X} = 0 \Rightarrow \mu = 0$ and $\bar{b} < 0$ :
   Then only the budget can be balanced by:

$$\begin{aligned} x &= \frac{\bar{b}}{q(\pi)} \pi \quad \text{with} \\ Var(x+y) &= \frac{\bar{b}^2}{q(\pi)} + \sigma_y^2. \end{aligned}$$

5. Case: $\mu_{\mathbf{1}_X} = 0 \Rightarrow \mu = 0$ and $\bar{b} \geq 0$ :

$$\begin{aligned} \hat{x} &= 0, \\ Var(x+y) &= \sigma_y^2. \end{aligned}$$

There is no solution for $\mu_{\mathbf{1}_X} = 0$ and $\mu \neq 0$.

In the first, most relevant case when $\bar{b}\mu_{\mathbf{1}_X} < \mu\mu_\pi$ the derivatives of the variance for $\bar{b}$ and $\bar{b}, \mu$ sound:

$$\begin{aligned} \sigma_{\bar{b}}^2 &= 2\frac{\bar{b}\mu_{\mathbf{1}_X} - \mu\mu_\pi}{M_\pi \mu_{\mathbf{1}_X} - \mu_\pi^2} < 0 \text{ and} \\ \sigma_{\bar{b}\mu}^2 &= -2\frac{\mu_\pi}{M_\pi \mu_{\mathbf{1}_X} - \mu_\pi^2} < 0. \end{aligned}$$

In the lemma the cases one and three are combined yielding the variance

$$Var(\hat{x}+y) = \Upsilon\left(\mu, \bar{b}\right) \frac{\left(\bar{b} - \mu\frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right)^2 \mu_{\mathbf{1}_X}}{Var_X(\pi)} + \mu^2 \frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu\mu_y + \sigma_y^2,$$

where the indicator function $\Upsilon\left(\mu, \bar{b}\right)$ is equal to one if and only if

$$\pi \nparallel \mathbf{1}_X \quad \text{and} \quad \mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \geq \bar{b}.$$

For $\pi \nparallel \mathbf{1}_X$ the indicator function becomes zero exactly at the point where the parabola

$$\left(\bar{b} - \mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right)^2$$

reaches its minimum at zero. At this point the derivative of this expression with respect to $\bar{b}$ or $\mu$ is zero either. Thus, standard deviation is continuously differentiable with a kink in its curvature only at $\bar{b} = \mu \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}$.
The standard deviation defines a norm in $\mathbb{R}^J$ (in $\mathbb{R}^{J-1}$ if $\mathbf{1}_X = \mathbf{1}$) of the portfolios, because $\sigma_A^2$ $\left(\sigma_{A\setminus\mathbf{1}}^2\right)$ is strictly positive definite and symmetric. Thus, $\sigma_{x+y} = \left(\theta_x^T \sigma_A^2 \theta_x - 2\mu\mu_y + \sigma_y^2\right)^{\frac{1}{2}}$ is a (if $\mathbf{1}_X = \mathbf{1}$ and $\sigma_y^2 = 0$ then weakly) convex function. Along the Capital Market Line $\sigma_{x+y}$ is minimized with respect to two constraints, which are linear in the two parameters $\mu$ and $\bar{b}$. Then the minimal standard deviation is itself a (if $\mathbf{1}_X = \mathbf{1}$ and $\sigma_y^2 = 0$ then weakly) convex function of those parameters. ■

### 2.11.3 Mutual Fund

**Proof.** of Corollary 47: A trader's optimization problem reads:

$$\max_{\mu^i \in \mathbb{R}} v^i \left(\mu^i + \mu_{w^i_{X^\perp}}, \sigma^2\left(\mu^i, q\left(w_X^i\right), \mu_{w^i_{X^\perp}}\right)\right).$$

Variance, $\sigma^2\left(\mu^i, q\left(w_X^i\right), \mu_{w^i_{X^\perp}}\right)$, is a differentiable parabola in mean:

$$\Upsilon^i \frac{\left(q\left(w_X^i\right) - \mu^i \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right)^2 \mu_{\mathbf{1}_X}}{Var_X\left(\pi\right)} + \left(\mu^i\right)^2 \frac{1 - \mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu^i \mu_{w^i_{X^\perp}} + \sigma^2_{w^i_{X^\perp}},$$

with a derivative for $\mu$ equal to:

$$-2\Upsilon^i \frac{\left(q\left(w_X^i\right) - \mu^i \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\right)\mu_\pi}{Var_X\left(\pi\right)} + 2\mu^i \frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} - 2\mu_{w^i_{X^\perp}}.$$

Therewith the first order condition:

$$-r^i = \frac{1}{2}\sigma_\mu^2,$$

transforms to:

$$\mu^i = \left(\frac{1-\mu_{\mathbf{1}_X}}{\mu_{\mathbf{1}_X}} + \Upsilon^i \frac{\mu_\pi^2}{\mu_{\mathbf{1}_X} Var_X\left(\pi\right)}\right)^{-1}\left(\mu_{w^i_{X^\perp}} + \Upsilon^i \frac{\mu_\pi q\left(w_X^i\right)}{Var_X\left(\pi\right)} - r^i\right).$$

If the consumer is not satiated, i.e. if $\Upsilon^i = 1$, the optimal mean sounds:

$$\mu^i = \frac{q\left(\pi\right) - Var_X\left(\pi\right)}{Var_X\left(\pi\right)}\left(\mu_{w^i_{X^\perp}} + \frac{\mu_\pi q\left(w_X^i\right)}{Var_X\left(\pi\right)} - r^i\right).$$

The resulting mean $\mu^i$ is applied to the Capital Market Line, see Lemma 42. This yields the optimal demand in $X$. The cases not mentioned in the main text are:

1. If $\mathbf{1}_X = 0 \neq \pi$ :

$$\hat{x}_X^i = \frac{q\left(w_X^i\right)}{q(\pi)}\pi \text{ and } \mu^i = 0.$$

2. If $\mathbf{1}_X = 0 = \pi$ :

$$\hat{x}_X^i = 0 \text{ and } \mu^i = 0.$$

3. If $\mathbf{1}_X \neq 0 = \pi$ consumers demand is either infinite or they are satiated in $\mathbf{1}_X$.

If the pricing asset $\pi$ is replaced by its positive multiple $\lambda\pi$, the number $\lambda > 0$ cancels down in every equation where $\pi$ is involved. ■

### 2.11.4 CAPM–Equilibrium

**Proof.** of Proposition 50: Let $\bar{R} < 0$ and consumers not being satiated. The conditions for non–satiation are shown afterwards with the equilibrium properties, which have to be established first.

Since $\bar{w}_X \parallel \mathbf{1}_X$ leads to a degenerated equilibrium this case will not explicitly be considered to save notation. However $\bar{w}_X \parallel \mathbf{1}_X$ is covered by the following calculations if one allows for zeros in the denominator to cancel down for zeros in the numerator, which is indeed harmless, because all expressions are continuous and converge for $\bar{w}_X \to \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}} \mathbf{1}_X$.

Thus, assume $\bar{w}_X \nparallel \mathbf{1}_X$. The equilibrium condition to solve for is $\sum_{i \in \mathbf{I}} x_X^i = \bar{w}_X$ implying $\sum_{i \in \mathbf{I}} \mu^i = \mu_{\bar{w}_X}$. With the aggregated quantities $\bar{w}_{X^\perp}$ and $\bar{r}$ those equations transform in the following way. The market clearing condition for mean:

$$\mu_{\bar{w}_X} = \frac{1}{\frac{q(\pi)}{Var_X(\pi)} - 1} \left( E\left(\bar{w}_{X^\perp}\right) + \frac{\mu_\pi q\left(\bar{w}_X\right)}{Var_X\left(\pi\right)} - \bar{r} \right),$$

transforms equivalently to

$$\left(\bar{r} - \mu_{\bar{w}}\right) Var_X\left(\pi\right) = \mu_\pi q\left(\bar{w}_X\right) - \mu_{\bar{w}_X} q\left(\pi\right). \tag{1}$$

Whereas the market clearing condition for consumption plans in $X$ :

$$\bar{w}_X = \frac{\mu_{\mathbf{1}_X} q\left(\bar{w}_X\right) - \mu_{\bar{w}_X}\mu_\pi}{Var_X\left(\pi\right)} \left(\pi - \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \mathbf{1}_X\right) + \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}} \mathbf{1}_X,$$

yields the pricing asset

$$\pi = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \mathbf{1}_X + \frac{Var_X\left(\pi\right)}{Cov_X(\bar{w}_X, \pi)} \left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}} \mathbf{1}_X\right). \tag{2}$$

Remind that for the pricing of any $x \in X$

$$\begin{aligned} E\left[\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right)x\right] &= E\left(\bar{w}_X x\right) - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mu_x \\ &= \frac{1}{\mu_{\mathbf{1}_X}} Cov_X(\bar{w}_X, x) \end{aligned}$$

holds. Now it is appropriate to simplify $q\left(\bar{w}_X\right)$, $q\left(\pi\right)$ and $q\left(w_X^i\right)$, and to substitute the resulting expressions in $\pi$ and $\mu^i$. The Equation (1) solves for $q\left(\pi\right)$ :

$$q\left(\pi\right) = \frac{\mu_\pi q\left(\bar{w}_X\right) + \left(\bar{r} - \mu_{\bar{w}}\right)\mu_\pi^2}{\bar{R}}. \tag{3}$$

Pricing $\bar{w}_X$ with Equation (2) reads:

$$\begin{aligned} q\left(\bar{w}_X\right) &= E\left(\bar{w}_X \pi\right) \\ &= \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\mu_{\bar{w}_X} + \frac{Var_X\left(\pi\right)}{Cov_X(\bar{w}_X, \pi)}\frac{1}{\mu_{\mathbf{1}_X}} Var_X\left(\bar{w}_X\right), \end{aligned} \tag{4}$$

which is equivalent to

$$Var_X\left(\pi\right) \cdot Var_X\left(\bar{w}_X\right) = \left[Cov_X(\bar{w}_X, \pi)\right]^2.$$

Inserting the price of $\pi$ from Equation (3) into $Var_X\left(\pi\right)$ yields

$$\begin{aligned} Var_X\left(\pi\right) &= q(\pi)\mu_{\mathbf{1}_X} - \mu_\pi^2 \\ &= \frac{\mu_{\mathbf{1}_X} q\left(\bar{w}_X\right) - \mu_{\bar{w}_X}\mu_\pi}{\bar{R}}\mu_\pi \\ &= \frac{Cov_X(\bar{w}_X, \pi)}{\bar{R}}\mu_\pi. \end{aligned}$$

Combining the two previous equations

$$\begin{aligned} \bar{R} &= \frac{\mu_\pi Cov_X(\bar{w}_X, \pi)}{Var_X\left(\bar{w}_X\right)} \quad \text{and} \\ \bar{R}^2 &= \frac{\mu_\pi^2 Var_X\left(\bar{w}_X\right)}{Var_X\left(\pi\right)} \end{aligned}$$

follow. These two equations determine any two of the three unknown terms $q(\bar{w}_X), q(\pi)$ and $\bar{r}$. With this the pricing asset from Equation (2) sounds

$$\pi = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \mathbf{1}_X + \frac{\mu_\pi}{\bar{R}} \left( \bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}} \mathbf{1}_X \right),$$

and the price of the market asset given by Equation (4) is

$$q(\bar{w}_X) = \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \left( \mu_{\bar{w}_X} + \frac{Var_X(\bar{w}_X)}{\bar{R}} \right),$$

which itself simplifies the price of the pricing asset from Equation (3):

$$q(\pi) = \frac{\mu_\pi^2}{\mu_{\mathbf{1}_X}} \left[ 1 + \frac{Var_X(\bar{w}_X)}{\bar{R}^2} \right].$$

Now the pricing of any $x \in X$ by $\pi$ reveals the famous $\beta-$pricing formula:

$$\begin{aligned} q(x) &= \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \left[ \mu_x + \frac{Cov_X(\bar{w}_X, x)}{\bar{R}} \right] \\ &= \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \mu_x + \frac{Cov_X(\bar{w}_X, x)}{Var_X(\bar{w}_X)} \left( q(\bar{w}_X) - \mu_{\bar{w}_X} \frac{\mu_\pi}{\mu_{\mathbf{1}_X}} \right). \end{aligned}$$

Since individual consumption plans are determined by the Mutual Fund Theorem 47, it remains to simplify the demand for mean derived in Appendix 2.11.3:

$$\begin{aligned} \mu^i &= -\frac{1}{\frac{q(\pi)}{Var_X(\pi)} - 1} \left( r^i - \frac{\mu_\pi q(w_X^i)}{Var_X(\pi)} - \mu_{w^i_{X^\perp}} \right) \\ &= \frac{\mu_\pi q(w_X^i)}{q(\pi) - Var_X(\pi)} - \frac{Var_X(\pi)}{q(\pi) - Var_X(\pi)} \left( r^i - \mu_{w^i_{X^\perp}} \right). \end{aligned}$$

With

$$q(\pi) - Var_X(\pi) = \frac{\mu_\pi^2}{\mu_{\mathbf{1}_X}} \frac{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)}{\bar{R}^2},$$

and

$$\frac{Var_X(\pi)}{q(\pi) - Var_X(\pi)} = \frac{Var_X(\bar{w}_X)\,\mu_{\mathbf{1}_X}}{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)},$$

as well as

$$\frac{q\left(w_X^i\right)}{q(\pi) - Var_X(\pi)} = \frac{\bar{R}^2\mu_{w_X^i} + \bar{R}Cov_X\left(\bar{w}_X, w_X^i\right)}{\mu_\pi\left(\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)\right)}$$

the demand for mean in $X$ is

$$\mu^i = \frac{\bar{R}^2\mu_{w_X^i} + \bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) + Var_X(\bar{w}_X)\,\mu_{\mathbf{1}_X}\left(\mu_{w_{X^\perp}^i} - r^i\right)}{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)}.$$

Expanding the right hand side by $\mu_{w_X^i}$ results in

$$\mu^i = \mu_{w_X^i} + \frac{\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) - R^i Var_X(\bar{w}_X)}{\bar{R}^2 + Var_X(\bar{w}_X)\left(1 - \mu_{\mathbf{1}_X}\right)}.$$

If $\bar{w}_X \parallel \mathbf{1}_X$ the fraction is zero by zero $X-$(co)variance. From the Mutual Fund Theorem 47 trader $i$'s consumption plan follows, in which the results derived so far are inserted:

$$\begin{aligned}
\hat{x}_X^{*i} &= \frac{q\left(w_X^i\right)\mu_{\mathbf{1}_X} - \mu^i\mu_\pi}{Var_X(\pi)}\left(\pi - \frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) + \frac{\mu^i}{\mu_{\mathbf{1}_X}}\mathbf{1}_X \\
&= \bar{R}\frac{\frac{\mu_\pi}{\mu_{\mathbf{1}_X}}\left[\mu_{w_X^i} + \frac{Cov_X\left(\bar{w}_X, w_X^i\right)}{\bar{R}}\right]\mu_{\mathbf{1}_X} - \mu^i\mu_\pi}{\mu_\pi Var_X(\bar{w}_X)}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) + \frac{\mu^i}{\mu_{\mathbf{1}_X}}\mathbf{1}_X \\
&= \frac{\bar{R}\left(\mu_{w_X^i} - \mu^i\right) + Cov_X\left(\bar{w}_X, w_X^i\right)}{Var_X(\bar{w}_X)}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) + \frac{\mu^i}{\mu_{\mathbf{1}_X}}\mathbf{1}_X.
\end{aligned}$$

The first fraction in the previous line simplifies by the solution for $\mu^i$ as follows:

$$\frac{Cov_X\left(\bar{w}_X, w_X^i\right)}{Var_X\left(\bar{w}_X\right)} - \frac{\bar{R}^2 Cov_X\left(\bar{w}_X, w_X^i\right) - \bar{R}R^i Var_X\left(\bar{w}_X\right)}{Var_X\left(\bar{w}_X\right)\left[\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)\right]} =$$

$$\frac{\bar{R}R^i + Cov_X\left(\bar{w}_X, w_X^i\right)\left(1-\mu_{\mathbf{1}_X}\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}.$$

Now re–arrange the equation for $\hat{x}_X^{*i}$ while taking the $L^2$–norm:

$$\begin{aligned}
\sqrt{\frac{Var_X\left(\hat{x}_X^{*i}\right)}{\mu_{\mathbf{1}_X}}} &= \frac{\bar{R}\left(\mu_{w_X^i} - \mu^i\right) + Cov_X\left(\bar{w}_X, w_X^i\right)}{\sqrt{Var_X\left(\bar{w}_X\right)\mu_{\mathbf{1}_X}}} \Leftrightarrow \\
\frac{\bar{R}\left(\mu^i - \mu_{w_X^i}\right)}{\sqrt{Var_X\left(\bar{w}_X\right)}} &= \frac{Cov_X\left(\bar{w}_X, w_X^i\right)}{\sqrt{Var_X\left(\bar{w}_X\right)}} - \sqrt{Var_X\left(\hat{x}_X^{*i}\right)}.
\end{aligned}$$

Obviously, an equilibrium is incompatible with over satiated consumers. The condition for traders to be not over satiated in the Quasi–Bond sounds by Lemma 35:

$$r^i\mu_{\mathbf{1}_X} + \mu^i\left(1-\mu_{\mathbf{1}_X}\right) - \mu_{\mathbf{1}_X}\mu_{w_X^i} \leq 0.$$

With the solution for $\mu^i$ and the definition of $R^i$ this inequality transforms to

$$R^i + \left(1-\mu_{\mathbf{1}_X}\right)\frac{\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) - R^i Var_X\left(\bar{w}_X\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)} \leq 0,$$

which is equivalent to

$$R^i\bar{R}^2 + \left(1-\mu_{\mathbf{1}_X}\right)\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) \leq 0.$$

If this equation is aggregated over $i \in \mathbf{I}$ it reads

$$\bar{R}\left[\bar{R}^2 + \left(1-\mu_{\mathbf{1}_X}\right)Var_X\left(\bar{w}_X\right)\right] \leq 0.$$

Thus, it holds that $\bar{R} \leq (<)\ 0$ if traders are not over satiated in $\mathbf{1}_X$ (and at least one trader is not satiated in $\mathbf{1}_X$). Therewith the individual constraint shows to be

$$-R^i \bar{R} \leq \left(1 - \mu_{\mathbf{1}_X}\right) Cov_X\left(\bar{w}_X, w_X^i\right),$$

where equality holds if and only if consumer $i$ is satiated.

Positive homogeneity of degree zero in $\pi$ obviously holds, because this property holds already for traders' optimal demand, see Corollary 47. ■

### 2.11.5 Positive State Prices and Monotonicity

**Proof. of Theorem 68:** Remind the pricing asset from Proposition 50:

$$\begin{aligned} \pi &= \mathbf{1}_X + \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) \text{ with} \\ \bar{R} &= \mu_{\mathbf{1}_X}\left(\bar{r} - \mu_{\bar{w}}\right) + \mu_{\bar{w}_X}. \end{aligned}$$

The suitably chosen asset in $\{\cdot\}$ –brackets from the orthogonal subspace $X^{\perp}$ is added to the pricing asset as follows:

$$\begin{aligned} &\pi + \left\{\mathbf{1}_{X^\perp} + \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\bar{w}_{X^\perp} - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_{X^\perp} + \eta\right)\right\} \\ &= \mathbf{1} + \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\bar{w} - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1} + \eta\right) \\ &= \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\frac{\mu_{\mathbf{1}_X}\left(\bar{r} - \mu_{\bar{w}}\right) + \mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1} + \bar{w} - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1} + \eta\right) \\ &= \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left[\left(\bar{r} - \mu_{\bar{w}}\right)\mathbf{1} + \bar{w} + \eta\right]. \end{aligned}$$

Applying the FTP yields that this pricing density should be positive $P$–$a.s.$ for some $\eta \in X^{\perp}$. This rules out arbitrage, which is the first result, since $\mu_{\mathbf{1}_X}/\bar{R}$ is negative if at least one agent is not satiated in any equilibrium (if $\mu_{\mathbf{1}_X} = 0$ agents are always weakly satiated).

For the second assertion one needs to have $\pi > 0$ :

$$\begin{aligned} 0 &< \mathbf{1}_X + \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) \qquad P-a.s. \\ &= \frac{\mu_{\mathbf{1}_X}}{\bar{R}}\left[\left(\bar{r}-\mu_{\bar{w}}\right)\mathbf{1}_X + \bar{w}_X\right] \qquad \Leftrightarrow \\ 0 &> \left(\bar{r}-\mu_{\bar{w}}\right)\mathbf{1}_X + \bar{w}_X \qquad P-a.s. \end{aligned}$$

Monotonicity of preferences in an equilibrium holds, if an increase in the pay–offs in an essential event $\mathbb{E} \in \mathcal{F}$ by a small amount of $\mathbf{1}_{\mathbb{E}}(\omega)$ also enhances utility at the equilibrium allocation. Thus, for all essential events $\mathbb{E} \in \mathcal{F}$ the derivative:

$$\left.\frac{\partial}{\partial\alpha}\right|_{\alpha=0} v^i\left(\mu_{x^{*i}} + \alpha P(\mathbb{E}), \sigma^2_{x^{*i}} + 2\alpha Cov\left(x^{*i}, \mathbf{1}_{\mathbb{E}}\right) + \alpha^2 P(\mathbb{E})\left(1-P(\mathbb{E})\right)\right)$$

should be positive. This inequality transforms to

$$\begin{gathered} \left(v_1^i \circ x^{*i}\right) P(\mathbb{E}) + 2P(\mathbb{E})\left(E\left(x^{*i}\middle|\,\mathbb{E}\right) - \mu_{x^{*i}}\right)\left(v_2^i \circ x^{*i}\right) > 0 \Leftrightarrow \\ E\left(x^{*i}\middle|\,\mathbb{E}\right) - \mu_{x^{*i}} < -\left(r^i \circ x^{*i}\right). \end{gathered}$$

Since this inequality shall hold for all essential events it is equivalent to say

$$x^{*i} - \mu_{x^{*i}}\mathbf{1} + r^{*i}\mathbf{1} < 0 \qquad P-a.s. \tag{*}$$

Inserting the equilibrium allocation from Proposition 50 the right hand side is equal to:

$$\begin{aligned} & w^i_{X^\perp} + \frac{\bar{R}R^i + Cov_X\left(\bar{w}_X, w^i_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}\left(\bar{w}_X - \frac{\mu_{\bar{w}_X}}{\mu_{\mathbf{1}_X}}\mathbf{1}_X\right) \\ & + \left[\mu_{w^i_X} + \frac{\bar{R}Cov_X\left(\bar{w}_X, w^i_X\right) - R^i Var_X\left(\bar{w}_X\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}\right]\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} \\ & + \left[r^i - \mu_{w^i_{X^\perp}} - \mu_{w^i_X} - \frac{\bar{R}Cov_X\left(\bar{w}_X, w^i_X\right) - R^i Var_X\left(\bar{w}_X\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}\right]\mathbf{1}. \end{aligned}$$

If the last row is split up into two consumption streams $[\ldots]\left(\mathbf{1}_X + \mathbf{1}_{X^\perp}\right)$ and the part belonging to the Quasi–Bond is added in form of $[\ldots]\,\mu_{\mathbf{1}_X}\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}$ to the second row the latter transforms to

$$\left[R^i + \left(1-\mu_{\mathbf{1}_X}\right)\frac{\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) - R^i Var_X\left(\bar{w}_X\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}\right]\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}} = \frac{R^i\bar{R} + \left(1-\mu_{\mathbf{1}_X}\right)Cov_X\left(\bar{w}_X, w_X^i\right)}{\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)}\bar{R}\frac{\mathbf{1}_X}{\mu_{\mathbf{1}_X}}.$$

Then this expression is added to the second term of the first row by taking $\bar{R} = (\bar{r}-\mu_{\bar{w}})\,\mu_{\mathbf{1}_X} + \mu_{\bar{w}_X}$ into account. Now by multiplying the whole inequality with $\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)$ it has equivalently been transferred to

$$\begin{aligned}&\left[\bar{R}R^i + Cov_X\left(\bar{w}_X, w_X^i\right)\left(1-\mu_{\mathbf{1}_X}\right)\right]\left(\bar{w}_X + (\bar{r}-\mu_{\bar{w}})\,\mathbf{1}_X\right)\\ &\quad + \left[\bar{R}^2 + Var_X\left(\bar{w}_X\right)\left(1-\mu_{\mathbf{1}_X}\right)\right]\left(w_{X^\perp}^i + \left(r^i-\mu_{w^i}\right)\mathbf{1}_{X^\perp}\right)\\ &\qquad - \left[\bar{R}Cov_X\left(\bar{w}_X, w_X^i\right) - R^i Var_X\left(\bar{w}_X\right)\right]\mathbf{1}_{X^\perp} < 0 \qquad P-a.s.\end{aligned}$$

Aggregating Equation (*) over $i \in \mathbf{I}$ yields in equilibrium:

$$0 < (\mu_{\bar{w}} - \bar{r})\,\mathbf{1} - \bar{w} \qquad P-a.s.$$

This is a strictly positive random variable if monotonicity holds individually. Hence, it must have a positive second cross–moment with the Quasi–Bond, if the latter is not zero:

$$E\left[\left((\mu_{\bar{w}} - \bar{r})\,\mathbf{1} - \bar{w}\right)\mathbf{1}_X\right] = (\mu_{\bar{w}} - \bar{r})\,\mu_{\mathbf{1}_X} - \mu_{\bar{w}_X} = -\bar{R} > 0.$$

Thus, it is a necessary condition for monotonicity that $\bar{R}$ is negative in an equilibrium. ■

# References


Aghion, Philippe, Hart, Oliver, & Moor, John. 1992. The Economics of Bankruptcy Reform. *Journal of Law, Economics, and Organization*, **8**(3), 523 – 546.

Agnew, R. A. 1971. Counter-Examples to an Assertion Concerning the Normal Distributions and a New Stochastic Price Fluctuation Model. *Review of Economic Statistics*, **38**, 381 – 383.

Allingham, Michael. 1991. Existence Theorems in the Capital Asset Pricing Model. *Econometrica*, **59**(4), 1169 – 1174.

Anscombe, F., & Aumann, R. 1963. A Definition of Subjective Probability. *Annals of Mathematical Statistics*, **34**, 199 – 205.

Arrow, K. J. 1974. *Essays in the Theory of Risk–Bearing*. Amsterdam: North–Holland.

Artzner, Philippe, Delbaen, Freddy, Eber, Jean-Marc, & Heath, David. 1998 (July). *Coherent Measures of Risk*. mimeo ETH Zürich.

Aspandiiarov, Sanjar, Bottazzi, Jean-Marc, & Rabinowitz, Neil. 1998 (December). *Weighted Value at Risk: Portfolio Theory and Risk Management – The Best of Both Worlds*. mimeo CERMSEM Université de Paris I.

Barndoff-Nielsen, O. 1978. Hyperbolic Distributions and Distributions on the Hyperbolae. *Scandinavian Journal of Statistics*, **5**, 151 – 157.

Basle Committee of Banking Supervision. 2001 (January). *The New Basle Capital Accord – Consultative Document*. www.bis.org.

Bawa, V. S. 1978. Safety–First, Stochastic Dominance, and Optimal Portfolio Choice. *Journal of Financial and Quantitative Analysis*, **13**, 255 – 271.

Bettzüge, Marc Oliver, Hens, Thorsten, Laitenberger, Marta, & Siwik, Thomas. 2000. On Choquet Prices in a GEI–Model with Intermediation Costs. *Research in Economics*, **54**, 133 – 152.

Billingsley, Patrick. 1995. *Probability and Measure*. Third edn. Probability and Mathmatical Statistics. New York: John Wiley and Sons.

Black, F. 1972. Capital Market Equilibrium with Restricted Borrowing. *Journal of Business*, **45**, 444 – 455.

Blume, L., & Easley, D. 1992. Evolution and Market Behavior. *Journal of Economic Theory*, **58**, 9 – 40.

Borch, K. 1969. A Note on Uncertainty and Indifference Curves. *Review of Economic Studies*, **36**, 1 – 4.

Bottazzi, Jean–Marc, Hens, Thorsten, & Löffler, Andreas. 1998. Market Demand Functions in the Capital Asset Pricing Model. *Journal or Economic Theory*, **79**, 192 – 206.

Cass, D., & Stiglitz, J. 1970. The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Selection: A Contribution to the Pure Theory of Mutual Funds. *Journal of Economic Theory*, **2**, 122 – 160.

Chamberlain, Gary. 1983. A Characterization of the Distributions That Imply Mean–Variance Utility Functions. *Journal of Economic Theory*, **29**, 185 – 201.

Chateauneuf, Alain. 1994. Modeling Attitudes Towards Uncertainty and Risk Through the Use of Choquet Integral. *Annals of Operations Research*, **52**, 3 – 20.

Clark, Stephen A. 1993. The Valuation Problem in Arbitrage Price Theory. *Journal of Mathematical Economics*, **22**, 463 – 478.

Copeland, Thomas E., & Weston, J. Fred. 1979. *Financial Theory and Coporate Policy*. Reading Mas.: Addison–Wesley.

Dalal, A. J., & Arshanapalli, B. G. 1993. Estimating the Demand for Risky Assets Via the Indirect Expected Utility Function. *Journal of Risk and Uncertainty*, **6**, 277 – 288.

Dana, Rose-Anne. 1999. Existence, Uniqueness and Determinacy of Equilibrium in C.A.P.M. with a Riskless Asset. *Journal of Mathematical Economics*, **32**, 167 – 175.

Dubey, Pradeep, Geanakoplos, John, & Shubik, Martin. 2000. *Default and Efficiency in a General Equilibrium Model with Incomplete Markets*. Discussion Paper.

Duffie, Darrell. 1988. *Security Markets: Stochastic Models*. Boston: Academic Press.

Elsinger, Helmut, & Summer, Martin. 1998. *Arbitrage and Optimal Portfolio Choice with Financial Constraints*.

Embrechts, Paul, Mikosch, Thomas, & Klüppelberg, Claudia. 1997. *Modelling Extremal Events: For Insurances and Finance*. Applications of Mathmatics, vol. 33. Berlin: Springer Verlag.

Epstein, Larry G. 1985. Decreasing Risk Aversion and Mean–Variance Analysis. *Econometrica*, **53**, 945 – 961.

Feldstein, M. S. 1969. Mean–Variance Analysis in the Theory of Liquidity Preference and Portfolio Selection. *Review of Economic Studies*, **36**, 5 – 12.

Friend, I., & Blume, M. 1975. The Demand for Risky Assets. *The American Economic Review*, December, 900 – 922.

Geanakoplos, John, & Shubik, Martin. 1990. The General Asset Pricing Model as a General Equilibrium with Incomplete Markets. *The Geneva Papers on Risk and Insurance Theory*, **15**, 55 – 72.

Geanakoplos, John, Zame, Bill, & Dubey, Pradeep. 1996. *Default, Collateral, and Derivatives.* mimeo Venice Conference on General Equilibrium Theory.

Gonedes, Nicholas J. 1976. Capital Market Equilibrium for a Class of Heterogeneous Expectations in a Two-Parameter World. *Journal of Finance*, **31**(1), 1–15.

Güth, Sandra, Ludwig, Sven, & Siwik, Thomas. 2000 (February). *Success of the Uninformed – Pure Chance or More?* mimeo Universität Bielefeld.

Hara, Chiaki. 1997 (December). *A Sequence of Pareto Improving Financial Innovations.* mimeo.

Hara, Chiaki. 1998 (December). *Prices of the Market Portfolio in the CAPM with Incomplete Markets.*

Hart, Oliver D. 1974. On the Existence of Equilibrium in a Securities Model. *Journal of Economic Theory*, **9**, 293 – 311.

Hens, Thorsten. 1998. Elements of General Equilibrium Analysis. *Chap. 5, pages 139 – 210 of:* Kirman, Alan (ed), *Festschrift in Honour of Debreu's 75th Birthday.* Oxford: Blackwell Publishers.

Hens, Thorsten, Laitenberger, Jörg, & Löffler, Andreas. 2000 (July). *On Uniqueness of Equilibria in the CAPM.* mimeo IEW Universität Zürich.

Hiroshi, Konno, & Hiroshi, Shirakawa. 1995. Existence of a Nonnegative Equilibrium Price Vector in the Mean–Variance Capital Market. *Journal of Mathematical Fiance*, **5**(3), 233 – 246.

Horn, Roger A., & Johnson, Charles R. 1985. *Matrix Analysis.* 1996 edn. Cambridge University Press.

Jurczenko, Emmanuel, & Maillet, Bertrand. 2000 (April). *The Three–Moment CAPM: Theoretical Foundations and an Asset Pricing Models Comparison.* mimeo Université de Paris I.

Kataoka, S. 1963. A Stochastic Programming Model. *Econometrica*, **31**, 181 – 196.

Laitenberger, Joerg. 1997. *Existence and Multiplicity of Equilibria in the Two–Factor–CAPM.* mimeo Universität Bonn.

Lajeri, Fatma, & Nielsen, Lars Tyge. 1994 (April). *Risk Aversion and Prudence: The Case of Mean–Variance Preferences.* mimeo INSEAD Fontainebleau.

Levy, H. 1994. Absolute and Relative Risk Aversion: An Experimental Study. *Journal of Risk and Uncertainty*, **8**, 289 – 307.

Lintner, John. 1965a. Security Prices, Risk, and Maximal Gains from Diversification. *Journal of Finance*, **20**(4), 587 – 615.

Lintner, John. 1965b. The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economic Studies*, **47**, 13 – 37.

Lintner, John. 1969. The Aggregation of Investor's Diverse Judgements and Preferences in Purely Competitive Security Markets. *Journal of Financial and Quantitative Analysis*, **4**, 347 – 400.

Lintner, John. 1970. The Market Price of Risk, Size of Market and Investor's Risk Aversion. *Review of Economics and Statistics*, **52**, 87 – 99.

Löffler, Andreas. 1995. *A Risk Aversion Paradox and Wealth Dependent Utility.* mimeo Freie Universität Berlin.

Löffler, Andreas. 1996a. *Theorie Des Consumption Based CAPM.* Ph.D. thesis, Freie Universität Berlin.

Löffler, Andreas. 1996b. Variance Aversion Implies Mean–Variance–Criterion. *Journal of Economic Theory*, **69**, 532 – 539.

Machina, M. J. 1982a. Expected Utility Analysis Without the Independence Axiom. *Econometrica*, **50**, 277 – 323.

Machina, M. J. 1982b. A Stronger Characterization of Declining Risk Aversion. *Econometrica*, **50**, 1069 – 1079.

Magill, Michael J. P., & Quinzii, Martine. 1996. *The Theory of Incomplete Markets*. Vol. 1. Cambridge, MA: MIT–Press.

Markowitz, H. M. 1952. Portfolio Selection. *Journal of Finance*, **7**, 77 – 91.

Markowitz, H. M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: Wiley.

Mas–Colell, Andreu. 1992. *Equilibrium Theory with Possibly Satiated Preferences*. Macmillan, London. Chap. 9, pages 201 – 213.

Mas–Colell, Andreu, Whinston, Michael D., & Green, Jerry R. 1995. *Microeconomic Theory*. Oxford University Press.

Merton, R. C. 1972. An Analytical Derivation of the Efficient Portfolio Frontier. *Journal of Financial and Quantitative Analysis*, **7**, 1851 – 1872.

Merton, R. C. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica*, **41**, 867 – 888.

Merton, Robert C. 1976. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Financial Economics*, **3**, 125 – 144.

Mossin, J. 1966. Equilibrium in a Capital Asset Market. *Econometrica*, **22**, 265 – 290.

Mukherji, Anjan. 1989. Quasi–Cancave Optimisation, Sufficient Conditions for a Maximum. *Economics Letters*, **30**, 341 – 343.

Nielsen, Lars Tyge. 1987. Portfolio Selection in the Mean–Variance Model: A Note. *Journal of Finance*, **XLII**(5), 1371 – 1376.

Nielsen, Lars Tyge. 1988. Uniqueness of Equilibrium in the Classical Capital Asset Pricing Model. *Journal of Fiancial and Quantitative Analysis*, **23**(3), 329 – 336.

Nielsen, Lars Tyge. 1989. Asset Market Equilibrium with Short–Selling. *Review of Economic Studies*, **56**, 467 – 474.

Nielsen, Lars Tyge. 1990a. Equilibrium in CAPM Without a Riskless Asset. *Review of Economic Studies*, **57**, 315 – 324.

Nielsen, Lars Tyge. 1990b. Existence of Equilirium in CAPM. *Journal of Economic Theory*, **52**, 223 – 231.

Nielsen, Lars Tyge. 1992. Positive Prices in CAPM. *Journal of Finance*, **47**(2), 791 – 808.

Oh, G. 1992. *Some Results in the CAPM with Non–Traded Endowments*. Tech. rept. Financial Markets Institute, Yale University.

Pilgrim, Jan. 1998. *Eine Numerische Analyse Evolutionärer Finanzmärkte*. M.Phil. thesis, Universität Bielefeld.

Polemarchakis, H., & Siconolfi, P. 1993. Competitive Equilibria Without Ordered Preferences or Free Disposal. *Journal of Mathematical Economics*, **22**, 85 – 99.

Pratt, J. W. 1964. Risk Aversion in the Small and the Large. *Econometrica*, **32**, 122 – 136.

Rabinovitch, Ramon, & Owen, Joel. 1978. Non–Homogeneous Expectations and Information in the Capital Asset Market. *Journal of Finance*, **33**(2), 575 – 587.

Reichling, Peter. 1999. Risikomessung Durch Volatilität, Value–at–Risk Oder Lower–Partial–Moment. *Pages 223 – 238 of:* R. Eller, W. Gruber, & Reif, M. (eds), *Handbuch Bankenaufsicht und Interne Risikosteuerungsmodelle*. Stuttgart: Schäffer–Poeschel Verlag.

Ross, S. A. 1981. Some Stronger Measures of Risk Aversion in the Small and the Large with Applications. *Econometrica*, **49**, 621 – 638.

Roy, A. D. 1952. Safety First and the Holding of Assets. *Econometrica*, **20**, 431 – 449.

Schmeidler, D. 1989. Subjective Probability and Expected Utility Without Additivity. *Econometrica*, **57**, 571 – 587.

Shafer, W., & Sonnenschein, H. 1982. Market Demand and Excess Demand Functions. *Chap. 14, pages 71 – 95 of:* Arrow, K., & Intrilligator, M. (eds), *Handbook of Mathematical Economics Vol. II.* Handbook of Mathmetical Economics, vol. II. Amsterdam: North–Holland.

Sharpe, W. F. 1963. A Simplified Model for Portfolio Analysis. *Management Science*, January, 277 – 293.

Sharpe, W. F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, **19**, 425 – 442.

Sharpe, W. F. 1970. *Portfolio Theory and Capital Markets.* New York: McGraw–Hill.

Telser, L. G. 1956. Safety First and Hedging. *Review of Economic Studies*, **23**, 1 – 16.

Tetens, J. N. 1789. *Einleitung Zur Berechnung der Leibrenten und Anwartschaften.* Leipzig.

Tobin, J. 1958. Liquidity Preferences as Behavior Toward Risk. *Review of Economic Studies*, **25**, 65 – 85.

Treynor. 1961. *Toward a Theory of the Market Value of Risky Assets.* Unpublished manuscript.

Tversky, Amos, & Kahneman, Daniel. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, **5**, 297 – 323.

Wakker, Peter, & Tversky, Amos. 1993. An Axiomatization of Cumulative Prospect Theory. *Journal of Risk and Uncertainty*, **7**, 147 – 176.

# 3 The Foreign Exchange Rate in Financial Markets

**Abstract:** The main topic of this chapter is the law of motion of foreign exchange rates in financial markets implied by no–arbitrage arguments. In contrast to an important part of the empirical literature, which has tested information efficiency of foreign exchange markets due to the validity of the uncovered interest parity, it is shown that this hypothesis is generically wrong in arbitrage–free markets. The exchange rate should be understood as a change between two numéraires. Thereby the no–arbitrage condition implies that the fx–rate must be equal to the ratio of the state price deflators corresponding to the involved currencies. The complete security markets jointly determine the state price densities of both numéraires. Using this result it turns out that the exchange rate returns move along the short rate differential and its linear projection onto the adjusted excess returns of risky assets – 'adjusted' by their prices of risk and volatilities. This projection is a premium over the short rate differential, which – in the nominal context – contains the real risk premium, a risk premium due to uncertain inflation and an expression evoked by the change to nominal numéraires. Introducing into the arbitrage pricing model the real entities of an example economy illuminates how prices of risk are formed in an equilibrium. In this equilibrium model prices of risk are convex combinations of each country's specific risk factors. Thereafter the theoretically derived representation is transformed into an econometric equation and applied to the daily DEM/US–$ exchange rate. The data show some significant evidence for a linear relationship between the returns from the foreign exchange market and the excess returns of assets. There is even more empirical support of an AR(1) process for the premium, which implies an ARIMAX(1,1,1) process for the exchange rate dynamics itself.

## 3.1 Introduction

In the past decades foreign exchange (fx) markets have become more and more liberalized. This was not only 'lubricating oil' for the growing interna-

tional trade, but – even more – caused an enormously increasing international risk–sharing in financial markets. Nowadays the largest amount of all foreign exchange transactions are not (directly) linked to international trade affairs, but to investments in foreign capital or stock markets. The total value of world's exports (goods and services) was $6.3 trillion in 1995 [(IMF, 1997) p.176]. In contrast the "total net *daily* currency trading has jumped to about $1.3 trillion".[73] Furthermore just 7% (17%) of London's (New York's) fx–transactions involve non–financial companies (The Economist, 1995).
Obviously from this perspective, the interest of financial institutions in understanding the behaviour of fx–markets has remarkably increased. Often exposures in foreign currencies are undertaken not only because of speculation on de–/revaluations but also to invest in 'more' profitable projects. This causes the needs to adequately understand the risk incorporated in foreign exchange rates.

### 3.1.1 The Uncovered Interest Parity

Although the economic literature on fx–markets had an enormous upswing from the beginning of the 90s there is no convincing econometric model in sight which satisfies the demand of financial institutions for unbiased forecasts of the exchange rate and its risk. (Frankel & Rose, 1995) surveyed[74] a variety of articles concerning the empirical behaviour of exchange rates. They emphasize as a stylized fact that fundamentals cannot explain the fx–rate changes in the short run [Ch. 1.4, p. 1705]. Even a random walk outperforms many models [Ch. 1.1, p. 1691 and Ch. 1.3.2, p. 1704]. However, one can observe a very close co–movement between exchange rates and forward rates. This raises the question whether these findings imply the unbiasedness of the forward expectation hypothesis (FEH[75]). In the pure, non–logarithmic

[73] It should be noted that the connection between fx–transactions and exports is usually not one–to–one in volume. For example a role over fx–forward could cause the multiple volume than the real hedged exposure. Nevertheless the fx–volume is remarkably about 75 times higher than the volume of trade on annual basis.

[74] Also the survey of (Taylor, 1995) underpins their results.

[75] 'FEH' denotes the hypothesis that the expected value of the exchange rate conditional on the information available today is equal to the (default free) forward rate. While

form the FEH would lead to the so called Siegel–Paradox: If the expected fx–rate were equal to the corresponding forward rate in both countries, Jensen's inequality would be violated for any probability measure regarded, provided the fx–rate is uncertain, see (Siegel, 1972). This negative result let econometricians take the logarithmic form of the FEH into consideration. But the UIP gives negative forecasts [Ch. 5(1) p. 1718].[76] (Hansen & Hodrick, 1980) tested the information efficiency of fx–markets by regressing the forecast error made by the UIP on the past realizations of this forecast error. If the UIP were true, no more information could be gained from the price–history than already contained in the present expectation error. But in some of the time series they found – using a linear regression – significant information in past realizations and the constant term. This is indeed a clue against the UIP and – as will be shown in the sequel – not against the information efficiency of fx–markets.

Although theoretical critique was raised against the UIP[77] (and the FEH respectively), its seeming validity is commonly held by practitioners [e.g. (Winters, 1999)], in some amount by macro–economists [p. 16 in (Lexton *et al.*, 1998) among others] and econometricians [(Chiang, 1988) p. 214: *"The notion of rational expectations with no risk premium can be expressed formally by the UIP"* (where *"the UIP"* appears as a formula)]. To the author's knowledge the literature lacks a thorough explanation, whether the FEH respectively the UIP might hold or not;[78] exemplary is the educa-

---

the UIP denotes the logarithmic form, which means that the expected value of the log–exchange rate equals the log–forward rate.

[76] See also (Lewis, 1995) for puzzles in fx–markets and (Hansen & Hodrick, 1980) as well as (Chiang, 1988) for empirical tests of the UIP. (Hodrick, 1987) and (Meese, 1989) survey the empirical literature about the FEH and the UIP.

[77] For example in (Siebert, 1989), who analysed the risk premium in an OLG–economy. (Schmidt, 1993) showed the biasedness of the FEH for risk neutral investors. And (McCallum, 1994) proposed a sound macro–economic model where neither the FEH nor the UIP hold.

[78] Only recently the author learned about the Ph.D.–thesis by (Trojani, 1999). Trojani analyses the FEH and the UIP in the context of arbitrage pricing, which is a topic in this chapter as well. It will be mentioned if his and the author's results coincide or diverge. Trojani gives some hints to unpublished working papers about the FEH and the UIP which are also cited here for reasons of completeness.

tional article (Breuer, 1996). But it should be emphasized that in not a few articles on continuous time finance the exchange rate is modelled in such a way that conclusions about the UIP or FEH could easily be drawn, see e.g. Ch. 7 of (Musiela & Rutkowski, 1997) and (Flesaker & Hughston, 1996). These authors seemed to be not aware of the forecasting–difficulty, since this issue is more related to macro–economics or econometrics than to their concern of pricing derivatives.

### 3.1.2 Outline of this Chapter

This essay provides a more complete picture in that: it is shown in an intuitively manner that the UIP– and FEH–hypothesis are non–generic; these assertions are proved via arbitrage–pricing theory; arbitrage–pricing is embedded in an equilibrium model; and, econometric results are provided sustaining the findings. The remainder of this section describes more accurately the procedure to the sequel.

**Economic reasons:** In the following section it is shown by very intuitive arguments that the UIP is generally a wrong hypothesis. In case of the FEH one has to distinguish between real and nominal economies. It turns out just by economic arguments that the FEH holds only between two economies with real numéraires if investors were risk neutral. In this case the foreign exchange rate as well as the interest rates are deterministic. In the nominal context the FEH is non–generic.

**FX and Arbitrage Pricing:** In the next section but one the previous line of argument is more precisely explained in an extended version of the Black–Scholes model of continuous time financial markets. As mentioned above the foreign exchange rate seems to be mainly determined by financial markets, i.e. by the agents' motivation to share risks, their speculation on different expectations, etc. Fundamental factors should play an important role, but those may affect the fx–markets more indirectly through the stocks of exporters and interest rates than directly by foreign trade. Foreign trade is burdened with high transaction costs. Even if domestic and foreign consumption or investment goods or services were more or less substitutes, an arbitrage ar-

gument by trade would hardly be applicable, because transaction costs leave too much freedom for prices to vary in arbitrage–free bounds. Moreover, short–sales are feasible only for very few goods and with high transaction costs in comparison to financial assets, where short–sales can easily be realized by professional traders. But transaction costs as well as the opportunity of short–sales are evident to arbitrage pricing. Thus the determination of the exchange rate ought to be tighter in financial markets than in export–markets. Therefore the approach seems to be reasonable to hedge the foreign exchange rate in the Black–Scholes framework.
Arbitrage–free pricing is based on given, observed price processes. Those could be explained in an equilibrium model, in which the prices of risk (the pricing–measure, respectively) are determined. To apply this idea of duplicating the fx–rate one should form a complete asset market in one country consisting of its own and of the foreign country's financial assets. For this purpose assets denominated in foreign currency have to be exchanged into the domestic currency. Since the exchange rate is not an asset by itself, the foreign locally riskless bank account, exchanged to the domestic currency, will be priced by duplication instead. Knowing the foreign bank account process this results indirectly in the pricing of the foreign exchange rate. This works if the risk–factors incorporated by the fx–rate can already be traded by other securities in financial markets. One arrives at a representation of the stochastic differential equation of the exchange rate. From this representation it is not a big step to prove that the UIP can only be valid for non–generic economies. This summarizes the main idea of the third section (3.3.2).
**Equilibrium foundation:** The arbitrage pricing model leaves open the determination of the prices of risk. Those prices of risk play an important role in the law of motion of the exchange rate. Since prices of risk are linked to the underlying risks of the economy, for instance the risk inherent in production and endowments, they could be learned only in a framework of a real economy. An example economy is introduced in the forth section (3.4), which is an extension of the model developed by (Zapatero, 1995). It turns out that the prices of risk are convex combinations of the country's

specific risk–factors. A reasonable conjecture from this section is that higher profitability of investment strategies, measured as excess return over the risk free short rate (which equals the price of their risk), could be inter–linked to a devaluation of the currency, if the project contains mainly country specific risks. Country fonds are examples of such investments. Therefore the fx–risk could reduce systematically, not only by chance, the overall profitability of investments in foreign countries. This hypothesis will be tested empirically in Section 3.5 introduced next.

**Econometric validation:** After this theoretical examination the representation of the fx–rate is transformed into an econometric equation, which is then applied to the real DEM/US–$ exchange rates. The estimation is accompanied with some difficulties, because the econometric equation involves unobservable expressions like the variance of assets. Those have to be estimated separately. The empirical results partly sustain the theoretical findings, but are far from being exciting. In the last subsection before this chapter is concluded, the idea for testing information efficiency of fx–markets is taken up from the literature. The fx–returns are estimated via an 'error'–correction model, in which the unobservable premium plays the role of the forecasting 'error'. This equation is similar to that applied by (Hansen & Hodrick, 1980) and works reasonably well. It implicitly presumes a relatively slower fluctuation of prices of risk in comparison to financial prices. It is found that the premium of fx–returns over the short rate differential is well described by an auto regressive process of order one, in short notation AR(1). This implies for the fx–rate an integrated auto regressive process with moving average, in which the short rate differential is the only exogenous factor, i.e. an ARIMAX(1,1,1)–process.

The final Section 3.6 summarizes the results of this chapter and draws the conclusions. The appendices contain most of the mathematical proofs and some mathematical issues to which has been referred in the main text.

## 3.2 The Change of numéraire

In quite a few journal articles, as for instance in (Winters, 1999) and (Lexton *et al.*, 1998), the uncovered interest parity is taken as a no–arbitrage condition or a condition of financial market equilibrium. Especially, there is a huge literature in applied econometrics testing the informational efficiency of fx–markets via the UIP or the FEH, for instance (Geweke & Feige, 1979), (Chiang, 1988), and (Hansen & Hodrick, 1980) among others. This can be justified only by means of no–arbitrage or equilibrium arguments. It will be shown – in this section more intuitively – why both underlying hypotheses imply a conceptual problem. In the next section a more technical arbitrage pricing model is provided to prove some of the claims made here and to develope an econometric equation. Throughout the chapter, markets are assumed to operate frictionless and competitive.

### 3.2.1 The real economy

Consider two economies each having a real numéraire to which all prices are normalized. These numéraires are associated with a pre–defined basket of commodities, which could also consist of only a single commodity each. Variables which are specific to a country are indexed with $i \in I := \{d, f\}$, where $d$ stands for the domestic and $f$ for the foreign economy. Furthermore the index $i$ means that any one of both countries is concerned even though the set $I$ is not always mentioned. Many equations turn out to be valid for the domestic as well as for the foreign country. To stress it but to circumvent double notations, by $\neg d$ it is meant $f$, which reads $f$ is non $d$, and vice versa. Thus, in all expressions one can substitute $\neg i = f$ and $i = d$ or $\neg i = d$ and $i = f$, respectively. The real exchange rate $X_t$ at date $t$ is consequently the price of the foreign numéraire in terms of the domestic numéraire, say $X_t$ is measured in EURO–basket per US–basket. 'Currency' will stand synonymously for the real numéraire as well as for the money. Denote with $F_{t,T}$ the forward contract singed in $t$ with settlement in $T$, which involves no start–up payments. The forward contract is the obligation to deliver one unit of the US–basket in exchange for $F_{t,T}$ units of the EURO–basket at time $T$. Since forward and spot exchange rate can be seen from either country let $X_t^i$ be defined such that $X_t = X_t^d = \frac{1}{X_t^f}$ holds. The same rule shall also apply to $F$ but not to any other expression. The real instantaneous compound interest rates are named $r_t^i$. Then the discount factor in $t$ of future payments in $T$ reads $b_{t,T}^i = \exp(-\int_t^T r_u^i du)$. Consequently, the corresponding up–count factor is $b_{T,t}^i = 1/b_{t,T}^i$, which is also called country $i$'s locally riskless savings or bank account. $B_{t,T}^i$ denotes the price in currency $i$ and period $t$ of a real zero coupon bond delivering one unit of the corresponding basket at maturity $T$. Throughout the chapter it is assumed that at every point in time zero bonds are traded for all times to maturity.[79] The expectation operator with respect to a probability measure $P$ and the information available in $t$ is denoted by $E_P^t(\cdot)$ and the conditional measure by $P_t$.

Assume that financial markets do not offer arbitrage opportunities. The no–arbitrage condition implies the covered interest parity[80] (abbreviated CIP):

$$F_{t,T} = X_t \frac{B_{t,T}^f}{B_{t,T}^d}. \tag{CIP}$$

Whereas the FEH, which is defined by:

$$E_P^t\left(X_T^i\right) = F_{t,T}^i, \tag{FEH}$$

might be justified at most as an equilibrium condition of a very special economy. Only the stronger requirement of an equilibrium may ensure the FEH, which is not implied by the no–arbitrage condition as it is argued in the next paragraph. In contrast to the (CIP), which is a relation between today's prices, the FEH relates future prices to today's prices. Thus, a prognosis of future prices is only meaningful with the FEH – provided that it holds –, which makes it very attractive in practice.

---

[79] It is sufficient if zero bonds can be replicated by existing assets. This assumption is implied by but weaker than complete asset markets.

[80] The duplication of the forward exchange rate works in the following way: Lend in period $t$ the present value, $B_{t,T}^f$, of one unit of the US–basket from period $T$; change this amount into the EURO–basket by the fx–rate $X_t$; and, purchase with the resulting amount the EURO–zero–coupon–bond $B_{t,T}^d$.

The argument for the FEH to hold in an equilibrium is based on risk neutrality. Consider a risk neutral investor, who has the opportunity to sell/buy forward contracts in any amount. As long as $E_P^t(X_T)/F_{t,T} > 1$ they would buy and in the case of $E_P^t(1/X_T)F_{t,T} > 1$ they would sell the forward contract, respectively, with the intention to undo the currency exchange on the spot market in $T$. Thus, the inequalities seem to be incompatible with an equilibrium. But the transactions $X_T/F_{t,T}$ and $F_{t,T}/X_T$ are not necessarily riskless, i.e. it may well be that both $E_P^t(X_T)/F_{t,T} > 1$ and $P^t(X_T/F_{t,T} \geq 1) < 1$ are valid. Therefore, the inequalities are not a violation of the no–arbitrage condition a priori. But the no–arbitrage condition has to hold in an equilibrium of an economy, in which agents have strictly monotone preferences. It *seems* to be reasonable that the FEH applies if investors were risk neutral. In any case Siegel (Siegel, 1972) showed the following 'contradiction':

**Lemma 76 (Siegel's Paradox)** *If the future exchange rate has positive variance, the FEH is a contradiction to Jensen's inequality regardless of the probability measure entering into the expectation operator.*

**Proof.** Suppose that $E_P^t(X_T^i) = F_{t,T}^i$ is an equilibrium condition for risk neutral investors $i \in I$. This implies that $E_P^t(1/X_T) = 1/E_P^t(X_T)$. However this is a contradiction to the positive variance of the exchange rate by Jensen's inequality since $1/x$ is a strictly concave function. ■

Of course this 'paradox' would not appear if the exchange rate is deterministic. And this is claimed to be the only solution to the 'problem' of risk neutral investors. Before a compelling argument can be given, one must analyze the equilibrium condition more deeply. Three critiques can be objected against the FEH:

1. The cash flow caused by the transactions $X_T^i/F_{t,T}^i$ takes place in the future at date $T$. But the individual valuation of this transaction happens in $t$, at the same time as the decision to buy or sell the forward contract is undertaken. Accurately the cash flow must be discounted by $b_{t,T}^i$. Furthermore this transaction needs an investment of one domestic or foreign currency unit in period $T$, so that the cash flow is

actually $X_T^i/F_{t,T}^i - 1$ without starting costs (or $X_T^i/F_{t,T}^i$ with an initial investment of $B_{t,T}^i$).

2. In the – somewhat more realistic – case of risk avers investors the market valuation of portfolio returns takes their risk aversion into account via risk premiums. More elegantly the prices of risky assets can be written as the expected value of their discounted cash flows with respect to the so called risk neutral probability measure, say $Q$, the measure which 'adjusts' $P$ to the risk aversion of market participants.

3. The valuation must be understood as *relative* to the numéraire, which is here the locally riskless bank account in one of either currencies. The real interest rates generally differ in both countries. Moreover the attitude towards risk might depend on the numéraire regarded. Therefore it is reasonable to change the measure according to the currency in which the cash flow happens to be, i.e. either $Q^d$ or $Q^f$. Note that the different measures do not exclusively belong to either the domestic or the foreign investors, but to the different numéraires![81] For example, the portfolio $X_T/F_{t,T} - 1$ is the obligation to receive or deliver the EURO–basket, whereas $F_{t,T}/X_T - 1$ is the reciprocal transaction in terms of the US–basket. To the first transaction the domestic measure applies and for the latter the foreign measure is appropriate.

In an equilibrium there is at least one 'pricing' probability measure $Q^i$ for each numéraire $i = d, f$ on which the investors can agree. Putting the critique together, for any marketed contingent claim $H^i$ with a price process $H_t^i$ measured in the respective currency and paying $H_T^i$ units in $T$ one arrives at the following no–arbitrage condition:

$$E_{Q^i}^t \left[ b_{t,T}^i H_T^i \right] = H_t^i \qquad P - a.s. \tag{NA}$$

[81] Since the foreign investor has monotone preferences in the EURO–basket as well it is misleading to speak from the viewpoint of a domestic investor only because the cash flow happens to be in the EURO–basket.

To postulate equilibrium prices is a rather demanding description of the price mechanism in financial markets. A much weaker and more realistic requirement is that prices do not offer arbitrage. An arbitrage opportunity is a non–negative marketed contingent claim which involves no investments while guaranteeing positive pay–offs with positive probability. In more technical terms this means there is a contingent claim with a price process $H^i$ and there are two periods of time $s < t$ for which $H_s^i \leq 0$ and $P_s\left(H_t^i \geq 0\right) = 1$ and at least one of both prices is $P$–essentially different from zero. It is a well–known result, that arbitrage is ruled out whenever a risk neutral probability measure exists (apart from mild restrictions on the trading strategies[82]). The risk neutral probability measure implies the valuation by (NA) for all marketed contingent claims. The reverse conclusion is also true under an additional fairly mild technical assumption.[83] Therefore one speaks of (NA) as the no–arbitrage valuation.

The existence of arbitrage prevents price processes from belonging to an equilibrium. However if arbitrage opportunities do not occur there could exist an economy sustaining these prices in an equilibrium. The simplest economy which could be constructed to sustain an arbitrage–free price system is a one–consumer economy, which has a no–trade equilibrium at the prices regarded. More sophisticated examples exist due to the theorems of Sonnenschein–Mantel–Debreu and their extensions to financial markets, see the classical survey (Shafer & Sonnenschein, 1982) and the more recent, less technical contribution (Hens, 1998).

In the special case of the forward contract the initial investments $H_t^i$ and final pay–offs $H_T^i$ are

$$H_t^i = E_{Q^i}^t\left[b_{t,T}^i\right] = B_{t,T}^i \quad \text{and} \quad H_T^i = X_T^i / F_{t,T}^i.$$

[82]Ponzi games are meant here, see next Section 3.3.2.

[83]The additional requirement is the non–existence of *approximate* arbitrage, see next section 3.3.2.

This implies by (NA)

$$E^t_{Q^i}\left[b^i_{t,T}X^i_T\right] = X^i_t B^{\neg i}_{t,T}.$$

The FEH must be shown from this perspective! The next lemma considers the case of risk neutrality. Investors are regarded here as risk neutral in one of the numéraires, if their von–Neumann–Morgenstern utility function is additive separable and affine with only time–dependent coefficients in this particular numéraire.

**Lemma 77** *If investors are risk neutral with respect to one of the numéraires, the FEH holds for this particular numéraire in any equilibrium*[84]*. In this case the corresponding interest rate is a deterministic process.*

**Proof.** Suppose investors were in one numéraire risk neutral and consider only the consumption in this particular numéraire. Their marginal rate of substitution (MRS) between the consumption of two periods of time would be independent of the level of consumption and therefore independent of the realized state of the world. Then in any equilibrium the continuously compound real interest rate must be equal to the investors instantaneous MRS of the corresponding good between two periods. Otherwise investors demand in this numéraire would be suboptimal. This leads to a non–stochastic interest rate, which might only be time–depended. Risk–neutrality in one currency means that either $Q^d = P$ or $P = Q^f$. The one–sided FEH follows then from the no–arbitrage valuation (NA), because the discount factor cancels down:

$$\begin{aligned} B^i_{t,T} &= E^t_{Q^i}\left[b^i_{t,T}X^i_T/F^i_{t,T}\right] \\ &= B^i_{t,T}E^t_{Q^i}\left[X^i_T\right]/F^i_{t,T} \Leftrightarrow \\ F^i_{t,T} &= E^t_P\left[X^i_T\right]. \end{aligned}$$

■

[84] *Implicitly it is assumed here, that if there exist multiple equilibria, each is stable over the equilibrium path. Once an equilibrium settled the investors have the rational expectation to stay in this particular equilibrium.*

**Corollary 78** *If investors are risk neutral in both currencies, the real exchange rate is a deterministic process in any equilibrium, which trivially implies the FEH.*[85]

**Proof.** By the previous lemma, the FEH would hold in any equilibrium with risk neutral investors. The exchange rate must be non–stochastic in this case as Siegel's Paradox shows.
This proof by contradiction is economically not insightful. Here is the direct argument:
The MRS between the numéraires in a given period would be also independent of the level of consumption and therefore independent of the realized state of the world. This is because investors have then linear utility functions in both numéraires with non–stochastic, but possibly time–dependent coefficients. The MRS between the numéraires in a particular period, which is the ratio of two non–stochastic coefficients, has to be equal to the spot price in equilibrium at that time. But this is the exchange rate of the two currencies. ■

However if investors are risk avers one can hardly object the three refutations to the FEH stated above. The no–arbitrage valuation (NA) does also not give any support for the logarithmic form of the FEH, which is the UIP[86]:

$$E_P^t \left( \ln X_T \right) = \ln F_{t,T}. \tag{UIP}$$

The UIP is much more popular than the FEH, because it is usually argued, that the UIP avoids Siegel's Paradox. With the definition of continuously compound zero rates:

$$R_{t,T}^i := -\frac{1}{T-t} \ln B_{t,T}^i,$$

and in conjunction with the CIP the UIP transforms to the following, well

[85] *This result does not stand in contrast to (Schmidt, 1993), who showed the biasedness of the forward as a predictor of the spot exchange rate for risk neutral agents, because he considered inflation risk. The corresponding result follows below.*

[86] Note that the basis of the logarithm is irrelevant by $\log_b x = \ln x / \ln b$.

known equation:

$$E_P^t(\ln X_T - \ln X_0) = (T-t)\left(R_{t,T}^d - R_{t,T}^f\right).$$

Actually, the naming "Uncovered *Interest Parity*" is due to the expression $R_{t,T}^d - R_{t,T}^f$ in this equation.[87]

Indeed it has been shown that the premises of Siegel's Paradox contradict, namely positive variance of the exchange rate and risk neutrality. Therefore the ad–hoc argument of avoiding Siegel's paradox in favour of the UIP is meaningless.

It has not been shown yet that there cannot exist exceptional economies in which accidentally the FEH or the UIP could hold in equilibrium (e.g. FEH and UIP obviously hold in any economy without risk). The claim here is that both hypothesis are not generic:

**Conjecture 79** *If investors are not risk neutral with respect to both numéraires, the FEH and the UIP do not hold in general.*

The proof of this conjecture with regard to the UIP is one of the main results in the next section (see Proposition 93).

Observe that any pay–off can be valued by no–arbitrage in either of both currencies. Therefrom one is able to calculate the change of measure, which has to be taken into account, if in the valuation of random returns the numéraire changes:

$$\begin{aligned}\forall \text{ marketed } H^i : H_t^i &= E_{Q^i}^t\left[b_{t,T}^i H_T^i\right] \\ &= X_t^i E_{Q^{\neg i}}^t\left[b_{t,T}^{\neg i} H_T^i / X_T^i\right]. \quad \text{(VAL)}\end{aligned}$$

[87] Often the UIP is re–formulated with discrete compound zero rates $\tilde{R}$, which are defined by $B_{t,T}^i = \left(1+\tilde{R}_{t,T}^i\right)^{-(T-t)}$. Then usually the approximation: $R_{t,T}^d - R_{t,T}^f = \ln\left(1+\tilde{R}_{t,T}^d\right) - \ln\left(1+\tilde{R}_{t,T}^f\right) \approx \tilde{R}_{t,T}^d - \tilde{R}_{t,T}^f$ is applied.

This is implied by the change of measure:[88]

$$\left.\frac{dQ^f}{dQ^d}\right|_t = \frac{X_t}{X_0}\frac{b_{0,t}^d}{b_{0,t}^f}.$$

Since the right hand side starts in one and is strictly positive the following lemma remains to be shown for establishing the change of measure .

**Lemma 80** *Given that all quantities admit integrability, the no–arbitrage valuation (NA) implies that* $\frac{X_t^i}{X_0^i}\frac{b_{0,t}^i}{b_{0,t}^{\neg i}}$ *is a* $Q^i$*–martingale.*[89]

**Proof.** This is an immediate implication of (NA) by taking the locally riskless bank account $1/b_{0,t}^{\neg i}$ as the asset $H_t^{\neg i}$ to be valued in currency $i$:

$$\frac{X_t^i}{X_0^i}\frac{b_{0,t}^i}{b_{0,t}^{\neg i}} = E_{Q^i}^t\left(\frac{X_T^i}{X_0^i}\frac{b_{0,T}^i}{b_{0,T}^{\neg i}}\right).$$

■

This means that under the risk neutral measure $Q^i$ the instantaneous return of the exchange rate $X_t^i$ is equal to that of $\frac{b_{0,t}^{\neg i}}{b_{0,t}^i}$, whose return on the other hand is the interest rate differential $r_t^i - r_t^{\neg i}$. Thus, the UIP holds with respect to the risk–neutral measures $Q^i$ only instantaneously.

Suppose that for each numéraire a so–called likelihood ratio process $z^i$ exists which changes the physical probability measure $P$ into the risk neutral measure $Q^i$, i.e. $E_{Q^i}^t(x_T) = E_P^t\left(x_T\ z_T^i/\ z_t^i\right)$. Since the price in $t$ of any contingent claim paying in $T$ is given by $H_t^i = E_P^t\left(H_T^i b_{t,T}^i\ z_T^i/\ z_t^i\right)$, $b_{0,t}^i z_t^i$ is called the state–price deflator. From the previous lemma follows:

$$E_P^t\left(\frac{Y_T z_T^d}{Y_t z_t^d}\right) = E_P^t\left(\frac{Y_t z_T^f}{Y_T z_t^f}\right) \quad P-a.s. \quad \text{with } Y_t = \frac{X_t}{X_0}\frac{b_{0,t}^d}{b_{0,t}^f}.$$

[88] The change of measure is not necessarily unique. All those changes of measure come into question, for which the right hand side is multiplied by a positive local martingale $u$, which starts in one and satisfies: $\forall$ marketed $H^i \perp_{Q^i} \left(u^i - 1\right)$.

[89] *This result resembles Lemma 88 proved in Section 3.3.2.*

Although it would not be unique by that argument, a reasonable choice for the exchange rate solving this equation is

$$X_t = X_0 \frac{b^f_{0,t}}{b^d_{0,t}} \frac{z^f_t}{z^d_t}.$$

This equation stresses best the role of the exchange rate as the vehicle which consistently transfers different numéraires into each other. Since the bank account $1/b^{\neg i}_{0,t}$ is not the only asset $H^{\neg i}_t$ valued this way, the proposed is indeed a unique solution if domestic and foreign markets were complete. Thus, in case of complete markets the UIP imposes the following restriction onto the exchange rate dynamics:

$$\begin{aligned} E_P\left(\ln X_T\right) - \ln X_0 &= E_P\left(\ln b^f_{0,t} z^f_t\right) - E_P\left(\ln b^d_{0,t} z^d_t\right) \\ &= \ln B^f_{t,T} - \ln B^d_{t,T} \quad \text{by UIP} \\ &= \ln E_P\left(b^f_{0,t} z^f_t\right) - \ln E_P\left(b^d_{0,t} z^d_t\right). \end{aligned}$$

Neither rational expectations nor absence of arbitrage can provide a justification to pull up 'ln' in front of the expectation operator from the right hand side of the first to the last equation. Risk neutrality can, but just because it implies deterministic processes. In the sequel it is shown, that this restriction turns the UIP into a non–generic property, even in incomplete markets.[90]

### 3.2.2 The nominal economy

Prices are observable only in nominal quantities. Money introduces further risk into the economy, namely the risk of inflation. In other words, nominal riskless assets are risky in real terms. The equilibrium conditions must pay attention to this fact. However the neutrality of money is assumed in the

[90] The conditions implying the UIP under the measures $P$ and $Q$, respectively, are quite different in (Trojani, 1999). Trojani's Proposition 4.7 demands that $b^i_{0,t} z^i_t$ and $b^i_{0,t}$, respectively, are conditionally uncorrelated with the exchange rate $X^i_t$ for the UIP to hold under $P$ and $Q$, respectively. This is questionable, because the logarithm in the UIP makes correlation irrelevant, and the exchange rate is generally not uncorrelated with those expressions.

following. This means that the real allocation is independent (not in the probabilistic sense!) of inflation.[91] It is abstracted from modelling an explicit money stock. Instead, let the price indices of the numéraires being labelled as $q_t^i$. Nominal terms are indexed by a hat on top. The nominal exchange rate is by arbitrage–free valuation

$$\hat{X}_t = X_t \frac{q_t^d}{q_t^f}.$$

A nominal bond pays at maturity the reciprocal of the price index in real terms, which is a random variable. The no–arbitrage pricing formula of nominal bonds and the nominal instantaneously compound interest rate[92] read:

$$\begin{aligned} \hat{B}_{t,T}^i &= q_t^i E_{Q^i}\left(b_{t,T}^i / q_T^i\right), \\ \hat{r}_t^i &= -\lim_{T \to t} \frac{\partial}{\partial T} \ln \hat{B}_{t,T}^i. \end{aligned}$$

The price of the nominal forward rate is again given by the no–arbitrage condition:

$$\begin{aligned} \hat{F}_{t,T} &= \hat{X}_t \hat{B}_{t,T}^f / \hat{B}_{t,T}^d \\ &= F_{t,T} \left[ E_{Q^f}\left(b_{t,T}^f / q_T^f\right) B_{t,T}^d / \left( E_{Q^d}\left(b_{t,T}^d / q_T^d\right) B_{t,T}^f \right) \right], \end{aligned}$$

which is different from the nominal price of the real forward: $F_{t,T} q_t^d / q_t^f$. Hence, in the nominal context the no–arbitrage valuation for contingent claim

[91] Herewith it is also assumed that the risk of inflation does not change the space of marketed contingent claims, i.e. the risk of inflation is already spanned by some real securities. But this assumption is harmless in the complete market context assumed here. In incomplete markets nominal risk might change the real equilibrium allocation, see e.g. (Geanakoplos & Mas–Colell, 1989).

[92] The valuation of bonds shows, that in the case of risk avers investors the nominal interest rate is in general not the sum of the real interest rate and the expected rate of inflation, which is another popular, but (generally) wrong 'no-arbitrage' hypothesis. Indeed it is shown in Lemma 94 below, that the nominal interest rate also contains a risk premium.

$\hat{H}^i$ denominated in currency $i$ is:

$$\hat{H}^i_t = q^i_t E^t_{Q^i}\left[b^i_{t,T}\hat{H}T^i/q^i_T\right]$$
$$= E^t_{\hat{Q}^i}\left[\hat{b}^i_{t,T}\hat{H}^i_T\right] \qquad \text{for } Q^i \to \hat{Q}^i$$
$$= \hat{X}^i_t E^t_{\hat{Q}^{\neg i}}\left[\hat{b}^{\neg i}_{t,T}\hat{H}^i_T/\hat{X}^i_T\right] \quad \text{for } \hat{Q}^i \to \hat{Q}^{\neg i}. \quad \text{(NomVal)}$$

The difference here is that there is a hat on the measures behind the second equality sign, which refers to the nominal point of view. Indeed the numéraires change now from the locally riskless real currency account, the baskets, to the risky nominal money account, the moneys US–$ and €. With this change, also the risk neutral measures must change from the real valuation (first equal sign) to the nominal valuation (second equal sign). Whereas the third equality sign corresponds to the change of measure between the moneys.

**Lemma 81** *Suppose the nominal no–arbitrage valuation by (NomVal) holds. Then, provided integrability, the likelihood ratio processes* $\frac{b^i_{0,t}q^i_0}{\hat{b}^i_{0,t}q^i_t}$ *and* $\frac{\hat{X}^i_t b^i_{0,t}q^i_0}{\hat{X}^i_0\hat{b}^{\neg i}_{0,t}q^i_t}$ *change the measure* $Q^i$ *to* $\hat{Q}^i$ *and* $\hat{Q}^{\neg i}$*, respectively, and* $\frac{\hat{X}^i_t\hat{b}^i_{0,t}}{\hat{X}^i_0\hat{b}^{\neg i}_{0,t}}$ *changes* $\hat{Q}^i$ *to* $\hat{Q}^{\neg i}$*.*

**Proof.** All likelihood ratio processes start in one and – integrability presumed – are strictly positive martingales by the no–arbitrage valuation (NomVal) with respect to $Q^i$ or $\hat{Q}^i$, respectively. ■

The valuation of the 'back–changed' forward, $\hat{H}^i_T = \hat{X}^i_T/\hat{F}^i_{t,T}$, with investment costs of one unit of currency $i$ in period $T$ gives the corresponding nominal no–arbitrage valuation:

$$E^t_{\hat{Q}^i}\left[\hat{b}^i_{t,T}\hat{X}^i_T\right] = \hat{B}^i_{t,T}\hat{F}^i_{t,T}.$$

Since the neutrality of money is assumed it is expected that the FEH holds for the real exchange rate if investors are assumed to behave risk neutral.

The nominal risk neutral measure $\hat{Q}^i$ gives the necessary degree of freedom to transform the nominal no–arbitrage valuation into the real FEH:

$$\begin{aligned}
E^t_{\hat{Q}^i}\left(\hat{b}^i_{t,T}\hat{X}^i_T\right) &= E^t_{Q^i}\left(\frac{b^i_{t,T}q^i_t}{\hat{b}^i_{t,T}q^i_T}\hat{b}^i_{t,T}X^i_T\frac{q^i_T}{q^{\neg i}_T}\right) && \text{by } \hat{Q}^i \to Q^i \\
&= E^t_P\left(\frac{b^i_{t,T}q^i_t}{q^{\neg i}_T}X^i_T\right) && \text{by } Q^i = P \\
&= E^t_P\left(1/q^{\neg i}_T\right)b^i_{t,T}q^i_tX^i_T && \text{by Lemma 77 and Corollary 78,} \\
\hat{F}^i_{t,T}\hat{B}^i_{t,T} &= \hat{X}^i_t\hat{B}^{\neg i}_{t,T} \\
&= X^i_t\frac{q^i_t}{q^{\neg i}_t}q^{\neg i}_tE^t_P\left(1/q^{\neg i}_T\right)b^{\neg i}_{t,T} && \text{by the same steps } \Leftrightarrow \\
X^i_T &= F^i_{t,T}\,. \qquad \checkmark
\end{aligned}$$

For the nominal exchange rate the FEH is indeed not an equilibrium condition for risk neutral investors, because *goods* count in investors' preferences and not money.[93] By arbitrage pricing the nominal exchange rate incorporates only the risk of inflation, because the real exchange rate is riskless. Assuming stochastic inflation, Siegel's Paradox applies to the FEH for the nominal exchange rate as well – even for risk neutral investors, whereas the UIP holds in a very exceptional case.

**Lemma 82** *In an economy with risk neutral investors the UIP holds for the nominal exchange rate if and only if*

$$E^t_P\left(\ln q^d_T\right) + \ln E^t_P\left(1/q^d_T\right) = E^t_P\left(\ln q^f_T\right) + \ln E^t_P\left(1/q^f_T\right).$$

[93] This is true for the 'homo oeconomicus', although one might have a different impression from real live. For instance, experiments undertaken by (Fehr & Tyran, 2000) show some degree of money illusion of participants: *...seemingly innocuous differences in pay-off representation cause pronounced differences in nominal prices inertia indicating the behavioural importance of money illusion.*

**Proof.** Suppose investors were risk neutral, i.e. $Q^i = P$. Then by arbitrage–free pricing of the exchange rate and Corollary 78 it holds that

$$\begin{aligned} E_P^t\left(\ln \hat{X}_T\right) &= \ln X_T + E_P^t\left(\ln q_T^d\right) - E_P^t\left(\ln q_T^f\right) \quad \stackrel{\ln \mathrm{FEH!}}{=} \\ \ln \hat{F}_{t,T} &= \ln\left[F_{t,T} E_P\left(b_{t,T}^f / q_T^f\right) B_{t,T}^h / \left(E_P\left(b_{t,T}^h / q_T^h\right) B_{t,T}^f\right)\right] \\ &\quad \text{by } Q^i = P \text{ and Lemma 77 this is} \\ &= \ln F_{t,T} + \ln E_P^t\left(1/q_T^f\right) - \ln E_P^t\left(1/q_T^h\right). \end{aligned}$$

Since the FEH holds for risk–neutral investors in the real economy, this equation transforms to the equation stated in the lemma. ■

If investors behave risk averse, nothing could explain why the FEH or the UIP shall hold. The reasoning against the UIP given for the real economy (right before this subsection) still holds in the nominal economy. Therefore one can claim:

**Conjecture 83** *The FEH and the UIP hold for the nominal exchange rate only in exceptional cases.*

The proof of this conjecture with regard to the UIP is one of the main results in the next section (see Proposition 93 and Lemma 94). For the FEH the conjecture has already been answered by this subsection.
So far it has been shown that the FEH and UIP are at best equilibrium conditions and not no–arbitrage conditions indeed. Moreover FEH and UIP might be valid only in exceptional economies. Herewith the mathematical finance model of the next section is intuitively prepared .

## 3.3 FX and Arbitrage Pricing

The notation of the previous section will be retained but extended here. In this section the valuation of the exchange rate is examined by the no–arbitrage condition in the Black–Scholes framework. For this reason it is not necessary to distinguish between the real and the nominal economy, because the no–arbitrage conditions are in either numéraires, i.e. baskets and moneys, the same. Thus, the exchange rate $X_t$ can be seen with both meanings, whereas one should be careful with the interpretation of the results. If appropriate the exchange rate is explicitly interpreted as nominal or real.

This section is divided into two parts. In the first subsection a version of the Black–Scholes model is introduced and – very briefly – the concepts of arbitrage–pricing theory are outlined to make this essay self–contained. The main results follow in the last part, where the exchange rate will be analysed in the model introduced first.

### 3.3.1 Arbitrage Pricing Model

#### Stochastic Environment

**Uncertainty:** The two economies have continuous trading financial markets at any time $t$ in a time interval $\mathbb{T} := [0, \mathcal{T}], 0 < \mathcal{T} < \infty$. Uncertainty is characterized by a $m$–dimensional standard Brownian motion $\left(W_t, \mathcal{F}_t^W\right)_{t\in\mathbb{T}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ with state space $\Omega$, the $\sigma$–algebra $\mathcal{F}$ containing all measurable events, and the measure $P$. The filtration $(\mathcal{F}_t^W)_{t\in\mathbb{T}}$ is generated by the Brownian motions $(W_t)_{t\in\mathbb{T}}$, where $\mathcal{F} = \mathcal{F}_{\mathcal{T}}^W$, and $\mathcal{F}_0^W$ is augmented by all $P$–null subsets of $\Omega$. Moreover, events in $\mathcal{F}_0^W$ are known from the beginning to be true or not. Thus, there is no uncertainty 'today' about today's realizations. All this is common knowledge.

**Stochastic Processes:** All equations involving stochastic terms shall hold $P$–almost surely. It will become clear from the context what variables will be random. Therefore the state $\omega \in \Omega$ is suppressed as an argument. Stochastic differential equations (SDEs) should be understood as the short hand notation of the integral equations. Summations of integrals are written as vector–multiplications of a row–vector of integrands with the column–vector of differentials.

Only Markovian Itô–processes of diffusion type[94] are considered:

$$dx^i(t) = \mu^{xi}(x(t), t)\,dt + \sigma_t^{xi}(x(t), t)\,dW_t,$$

whose dimension is $n^i, i = d, f$, and which start in a constant vector $x_0^i$. Drift $\mu^{xi} : \mathbb{T} \times \mathbb{R}^{n^d+n^f} \to \mathbb{R}^{n^i}$ and diffusion $\sigma^{xi} : \mathbb{T} \times \mathbb{R}^{n^d+n^f} \to \mathbb{R}^{n^i} \times \mathbb{R}^m$ of the country specific vector–process $x^i$ are Borel–measurable functions of $t$ and $x$, where $x_t$ denotes the joint column–vector $\left(\left(x_t^d\right)^T, \left(x_t^f\right)^T\right)^T$. The *index* time and the vector–process $x$ as an *index* are also meant to be an *argument*, i.e. let: $x_t \equiv x(t)$ and $\mu_t^x \equiv \mu_t^{xi}(x_t) \equiv \mu^{xi}(t, x(t))$ hold further on. Especially in SDEs one can suppress the time index, i.e. make use of the short hand notation $\mu^{xi} \equiv \mu^{xi}(t, x_t)$. The superscript $\cdot^x$ shall specify the process in concern and $i \in I$ the country the process belongs to. Without a superscript $\cdot^x$ they belong to the financial assets defined afterwards. Indices are not separated by comma if they appear in the superscript.

Note that the conditional expectation $E_P(\cdot\,|\mathcal{F}_t)$ is still indicated by $E_P^t(\cdot)$, but the filtration is now limited to the generated filtration of observed quantities $x$, i.e. $(\mathcal{F}_t) := (\mathcal{F}_t^x)$. Let $P_t$ denote the conditional probability measure $P|_{\mathcal{F}_t}$ and $\lambda$ the Lebesgue–measure on $(\mathbb{T}, \mathbb{B}(\mathbb{T}))$. Let the set of square integrable, $\left(\mathcal{F}_t^W\right)-$adapted and $\mathcal{F} \otimes \mathbb{B}(\mathbb{T})-$measurable processes be denoted by $\mathcal{L}^2(P \times \lambda)$ and that of bounded processes by $\mathcal{L}^\infty(P \times \lambda) \subset \mathcal{L}^2(P \times \lambda)$. Likewise is $L_t^2(P)$ and $L_t^\infty(P)$ defined for $\mathcal{F}_t-$measurable random variables. A measurable event $\mathbb{E} \in \mathcal{F}$ is called $P$–essential if it occurs with positive probability, $P(\mathbb{E}) > 0$.

For all given stochastic differential equations the existence of a unique strong solution is assumed, which is then continuous, $\left(\mathcal{F}_t^W\right)-$adapted[95], and Markovian. For this it suffices to assume that

[94] Itô–processes: trend and diffusion are $\left(F_t^W\right)$–adapted, jointly $F \otimes \mathbb{B}(\mathbb{T})$ measurable random processes; diffusion type: $\mu, \sigma$ are adapted to the filtration $(F_t^x)$ generated by $x$; Markovian: $\mu, \sigma$ are independent of the past ("anticipative").

[95] This implies $F \otimes \mathbb{B}(\mathbb{T})-$measurability of $x, \mu$ and $\sigma$ as well as progressive measurability

$\mu$ and $\sigma$ satisfy the linear growth conditions uniformly in $\mathbb{T}$ and are locally Lipschitz. Let $\|A\|$ denote the Euclidean norm of a matrix or a vector $A : \|A\| = \left[tr\left(A^T A\right)\right]^{1/2}$, where the upper index $T$ denotes the transpose of a matrix or a vector. Those conditions are defined as follows [see (Chung & Williams, 1990) on p. 229]:

**Definition 84** *A function $f : \mathbb{T} \times \mathbb{R}^N \rightarrow \mathbb{R}^M$ is said to satisfy the linear growth condition in $x$ uniformly in $t$ if there is a $k > 0$ such that for all $t \in \mathbb{T}$ and $x \in \mathbb{R}^N$ :*

$$\|f(t,x)\|^2 \leq k\left(1 + \|x\|^2\right).$$

*A function $f : \mathbb{T} \times \mathbb{R}^N \rightarrow \mathbb{R}^M$ is said to be locally Lipschitz if for all $K > 0$ there is a constant $k_K > 0$ such that for all $s, t \in \mathbb{T}$ and $x, y \in \mathbb{R}^N$ :*

$$\begin{aligned} \|(s,x)\| + \|(t,y)\| &\leq K \quad \textit{implies} \\ \|f(s,x) - f(t,y)\| &\leq k_K \|(s,x) - (t,y)\|. \end{aligned}$$

**Assumption** Trend $\mu$ and diffusion $\sigma$ satisfy the linear growth conditions and are locally Lipschitz.

For the purpose of a strong solution the growth condition for the trend $\mu$ can be replaced by the weaker requirement $x^T \mu(t,x) \leq k\left(1 + \|x\|^2\right)$, but the stronger assumption is needed below to establish the existence of a martingale measure.

**Critical Remarks**

The next four remarks shall take a critical view on the mathematical assumptions made so far.

**Modelling:** Why analyze foreign exchange rates in an environment with continuous trading, a continuous state space and with continuous sample

of $\mu$ and $\sigma$ as functions of the Brownian motions.

path? The answer to this question has two parts, because it could be done more general or simpler. More complicated processes, like semi–martingales in infinite time, are not included because some of the results needed in the following (like the existence of an equivalent martingale measure) are so far unproven, and it unnecessarily complicates the calculations.[96] Simpler approaches are two or multiperiod discrete time, discrete state space models. The notation of the latter is too fussy, but it is also appropriate. A two period model is too limited for an econometric specification, since it does not reflect the flow of information appropriate for time series analysis.

**Observability:** It is important to stress which quantities are observable, because realizations are information revealing. The only observable quantities in this context are the country specific processes and not the Brownian motions themselves. An important consequence is that expectations cannot condition on the Wiener filtration $\left(\mathcal{F}_t^W\right)$ but only on the filtration generated by the observable processes, $(\mathcal{F}_t^x)$, which is not finer: $\forall t \in \mathbb{T} : \mathcal{F}_t^x \subseteq \mathcal{F}_t^W$. Only if the diffusion $\left(\left(\sigma^{xd}\right)^T, \left(\sigma^{xf}\right)^T\right)^T$ is of *full rank m* the information structure will be not coarser, which implies $\mathcal{F}_t^x = \mathcal{F}_t^W$, see (Liptser & Shiryaev, 1977). This corresponds to complete markets if $x$ were defined as asset prices, an assumption which is made in the sequel, see (Harrison & Kreps, 1979).

Moreover the functional form of drift and diffusion is common knowledge. This is actually not a realistic assumption, but if drift and diffusion were not known, any strategies resting upon this information would be unavailable. For instance, imagine an option on a stock, whose expected average volatility until maturity is unobservable, because the functional form of the diffusion is unknown. The option could be priced only by an estimated or implicit volatility. Arbitrage strategies which could exploit any deviation from the true average volatility are unavailable. The error caused by any approximative 'arbitrage' strategy prevents it from being a *true* arbitrage strategy. In this case only 'simple' arbitrage strategies are possible, e.g. if

[96] A generalization to Itô–processes or semi–martingales might be possible with more technical effort. (Huang & Pagès, 1992) have shown that under some other mild technical conditions the model can be extended to an infinite horizon setup. However, to show the qualitative results the choice made in the text is absolutely sufficient.

the put/call–parity is violated. Finance models therefore idealize the world presuming too much knowledge. In real trading investors have to learn the law of motion – usually from historical data by econometric methods for a presumed model.

**Markovity:** The interpretation of the Markov property is generally seen as evident in financial markets, since it stresses that prices in $t$ reflect all information available up to time $t$. (Fama, 1984) distinguishes between three levels of market efficiency. He calls prices weakly efficient if they reflect all information of their past realizations, semi–strongly efficient if they contain additionally all publicly known information and strongly efficient if also private information is mirrored in prices. Markovity is a convincing view if realizations along a path before period $t$ do not affect fundamentals any more when the realization in $t$ is learned since they are 'sunk'. Then only the current state is the starting point for the future development. E.g. in a Markovian sense it is unimportant how a firm gets into a position, for its future development only the position it is currently into counts, which is fully reflected in the current price. This is for example not valid for GARCH$(l, k)$–volatilities, i.e. volatilities which depend on past volatilities and shocks. This property, often viewed as describing a stylized fact of asset prices very well, is obviously not Markovian if $(l, k) > (1, 1)$ and preserves under arbitrary changes of measure. Even if Markovity is not essential to this analysis, this chapter is limited to Markovian SDEs for simplicity.

**Strong Solution of SDEs:** The distinction between strong and weak solutions of a stochastic differential equation is that the former requires a solution $x$ for a given filtered probability space, an adapted Brownian motion and a specified (possibly random) starting value $x_0$, whereas the latter concept counts also the filtered probability space and the Brownian motion (not the distribution of the starting value) as a part of the solution. The concept of a weak solution could make sense for instance in modelling self–fulfilling prophecies: If agents believe in a particular law of motion and act in a way as if it were valid, then it might endogenously settle down in a weak solution. However, a predetermined probability space is taken here as an exogenous fundamental characteristic of the economy, e.g. determined by the risks of some production technologies. This restriction requires the existence of a strong solution. The existence and uniqueness result is for example proven in [(Chung & Williams, 1990), Theorem 10.6 on p. 229 with the Extensions on p. 234]. In the one–dimensional case these conditions can considerably be weakened, see [(Yamada & Watanabe, 1971) in (Karatzas & Shreve, 1988), Proposition 2.13, p. 291]. This is important for some processes involving square root diffusions which are not locally Lipschitz. The strong Markov property of the solution with respect to its generated filtration is established in [(Chung & Williams, 1990) Theorem 10.12 on p. 242].

**Financial Markets** Beside the locally riskless savings accounts, $1/b^i$, markets consist of $\hat{n}^d$ domestic and $\hat{n}^f$ foreign risky assets, which could be bonds or stocks for example. Asset prices in local currency are denoted by $H^i_k,\ k = 1, \ldots, \hat{n}^i$. Without a numbering index the column–vector of the corresponding variables is meant, e.g. the vector–process of asset prices is $H^i$. With $\left(n^d, n^f\right) \leq \left(\hat{n}^d, \hat{n}^f\right)$, where $n^d + n^f \leq m$, any selection of domestic and foreign assets is denoted. The assets in the selection shall not span the locally riskless bank account in their currency and should have a diffusion matrix with maximal rank $n^d + n^f$, so that there are no redundancies in this set. The short–rates $r^i$ are assumed to be bounded processes. The connection between the two financial markets is possible through a strictly positive exchange rate $X$, $X_t > 0\ P \times \lambda - a.s.$ The aim is to derive an explicit formula for the trend and the volatility of the exchange rate depending on the price processes of foreign and domestic assets and the short rates. For this reason, an Itô–process of diffusion type is assumed for the exchange rate. This choice is verified in the next subsection. Also assume Itô–processes of diffusion–type for $H^i$ and $r^i$ as well. Thus, with

$$
\begin{aligned}
y_t &= \left(\left(H^d_t\right)^T, r^d_t, \left(H^f_t\right)^T, r^f_t, X_t\right)^T \quad \text{and} \\
f^y\left(t, y_t\right) &= \Big[\left(f^d\left(t, y_t\right)\right)^T, \left(f^{rd}\left(t, y_t\right)\right)^T, \\
&\qquad \left(f^f\left(t, y_t\right)\right)^T, \left(f^{rf}\left(t, y_t\right)\right)^T, \left(f^X\left(t, y_t\right)\right)^T\Big]^T
\end{aligned}
$$

for $f=\sigma,\mu$, which satisfy the linear growth conditions and are locally Lipschitz (in the range of $y$), the equation

$$dy_t = \mu_t^y dt + \sigma_t^y dW_t$$

holds $P \times \lambda - a.s.$ Note that trend and diffusion are column–vectors and matrices of desired size. Hence, a unique strong solution is guaranteed.
It turns out that a re–definition of exchange rate's SDE simplifies the notion considerably. From now on let its trend be $X_t\mu_t^X$ and its diffusion be $X_t\sigma_t^X$. Since the situation should be regarded as symmetric between domestic and foreign financial markets, above linear growth conditions and local Lipschitz continuity are also assumed to hold if the SDE of $X$ is replaced by that for $X^{-1}$. The SDE of the latter process has trend $\frac{1}{X_t}\left(\left\|\sigma_t^X\right\|^2-\mu_t^X\right)$ and diffusion $\frac{1}{X_t}\sigma_t^X$, for which those conditions shall hold too. A zero or infinite exchange rate would offer arbitrage. Therefore it is assumed that the exchange rate is bounded away from zero and infinity.
Whenever in the sequel an assumption holds for all $y \in Range\,(H)$, this range describes the subset of $\mathbb{R}^{\hat{n}^d+\hat{n}^f+3}$ in which the dimensions corresponding to the interest rates and the exchange rate are closed intervals, i.e. the space these processes actually live in $P\times\lambda - a.s.$ This reflects the variables' smaller range particularly used in conditions. Any function, e.g. 'ln', in front of a vector $Y=(Y_i)_{i\in I}$ is applied to all its entries, i.e. $\ln Y = (\ln Y_i)_{i\in I}$. A bold $\mathbf{0}$ respectively $\mathbf{1}$ stands for a column–vector or a matrix filled with zeros respectively ones of appropriate length.
Let the vector process $\hat{H}^i$ denote the prices of the assets in currency $i$, that is $\hat{H}^i = \left(1/b^i, (H^i)^T, X^i/b^{\neg i}, X^i\left(H^{\neg i}\right)^T\right)^T$, whose trend and diffusion are denoted by $\hat{\mu}^i$ and $\hat{\sigma}^i$, respectively. Provided that the martingale measures exist with respect to both numéraires they are named $Q^i$. In this case the discounted €– respectively \$–valued securities, $\hat{H}_t^i b_{0,t}^i$, are $Q^i$–martingales.

**Remark 6 (Bonds)** *If both markets consist only of bonds, the underlying model could be of the Heath–Jarrow–Morton type, see (Heath* et al.*, 1992), i.e. one could assume a family of stochastic processes representing forward* *rate movements. These processes uniquely specify the spot rate process and any bond price process, see for example (9) in (Heath* et al.*, 1992). One can enlarge the economy and include assets into the model by adding SDEs of diffusion type. The main observation is the fact that the conditions on existence and uniqueness of a martingale measure (see C4 and C5 in (Heath* et al.*, 1992)) are still satisfied for this enlarged economy. Therefore, from now on the phrase 'asset' is used to denote a traded bond as well.*[97]

**Technical Issues of Arbitrage Pricing**

To make this article self–contained some well known definitions and results are introduced from continuous time Finance. This is done in a brief manner since it is treated in more detail in some good text books. In the next paragraph, 'Preliminaries', the 'proofs' provide only some selected arguments, whenever a reference is also cited, which ought to convince the reader. For the mathematical details refer to [(Musiela & Rutkowski, 1997) Ch. 10 and (Duffie, 1996) Ch. 6] as well as the references cited there.

**Preliminaries** A trading strategy $\theta$ is a $\mathbb{R}^{2+\hat{n}^d+\hat{n}^f}$–valued, progressively measurable process with an accompanying value process $v^{i\theta} = \theta^T \hat{H}^i$. A trading strategy is admissable if its discounted value process $\theta^T \hat{H}^i b^i$ is $P \times \lambda-$almost surely bounded from below by a constant credit constraint[98] and if each of its components $\theta_k \hat{H}^i_k$ is in $\mathcal{L}^2 (P \times \lambda)$. A consumption plan is a pair $(c, C)$ which consists of a meantime per period, progressively measurable consumption process $c \in \mathcal{L}^2 (P \times \lambda)$ and a final consumption $C \in L^2_T (P)$ . A consumption plan is said to be feasible if an admissable trading strategy $\theta$ exists which finances the consump-

[97] *This hint is due to Andreas Löffler.*

[98] The credit constraint is allowed to depend on the trading strategy only.

tion plan[99], i.e.:

$$\begin{aligned} v_t^{i\theta} &= v_0^{i\theta} + \int_0^t \theta_u^T d\hat{H}_u^i - \int_0^t c_u du \quad \text{and} \\ v_{\mathcal{T}}^{i\theta} &= C \qquad P \times \lambda - a.s. \end{aligned}$$

A self–financing trading strategy re–invests all trading gains made between 0 and $\mathcal{T}$, i.e. $c = 0 \ P \times \lambda - a.s.$; a financing $\theta$ is called here meantime self–financing between $s$ and $t$ if $(c_u)_{u \in [s,t]} = 0$. Asset markets are named complete if all consumption plans $(c, C) \in \mathcal{L}^\infty (P \times \lambda) \times L^\infty (P)$ are feasible. Note that the set $\mathcal{L}^\infty (P \times \lambda) \times L^\infty (P)$ is dense in $\mathcal{L}^2 (P \times \lambda) \times L^2 (P)$ but smaller.[100] The two following results show how to ensure complete financial markets in the connected asset markets of the two economies.

**Lemma 85** *Financial markets are complete if and only if $P \times \lambda -$ almost surely there is a set of assets such that the rank of their diffusion matrix is equal to the dimension of the Brownian processes* and *in their span lies a locally riskless bank account.*

**Corollary 86** *Assuming the existence of a locally riskless bank account, financial markets are complete if and only if the diffusion matrix*

$$\begin{pmatrix} \sigma_t^d \\ \sigma_t^f \\ \sigma_t^X \end{pmatrix} \text{ has rank } m \ P \times \lambda - a.s.$$

[99] Depending on agents' utility functions the consumption plan is sometimes restricted to be non–negative $P \times \lambda - a.s.$

[100] Completeness holds at most for $\mathcal{L}^1 (Q^i) \times L^1 (Q^i) \cap \mathcal{L}^2 (P) \times L^2 (P)$, see (Dothan, 1990), but this involves the martingale measure, which is an endogenous quantity. Nevertheless, $\mathcal{L}^\infty (P) \times L^\infty (P)$ is also dense in $\mathcal{L}^1 (Q^i) \times L^1 (Q^i)$, because $P \sim Q^i$. For $\mathcal{L}^1 (Q^i) \times L^1 (Q^i) \subset \mathcal{L}^2 (P) \times L^2 (P)$ it is sufficient that the likelihood ratio process of $Q^i$ with respect to $P$ is in $\mathcal{L}^2 (P)$. This is guaranteed by the Assumption (ApprArb) at the end of this section so that completeness holds indeed for $\mathcal{L}^2 (P) \times L^2 (P)$. An assumption ensuring this property is introduced in this subsection.

**Remark 7** *The requirement '$P - a.s.$' can always be replaced in the sequel by the stronger, sufficient condition '$\forall y \in Range(H)$', since outside this range the processes stay with probability zero. Nevertheless, the latter might be easier to verify for the exogenous functionals, because it does not require a picture of $P$, i.e. in which region the processes actually live with probability one.*

**Proof. of Lemma 85:** If the rank of the diffusion matrix were smaller than $m$ not all the risk of the Brownian motions could be spanned by the assets. Without a spanned bank account riskless income streams are unavailable. For details refer to Theorem 6.6 in (Karatzas & Shreve, 1998), p. 24. ■

**Proof. Proof of Corollary 86:** The diffusion matrix of $\hat{H}_t^i$ is

$$\hat{\sigma}^i = \begin{pmatrix} \mathbf{0}^T \\ \sigma_t^i \\ X^i/b^{\neg i}\sigma_t^X \\ X^i\left(\sigma_t^{\neg i} + H_t^{\neg i}\sigma_t^X\right) \end{pmatrix}.$$

The matrix in the two bottom rows is the diffusion of country $\neg i$'s assets exchanged to currency $i$, which comes from the stochastic differential equation of the product $X^i\left(1/b^{\neg i}, \left(H^{\neg i}\right)^T\right)^T$. Without the first row this matrix transforms by simple matrix operations into the matrix given in the corollary. Since the locally riskless bank account of currency $i$ is in $\hat{H}_t^i$ (first row), complete markets are assured by Lemma 85. ■

An arbitrage opportunity is a trading strategy $\theta$ such that it is meantime self–financing between some $s, t \in \mathbb{T}, s < t$, and there is an $P$–essential event $\mathbb{A} \in \mathcal{F}_s$ such that:

$$\begin{aligned} & v_s^{i\theta} \leq 0 \text{ on } \mathbb{A} \quad \text{and} \quad v_t^{i\theta} \geq 0, \quad P_s|_{\mathbb{A}} - a.s. \\ & \text{as well as } \; P_s|_{\mathbb{A}}\left(v_t^{i\theta} > 0\right) > 0 \quad \text{holds.} \end{aligned}$$

Markets are free of arbitrage if there is no arbitrage generated by an admissable portfolio strategy.[101] For an arbitrary but admissable and self–financing portfolio strategy $\theta$ let $\mu^{i\theta}$ and $\sigma^{i\theta}$ be the trend and the diffusion of the value process $v^{i\theta}$ in currency $i$. The prices of risk in currency $i$, denoted by $-\xi^i$, are the processes which solve for all admissable $\theta$ the equation

$$\mu_t^{i\theta} = r_t^i v_t^{i\theta} - \sigma_t^{i\theta} \xi_t^i \quad P \times \lambda - a.s.$$

The prices of risk[102] determine by Girsanov's Theorem [see for instance Ch. 3.5. in (Karatzas & Shreve, 1988)] the Radon–Nikodym derivative $z_T^i$ of the martingale measure $Q^i$ with respect to the physical measure $P$, provided that the martingale measure exists:

$$\left.\frac{dQ^i}{dP}\right|_{\mathcal{F}_t} = z_t^i, \text{ where } z_t^i = \mathcal{E}_t\left(\xi^i \bullet W\right),$$

where $\mathcal{E}_t$ stands for the stochastic exponential defined by

$$\mathcal{E}_t\left(\xi^i \bullet W\right) = \exp\left\{\int_0^t \left(\xi_u^i\right)^T dW_u - \frac{1}{2}\int_0^t \left\|\xi_u^i\right\|^2 du\right\}.$$

$(z_t^i)_{t\in\mathbb{T}}$ is also called the likelihood ratio process or – if in $\mathcal{L}^2(P \times \lambda)$ – the pricing asset. Established in the next paragraph is the connection between arbitrage and the existence of prices of risk, the martingale measure and a pricing asset. For this reason an assumption is developed at the beginning of that paragraph which guarantees that the likelihood ratio process is indeed a pricing asset. This simplifies the subsequent analysis.

Let for all $t \in \mathbb{T}$ the three matrix-processes $V^i(t, y_t) \in \mathbb{R}^{\hat{n}^d+\hat{n}^f+2} \otimes$

[101] Attainability of portfolio processes rules out some pathological arbitrage strategies known as Ponzi schemes. These doubling strategies could also be ruled out by $Q-$square integrable value processes [(Duffie, 1996), Ch. 6D.], but this involves the martingale measure before its existence is proven.
Since the riskless bank account is traded, the definition of arbitrage covers also the case of a strictly negative investment with non-negative payoffs.

[102] The minus sign of the prices of risk is usually suppressed in the text.

$\mathbb{R}^{n^d+n^f+2}, \Sigma^i(t,y_t) \in \mathbb{R}^{\hat{n}^d+\hat{n}^f+2} \otimes \mathbb{R}^m$ and $Y^i(t,y_t) \in \mathbb{R}^m \otimes \mathbb{R}^m$ build a singular value decomposition of $\hat{\sigma}^i(t,y_t) = V_t^i \Sigma_t^i (Y_t^i)^T$, where $\Sigma^i$ contains the singular values $\phi_j^i,\ j = 1\dots m$ (see Appendix (3.9) for the definition of a singular value decomposition). Since $V_t$ spans the entire $\mathbb{R}^{\hat{n}^d+\hat{n}^f+2}$, there is a unique $\zeta^i(t,y_t) \in \mathbb{R}^{\hat{n}^d+\hat{n}^f+2}$, which solves

$$\hat{\mu}^i(t,y_t) - r_t^i \hat{H}_t^i = V(t,y_t)\,\zeta^i(t,y_t)$$

for the excess return $\hat{\mu}^i(t,y_t) - r_t^i \hat{H}_t^i$ of assets $\hat{H}^i$ denominated in currency $i$.

**Assumption (ApprArb):** Say that *'approximate arbitrage is ruled out'* if $\exists K > 0$ such that $\forall i \in I, t \in \mathbb{T}, j = 1\dots m, y \in Range(H)$ :

$$\left|\zeta_j^i(t,y)\right| \leq K \left|\phi_j^i(t,y)\right|.$$

The following lemma summarizes some important results the subsequent analysis relies on.

**Lemma 87** *(a) In complete markets at most one price of risk and one martingale measure exist.*
*(b) Markets are free of arbitrage if and only if there is a price of risk.*
*(c) Suppose that 'approximate arbitrage is ruled out'. Then markets are free of arbitrage opportunities if and only if a martingale measure exists with a pricing asset in* $\mathcal{L}^2(P \times \lambda)$.
*(d) Domestic markets are free of arbitrage if and only if foreign markets are.*

The proof of parts (a) and (b) just spells out the argumentation given in (Duffie, 1996), Ch. 6, whereas the proof for (c) only applies a mathematical lemma to Assumption (ApprArb). In this proof the Moore–Penrose inverse appears for the first time, see Appendix 3.9 for the definition. The Moore-Penrose inverse of a matrix $A$ will be denoted by $A^+$.

**Proof.** (a) Consider the trading/consumption strategies which consist in buying a single asset in 0 and consuming its pay–offs in $\mathcal{T}$. Then the vector equation $\hat{\mu}^i = r^i \hat{H}^i - \hat{\sigma}^i \xi^i$ has the solutions $\xi^i = -\left(\hat{\sigma}^i\right)^+ \left(\hat{\mu}^i - r^i \hat{H}^i\right) + u$ with $u \in \mathcal{N}\left(\left(\hat{\sigma}^i\right)^T\right)$. Then $\mathcal{N}\left(\left(\hat{\sigma}^i\right)^T\right) = \{\mathbf{0}\}$ holds if $\hat{\sigma}^i$ has full column–rank $m$, yielding a unique solution. Provided that a martingale measure exist there is a one to one correspondence between the martingale measure and the price of risk by Girsanov's theorem.

(b) Suppose a solution for a price of risk does not exist on an essential set in $\mathcal{F} \otimes \mathbb{B}(\mathbb{T})$. Thus $\hat{\mu}^i - r^i \hat{H}^i \notin \mathcal{R}\left(\hat{\sigma}^i\right)$ and because of $rank\left(\hat{\sigma}^i\right) \leq m \leq 1 + \hat{n}^d + \hat{n}^f$ there is an admissable, self–financing trading strategy $\theta$ for which $\theta^T\left(\hat{\mu}^i - r^i \hat{H}^i\right) > 0 = \theta^T \hat{\sigma}^i$ on this particular set. The value process increases more than the locally riskless bank account and is itself locally riskless, which constitutes an arbitrage opportunity.

To the converse, if there is a price of risk any portfolio with positive excess return $\theta^T\left(\hat{\mu}^i - r^i \hat{H}^i\right) > 0$ cannot be locally riskless $\theta^T \hat{\sigma}^i \neq 0$. But any portfolio strategy which is not locally riskless on an essential set in $\mathcal{F} \otimes \mathbb{B}(\mathbb{T})$ cannot offer free lunch, since the local martingale part in the value process $v^{i\theta}$ allows for negative pay–offs with positive probability. Because an arbitrage strategy starts with a non–positive investment, a zero excess return accompanied by zero risk cannot result in positive pay–offs, if the strategy is self–financing. Thus, arbitrage opportunities are ruled out.

(c) Part (b) establishes the equivalence between the no–arbitrage condition and the existence of a price of risk. A price of risk induces a martingale measure, if the premises of Girsanov's Theorem are satisfied. Those are (i) square–integrability of the price of risk and (ii) the martingale property of the likelihood ratio process. Both are implied by Novikov's condition, which is square–integrability of the likelihood ratio process itself. Thus, the existence of a pricing asset is equivalent to Novikov's condition, which has to be verified now. Because markets are free of arbitrage $\Sigma_t^i \left(Y_t^i\right)^T \xi_t^i = \zeta_t^i$ holds $P \times \lambda - a.s.$ Thus $\xi = Y \Sigma^+ \zeta + u$

and $\|\xi\|^2 = \sum_{j=1}^{rank(\hat{\sigma})} \zeta_j^2 / \phi_j^2 + u^T u$, where $u = Y\left(I - \Sigma^+ \Sigma\right)\theta$ for some arbitrary $\theta \in \mathbb{R}^m$. For showing existence $u$ can be assumed to be zero to minimize the norm. Then the norm of $\xi$ is bounded by Assumption (ApprArb). For bounded processes $\xi$ (Chung & Williams, 1990) have shown, that the stochastic exponential $\mathcal{E}_t\left(\xi \bullet W\right)$ satisfies Novikov's condition [see their example following Theorem 6.5. on p. 126 ibidem]. (d) Since the exchange rate is a strictly positive scalar it does neither generate nor eliminate arbitrage opportunities. ■

**Remark 8** *The relevant cases for applying Assumption (ApprArb) occur, when $\hat{\sigma}$ changes its rank. Note, the norm of the excess returns is*

$$\left\|\hat{\mu}^i - r^i \hat{H}^i\right\|^2 = \sum_{j=1}^{rank(\hat{\sigma})} \left(\zeta_j^i\right)^2.$$

*The Assumption (ApprArb) states, whenever the diffusion matrix $\hat{\sigma}$ is about to change its rank, i.e. if $\phi_j$ is zero on a set $\mathbb{A}$ in $\mathcal{F} \otimes \mathbb{B}(\mathbb{T})$ and small in the neighbourhood of this set, the length $\left|\zeta_j\right|$ of the orthogonal projection of excess returns onto the corresponding eigenvector in $V_j$ must be zero on $\mathbb{A}$ (by arbitrage, if $\mathbb{A}$ is essential) and small in its neighbourhood (by no approximate arbitrage). If the latter condition does not hold, the portfolio $V_j / \left|\zeta_j\right|$ yields a constant return although its volatility $\left(\phi_j\right)^2 \big/ \left(\zeta_j\right)^2$ converges to zero in the neighbourhood of $\mathbb{A}$ the tighter $(\omega, t) \notin \mathbb{A}$ is to $\mathbb{A}$.*
*This is not an rigorous argument for establishing an approximate arbitrage strategy, because this concept has not been defined yet and it is intended to do so only verbally: An approximate arbitrage opportunity is a sequence of admissible, self–financing portfolio strategies whose investment costs are never larger and whose final pay–off is never smaller than a corresponding pay–off sequence which itself converges to an unfeasible free lunch 'opportunity'. Clark shows [in (Clark, 1993)] the equivalence between the existence of a pricing asset and non–existence of approximate arbitrage. It concerns approximate arbitrage opportuni-*

*ties* inside *the marketed subspace [Assumption A.2 in conjunction with Theorem 2, ibidem], which yields a continuous linear functional and thus allows for applying a variant of Riesz' representation,* as well as *approximate arbitrage considering the* entire *consumption space [Assumption A.5 in conjunction with Theorem 7, ibidem], which then allows for a continuous extension of the pricing functional to a positive pricing asset on the whole* $\mathcal{L}^2 (P \times \lambda)$. *Note that Assumption (ApprArb) might be violated even in complete markets, i.e. violated in a neighbourhood of a* $P \times \lambda$*–zero set on which markets are incomplete. From this point of view here only a proof is given that Assumption (ApprArb) is sufficient to the existence of a pricing asset and therefore to rule out approximate arbitrage* inside *the marketed subspace. However, the example at the beginning of this remark gives rise to the supposition that a similar condition would also be necessary. Nevertheless, Assumption (ApprArb) is a sufficient condition on the exogenous drift and diffusion functionals, and therefore more useful than Novikov's indirect condition.*

The last issue of attention is square–integrability of involved processes. If trend and diffusion of an Itô–process of diffusion type satisfy the linear growth conditions this process is not only square–integrable but all higher moments do exist either [see (Karatzas & Shreve, 1988) Problem 3.15 on p. 306 and its solution on p. 389]. The reason for assuming the exchange rate to be bounded from infinity and away from zero was that square–integrability is preserved if domestic assets are exchanged to the foreign currency and vice versa.

Since in the next subsection replacements are found for $\mu^X$ and $\sigma^X$ by functions of $\mu^i, \sigma^i$ the constraints imposed on $\mu^X$ and $\sigma^X$ go over to further restrictions on $\mu^i, \sigma^i$. It is not analysed whether those restrictions are superfluous or are even stronger than those already assumed for $\mu^i, \sigma^i$. It suffices to know that there exist processes satisfying those conditions jointly.[103]

[103]For instance geometric Brownian motions with constant coefficients for all assets and bounded interest rates.

### 3.3.2 Pricing the FX–rate

This section is subdivided into five parts. The first part establishes the fx–rate as a change between two numéraires. In the second subsection the trend and the diffusion of the fx–rate are determined by hedging the foreign bank account and other foreign assets respectively. After that the UIP is shown to be non–generic in the set of possible prices of risk. Then the influence of price inflation on the fx–rate is analysed. In the last part the reduced form equations, which are applied in the empirical Section 3.5, are derived from previous results.

#### Change of Numéraires

With a minimum of $m+1$ assets complete markets could be achieved. In this case prices of risk are uniquely specified in both currencies if arbitrage opportunities are ruled out. By changing the numéraires via the fx–rate also the martingale measures change from $Q^i$ to $Q^{\neg i}$. This implies that the short rates and the prices of risk completely determine the returns of the fx–rate under the martingale measures. The exchange rate inter–links both prices of risk in the way the following lemma states.

**Lemma 88 (SDE of the exchange rate)** *If in both countries a locally riskless savings account* $(b^i)^{-1}$ *is spanned by existing assets, the SDE of the exchange rate under the martingale measure* $Q^d$ *is:*

$$d \ln X \overset{Q^d}{=} \left(r^d - r^f\right) dt + \sigma^X d\tilde{W}^d.$$

*Suppose that the economy shows to have arbitrary (possibly equilibrium) prices of risk* $\xi^d$ *and* $\xi^f$*. Complete markets are free of arbitrage if and only if* $\sigma^X = \xi^f - \xi^d$*, i.e. the exchange rate solves the SDE*

$$dX = X\left(r^d - r^f - \left(\xi^f - \xi^d\right)^T \xi^d\right) dt + X\left(\xi^f - \xi^d\right)^T dW.$$

**Proof.** The trend $\mu_k^i$ of any asset $H_k^i$ is determined by its price of risk $\sigma_k^i \xi^i$ and the short rate $r^i$. Exchanging an asset to currency $\neg i$ involves the trend and the diffusion of the exchange rate. This constitutes an equation with an unique solution for $\left(\mu^X, \sigma^X\right)$ given $r^i$ and $\xi^i$. For details see Appendix 3.7.1. ■

This result was introduced in the literature mainly with the purpose of pricing fx–derivatives and can be found for instance in (Amin & Jarrow, 1991), (Flesaker & Hughston, 1996), and in [(Musiela & Rutkowski, 1997), Chapters 7 and 17]. With regard to the foreign exchange rate premium, which is $\left(\xi^f - \xi^d\right)^T \xi^d$, this result was derived in (Saá–Requejo, 1994), see also Theorem 4.11 and Corollary 4.17 in (Trojani, 1999).

If risk neutrality predominates with respect to both numéraires the real exchange rate is locally deterministic since both prices of risk have to be zero. This is not the case for the nominal exchange rate (see below). To be globally deterministic, the interest rate differential must be non–stochastic, which is an equilibrium condition for risk neutral investors and cannot be shown in an arbitrage pricing model. [104]

**Hedging the FX–rate**

Trend and diffusion of the exchange rate are to be determined. The idea is straight forward. Trend and diffusion of the exchange rate offer $m+1$ degrees of freedom. With any asset over the $m+1$ assets needed to complete markets one could solve for one variable out of $\left(\mu^X, \sigma^X\right)$ depending on the remaining unknowns. It is shown that with at least $2m + 2$ asset price processes the exchange rate is uniquely determined by existing assets. To follow this idea it is necessary to assume complete asset markets.

**Assumption (Complete Markets)** A locally riskless bank account is spanned for the domestic currency. Moreover, $n^d$ domestic and $n^f$ foreign assets exist such that their diffusion matrices and the diffusion

[104] However (Heath *et al.*, 1992) model stochastic forward rates unter the risk neutral *probability measure.*

vector of the exchange rate guarantee that the matrix:

$$\begin{pmatrix} \sigma_t^d \\ \sigma_t^f + H_t^f \sigma_t^X \end{pmatrix} \text{ is regular } P \times \lambda - a.s.$$

Because the prices of foreign assets have to be exchanged to the domestic currency first, the volatility of the foreign assets is adjusted by adding up the diffusion of the exchange rate times the foreign asset prices, which comes from Itô's rule.

In complete markets one is now in the position to price any asset via the linear pricing rule. But the exchange rate is not an asset by itself. To solve for the unknowns $\left(\mu^X, \sigma^X\right)$ one has to price an asset which involves the exchange rate, thus, a foreign asset. A natural choice is the foreign locally riskless bank account, which has in the domestic currency the same diffusion vector as the exchange rate. For this idea to go through one further assumption is needed.

**Assumption (FX–Spanning)** The diffusion of the exchange rate is spanned by the assets which define a complete domestic asset market:

$$\left(\sigma_t^X\right)^T \in \mathcal{R}\left(\begin{pmatrix} \sigma_t^d \\ \sigma_t^f \end{pmatrix}^T\right) \text{ holds } P \times \lambda - a.s.$$

Intuitively speaking, the exchange of the foreign assets into the domestic currency introduce the risk incorporated by the exchange rate, i.e. by adding up the volatility sub–matrix $H^f \sigma^X$. But if the risk of the exchange rate is not already spanned prior to the transformation of the foreign assets into the domestic currency this would lead to a degeneration in the valuation of the foreign bond. Because then the foreign bond can only be spanned by itself, which would not help in solving for the unknown components of $\left(\mu^X, \sigma^X\right)$.

**Proposition 89 (FX–Flow)** *Suppose that domestic financial markets are complete and that FX–spanning holds. Then asset prices are free of arbitrage if and only if the following relation between the trend and the diffusion vector*

$\left(\mu^X, \sigma^X\right)$ *of the exchange rate is satisfied* $P \times \lambda - a.s.$*:*

$$\mu^X = r^f - r^d + \beta \begin{pmatrix} \mu^d - r^d H^d \\ \mu^f - r^f H^f + \sigma^f \left(\sigma^X\right)^T \end{pmatrix}$$

$$\text{with } \beta = \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1}.$$

**Proof.** The mathematics of the proof is given in the appendix. Its general idea is already explained by the introduction of the previous two assumptions in an intuitive way: The $n^d$ domestic and $n^f$ foreign risky assets plus the domestic bond form a complete asset structure. Then the foreign bond measured in domestic currency is a redundant asset and can be valued by the linear pricing rule. It is then possible to derive two expressions for the price of the foreign bond. Setting those equal, the equation involves the unknown components $\left(\mu^X, \sigma^X\right)$, which can be solved for $\mu^X$. By arbitrage the result follows. See Appendix 3.7.2 for details. ■

**Remark 9 (Spanning)** *In the literature sometimes non–redundant derivatives are priced in incomplete markets. Then the martingale measure will not be unique. With more Brownian motions than assets the inverse in* $\beta$ *could be any right–inverse. One has to assume a certain measure to get a concrete result, which corresponds in this model to a concrete inverse, e.g. the Moore–Penrose inverse in case of the minimal risk approach. In many finance articles this is the minimal martingale measure introduced by (Föllmer & Schweizer, 1991), but also extreme measures are suggested, e.g. in (El–Karoui & Quenez, 1995). It is preferred here to avoid this approach, because of its indeterminacy, which could just be solved in a much more elaborated equilibrium model. In the next chapter such an equilibrium model is introduced, but with complete financial markets. Nevertheless emphasized is that the above result is in line with the orthogonal projection approach, because one can choose a basis of spanning Brownian motions such that* $X$ *uses just additionally some Brownian motions orthogonal to those used in* $H^d$ *and* $H^f X$*. These additional Brownian motions are not affected by the change of measure via the minimal risk approach.*

The diffusion of the exchange rate is necessary to compute the prices of risk belonging to one numéraire as long as assets denominated in the other numéraire are needed to complete asset markets. In other words, the less assets have to be exchanged to achieve complete asset markets, the 'lower' is the influence of $\sigma^X$ on $\xi^i$ and the more are the degrees of freedom in $\sigma^X$ restricted by Lemma 88. The next proposition exploits the idea of introducing more redundant securities, which are not redundant considering only securities denominated in the same numéraire, for instance a new domestic investment fond containing foreign assets. This leads to a more accurate determination of the diffusion of the exchange rate, which can be seen as a generalization of the well known purchasing power parity (PPP).

**Proposition 90 (Generalized PPP)** *Suppose there are two ways to achieve complete domestic asset markets. Either by $n^d + \bar{n}$ domestic and $n^f$ foreign assets or by $n^d$ domestic and $n^f + \bar{n}$ foreign assets. The two possibilities differ only in the substitution of $\bar{n}$ foreign assets by the same number of domestic assets, whereas all other assets remain the same. The changing assets are denoted by a bar on top. For both completions FX–spanning shall hold.*

*Under these assumptions markets are free of arbitrage opportunities if and only if the diffusion of the exchange rate satisfies the following equation up to $m - \bar{n}$ degrees of freedom:*

$$\begin{aligned}
\left(\sigma^X\right)^T &= \begin{pmatrix} \bar{\sigma}^d \\ \sigma^d \end{pmatrix}^+ \begin{pmatrix} \bar{\mu}^d - \bar{H}^d r^d \\ \mu^d - H^d r^d \end{pmatrix} + y^d \\
&\quad - \begin{pmatrix} \bar{\sigma}^f \\ \sigma^f \end{pmatrix}^+ \begin{pmatrix} \bar{\mu}^f - \bar{H}^f r^f \\ \mu^f - H^f r^f \end{pmatrix} + y^f \quad P \times \lambda - a.s. \\
&\quad \textit{for some arbitrary } y^i \in \mathcal{N} \begin{pmatrix} \bar{\sigma}^i \\ \sigma^i \end{pmatrix},
\end{aligned}$$

*where the superscript + stands for the Moore–Penrose inverse.*

**Proof.** See Appendix 3.7.3. ■

One can call this result a generalization of the purchasing power parity because there are some consumption processes independently priced in both currencies which are actually the same if exchanged to a particular currency. This restricts the arbitrage–free choice for $\sigma^X$. Note that if both markets are independently complete, i.e. $n^d = n^f = 0$ and $\bar{n} = m$, the diffusion is uniquely determined by the difference of the prices of risk:

$$\begin{aligned}\left(\sigma^X\right)^T &= \left(\bar{\sigma}^d\right)^{-1}\left(\bar{\mu}^d - \bar{H}^d r^d\right) - \left(\bar{\sigma}^f\right)^{-1}\left(\bar{\mu}^f - \bar{H}^f r^f\right) \\ &= \xi^f - \xi^d.\end{aligned}$$

This result concurs with Lemma 88, where unique prices of risk have been assumed in advance. Here those are derived from asset prices. This again emphasizes that the exchange rate is not an asset by itself, since the diffusion of a redundant asset is arbitrary if its return is appropriately determined by arbitrage pricing.

**The Forward Rate as a Biased Predictor**

The UIP is based on the differences of the log–exchange rate. The following corollary establishes the SDE of the log–fx–rate.

**Corollary 91 (Log–Exchange Rate)** *Suppose the economy admits two arbitrary (possibly equilibrium) prices of risk $\xi^d$ and $\xi^f$. Complete markets are free of arbitrage if and only if the logarithmic exchange rate follows the SDE*

$$d\ln X = \left(r^d - r^f + \frac{1}{2}\left(\left\|\xi^d\right\|^2 - \left\|\xi^f\right\|^2\right)\right)dt + \left(\xi^f - \xi^d\right)dW. \qquad \text{(LogFx)}$$

**Proof.** The first part follows by Itô's rule

$$d\ln X = \frac{dX}{X} - \frac{d\langle X, X\rangle}{2X^2}.$$

The trend of $dX/X$, $r^d - r^f - \xi^d\left(\xi^f - \xi^d\right)$ as derived in Lemma 88, is adjusted by $-1/2\left\|\xi^f - \xi^d\right\|^2$. This results in the trend given in the corollary. ■

Now the question is considered whether the uncovered interest parity could be satisfied. The next definition calls the UIP in mind. Let $\triangle_{t,\tau} \ln X := \ln X_\tau - \ln X_t$.

**Definition 92** *The exchange rate satisfies the UIP if* $\forall t, \tau \in \mathbb{T}, t \leq \tau$ :

$$E_P^t [\triangle_{t,\tau} \ln X] = \ln B_{t,\tau}^f - \ln B_{t,\tau}^d.$$

The claim here is that any economy is specified by some prices of risk $\xi^d$ and $\xi^f$. These prices could be arbitrary processes in $\mathcal{L}^2 (P \times \lambda)$ while satisfying Novikov's condition. Then the UIP should hold for arbitrary prices of risk. Otherwise it could not be valid for all economies. This kind of arbitrariness of the UIP would not be a very demanding property, because one cannot be assured whether the UIP holds for the economy in concern. 'Unfortunately' this is the result the next proposition states, which proves Conjectures 79 and 83 made in Section 3.2.

**Proposition 93 (Non–Viability of the UIP)** *Suppose the short rates have continuous, bounded sample paths. Assume moreover that the two economies are specified by arbitrary prices of risk* $-\xi^i : \mathbb{T} \times \Omega \to \mathbb{R}^m$ *in* $\mathcal{L}^2 (P \times \lambda)^m$ *implying the existence of suitable pricing assets* $z^i$ *and martingale measures* $Q^i$. *Then the UIP holds only for economies of Lebesgue–measure zero in the range of* $\xi^i$. *By assuming the UIP the dimension of the range shrinks by one.*

**Proof.** This result is rather intuitive, since the UIP states the equality of

$$\begin{aligned} E_P^t [\Delta_{t,\tau} \ln X] &= E_P^t \left[\ln b_{t,\tau}^d\right] - E_P^t \left[\ln b_{t,\tau}^f\right] + \\ &\quad \frac{1}{2} E_P^t \left[\int_t^\tau \left\|\xi_u^d\right\|^2 - \left\|\xi_u^f\right\|^2 du\right] \\ &= \ln E_{Q^d}^t \left[b_{t,\tau}^d\right] - \ln E_{Q^f}^t \left[b_{t,\tau}^f\right] = \ln F_{t,\tau} - \ln X_t, \end{aligned}$$

by Corollary 91. From the first to the bottom line $E_P^t (\ln (\cdot))$ changes to $\ln E_{Q^i}^t (\cdot)$ with respect to the discount factors $b_{t,\tau}^i$. In other words, the UIP in the bottom line suggests to take the logarithms of the expected values with respect to the martingale measures, whereas the arbitrage pricing in the first line favours the expected values with respect to the physical measure of logarithms plus a correction term. This difference can only hold by a suitable correction via the prices of risk, which is restrictive. In Appendix 3.7.4 it is shown that this equation is typically not satisfied, and it is in fact an one dimensional restriction imposed on the prices of risk. ■

(Trojani, 1999) derived the stochastic differential equations for the interest rates and the exchange rate such that the UIP holds exactly at every point in time. Thus, even though prices of risk satisfying the UIP are non–generic they can exist.

The forward price is in general a biased predictor of the underlying, not only in fx–markets. From experience this is a popular misinterpretation of forward prices in practice. The intention is not to rigorously prove this here, but to provide an intuitive argument.

Let the state price density be defined by $z_t^i = \frac{dQ^i}{dP}\Big|_{\mathcal{F}_t}$. The arbitrage–free forward price $F_{t,T}^y$ in $t$ of a traded underlying $y$ with settlement in $T$ is:

$$F_{t,T}^y = \frac{y_t}{B_{t,T}}.$$

The hypothesis $E_P^t(y_T) = F_{t,T}^y$ is equivalent to

$$\begin{aligned} B_{t,T} E_P^t(y_T) &= B_{t,T} F_{t,T}^y = y_t \Leftrightarrow \\ E_P^t(b_{t,T} z_T) E_P^t(y_T) &= E_P^t(y_T b_{t,T} z_T). \end{aligned}$$

This implies that the forward expectation hypothesis is only true if the underlying asset is uncorrolated with the state price functional $b_{0,T} z_T$ (see the last equation). Moreover, $y_t$ has to be a submartingale, which instantaneously grows at an expected rate equal to the instantaneous forward rate $f_{t,T,T} = -\frac{\partial}{\partial T} B_{t,T}$. Summarizing, the FEH may apply only to very special assets.

There are forward agreements on underlyings, which are not traded on the spot market, e.g. forward interest rate contracts. For those the forward expectation hypothesis sounds $f_{t,T_1,T_2} = E_P^t(r_{T_1,T_2})$, where on the right hand side stands the forward rate in $t$ for the future period $T_1$ to $T_2$ and on the left hand side is the expected value of the compound yield for this period. Those are defined by

$$f_{t,T_1,T_2} = \frac{1}{T_2 - T_1} \ln \frac{B_{t,T_2}}{B_{t,T_1}} \quad \text{and} \quad r_{T_1,T_2} = \frac{1}{T_2 - T_1} \ln B_{T_1,T_2}.$$

The hypothesis implies:

$$\ln \frac{E_Q^t (b_{t,T_2})}{E_Q^t (b_{t,T_1})} = E_P^t \left( \ln E_Q^{T_1} (b_{T_1,T_2}) \right),$$

which is generally true only if the short rate is deterministic.

### Risk of Price Inflation

The no–arbitrage conditions are in real and nominal terms the same. Nevertheless the short rate process and the prices of risk definitely change form the real to the nominal point of view. Moreover in real financial markets prices are observed in nominal quantities. To make this step money is introduced in a fairly abstract sense. Assumed is the neutrality of money even in the short run, i.e. money just normalizes prices. This will be done by the following two stochastic processes of price indices:

$$dq_t^i = \mu_t^{qi} q_t^i dt + \sigma_t^{qi} q_t^i dW_t, \quad q_0^i > 0.$$

**Lemma 94** *The instantaneous nominal interest rates $\hat{r}^i$, nominal prices of risk $\hat{\xi}^i$ and the nominal exchange rate $\hat{X}$ are given by:*

$$\begin{aligned} \hat{r}_t^i &= r_t^i + \mu_t^{qi} + \sigma_t^{qi} \left( \hat{\xi}_t^i \right)^T, \\ \hat{\xi}_t^i &= \xi_t^i - \sigma_t^{qi}, \\ \hat{X}_t &= X_t \frac{q_t^d}{q_t^f} = X_0 \frac{q_t^d}{q_t^f} \frac{b_{0,t}^f z_t^f}{b_{0,t}^d z_t^d} = X_0 \frac{\hat{b}_{0,t}^f \hat{z}_t^f}{\hat{b}_{0,t}^d \hat{z}_t^d}. \end{aligned}$$

**Proof.** See Appendix 3.7.5. ■

Note that the nominal short rate is not only the sum of the real short rate and expected inflation, but incorporates also an additional risk bonus or premium depending on whether the diffusion of the price index points into the direction or the opposite direction of the real price of risk, respectively. The nominal price of risk is the real price of risk minus the diffusion of the price index. The nominal risk premium is sensitive to insurance against the risk of inflation. As mentioned in the previous section, risk neutrality must be considered in real terms. In this case the real price of risk is zero in the corresponding numéraire in which investors are risk neutral. But the nominal price of risk is still equal to the diffusion of the price index.

**Reduced Form Equations**

The last claim in this section is a preparation of the empirical analysis made in the next section but one. An estimation turns out to be much more cumbersome and less efficient if the equation could not be solved for the endogenous variable on the one hand and for the exogenous variables on the other hand. For this reason stated in the next corollary is an orthogonal projection of the differential of the log–exchange rate onto the differentials of assets. The empirical section is limited to geometric Brownian assets. This considerably simplifies the reduced form equation in discrete time and also the estimation thereby. This simplification is undertaken in the following corollary.

**Corollary 95 (Reduced Form Equation)** *Suppose that domestic financial markets are complete and that FX–spanning holds. Let* $\beta = \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1}$.

*(1) Asset prices are free of arbitrage if and only if the log–exchange rate is the solution to the SDE*

$$d\ln X = \left( r^d - r^f + \beta \begin{pmatrix} \mu^d - r^d H^d - \frac{1}{2}\sigma^d \left(\sigma^X\right)^T \\ \mu^f - r^f H^f + \frac{1}{2}\sigma^f \left(\sigma^X\right)^T \end{pmatrix} \right) dt + \sigma^X dW,$$

*which implies the orthogonal projection on asset prices*

$$d\ln X = \left(r^d - r^f\right) dt + \beta \left[ d \begin{pmatrix} H^d \\ H^f \end{pmatrix} - \begin{pmatrix} r^d H^d - \frac{1}{2}\sigma^d \left(\sigma^X\right)^T \\ r^f H^f + \frac{1}{2}\sigma^f \left(\sigma^X\right)^T \end{pmatrix} dt \right].$$

*(2) Suppose additionally that assets are geometric Brownian:*

$$dH_k^i(t) = \mu_k^i(t) H_k^i(t) dt + \sigma_k^i(t) H_k^i(t) dW_t, \; k = 1, \ldots, \hat{n}^i, \; i \in I.$$

*Then by arbitrage the log–exchange rate can be projected in the following way:*

$$\begin{aligned} d\ln X &= \left(r^d - r^f\right) dt + \beta \left[ d\ln \begin{pmatrix} H^d \\ H^f \end{pmatrix} - \begin{pmatrix} r^d \mathbf{1} \\ r^f \mathbf{1} \end{pmatrix} dt + \right. \\ & \left. \begin{pmatrix} -\sigma^d \\ \sigma^f \end{pmatrix} \sigma^X dt + \frac{1}{2} \left( \left\|\sigma_1^d\right\|^2, \ldots, \left\|\sigma_{n^f}^f\right\|^2 \right)^T dt \right]. \end{aligned}$$

**Proof.** The results follow from a suitable transformation and Itô's rule. See Appendix 3.7.6 for more details. ■

The second SDE has an intuitive interpretation. First, one observes a *kind* of the uncovered interest parity, $r^d - r^f$. Note that

$$\begin{aligned} \beta &= \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^T \left[ \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^T \right]^{-1} \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \\ &= \langle \ln X, \left(H^d, H^f\right) \rangle \langle \left(H^d, H^f\right), \left(H^d, H^f\right) \rangle^{-1} \end{aligned}$$

is the vector of coefficients which results from a linear orthogonal projection from $\ln X$ on $\left(H^d, H^f\right)$. Hence the projection on asset returns above the locally riskless return corrected by a diffusion term, i.e. $dH^i - r^i H^i \mp \frac{1}{2}\sigma^i \left(\sigma^X\right)^T$, explain the second part in $d\ln X$. Here the excess returns $dH^i - r^i H^i$ are adjusted by an expression which looks like a risk premium. The 'risk premium' is equal to the covariance of assets with the fx–rate, $\sigma^i \left(\sigma^X\right)^T$. The projection suggests some rough estimation about the influence of asset returns on the exchange rate differential. The influence of domestic/foreign assets is the most in the negative/positive direction, if they have low/high covariance, low

variance and high excess return.
Consider now the question how fundamental values of an economy determine the interest rates and prices of risk, which themselves specify the exchange rate. This will help to formulate an empirical hypothesis about the projection given in the last corollary.

## 3.4 Equilibrium Foundation of FX–Rates

At the beginning the question stands which functionality the exchange rate serves in or between economies. Imagine an economy which is split up into two parts. In one part a new currency is introduced. In this sub–economy prices are normalized now with regard to this new numéraire. The exchange rate serves as the price ratio between the old and the new numéraire. Since the real economy stays the same, it is not expected that any real quantity changes, provided that markets operate frictionless.[105] International markets became more integrated, so that this perfectionist view in this section might be justified as a simplifying assumption. The methodological shortcoming of this view is that it can neither explain the existence of two particular numéraires nor the selection of a particular basket nor the choice of money as the numéraire.[106]
This section has three matters of concern. First, the arbitrage pricing model is enriched with real entities of the economy. It will help to understand which fundamentals influence the exchange rate in which direction. This cannot be derived from the pure arbitrage pricing view as dealt with above. In the

[105] If there are frictions in the market the change of numéraire could have indeed real impacts. A highly stylized example is a monopolist, who is maximizing his profits by choosing the optimal amount of a single output. Suppose he considers the influence of his production decision on the price of the produced good, but neglects it on the factor prices. This is a typical bounded rational behaviour. Then making his output good the numéraire would eliminate his influence on the price, which leads to a competitive outcome. For a more rigorous analysis of numéraire changes and their impact on the exchange rate as well as on the real allocation see (Hens *et al.*, 1999).
[106] Certain numéraires seem to be chosen for reducing transaction costs (beside of historical reasons). It is questionable that money generates utility directly, like it was assumed in (Basak & Gallmeyer, 1999). It is 'just' a technology which helps to allocate goods more efficiently.

next section the projection of the foreign exchange rate changes is estimated, which was derived in Corollary 95 at the end of the last section. For this reason one has to identify a set of assets which best reflect the uncertainty of the economy. Having done this, formulated is a hypothesis of how the coefficients of the representation, $\beta$, might look like.
This section focuses on an example economy which is both special but also extensive. In this sense the results should be seriously interpreted in a qualitative and not in a quantitative way. The advantage of this example will be an 'almost' closed form solution. More general economies can be treated with a gradient approach, if they offer the property of differentiability. In (Karatzas *et al.*, 1990) a one good economy is analyzed in this way.
The example given here is a generalization of the economy presented by (Zapatero, 1995). Zapatero does not take into account some features included here, e.g. population growth and production. This leads in his model to non–stochastic interest rates and prices of risk. While the following example economy widens the intuition where the randomness in these variables could come from. Although Zapatero made a very good starting point, he surprisingly left open the determination of the exchange rate at period zero. This is quite important because it shows which paying streams are responsible for the settlement of the exchange rate.

### 3.4.1 Equilibrium Model

Introduced now are all the ingredients of a real economy into the finance model. An example economy is chosen, where an equilibrium is explicitly computable. The assumptions will be made for this reason only while not being too unrealistic.

Consider an economy with two representative agents $i \in I$, two consumption goods understood as the countries' numéraires and two firms, which use only the labor supply of the local population as an input factor. Each country's representative agent offers his labor to the home industry. The agents are scaled by the population size of each country. Good markets and financial markets are frictionless integrated, whereas labor markets are fully separated.

All markets operate competitive.

**Firms:** There are two firms $i \in I$ each with a linear production technology in their labor demand $L^i$. Real dividends $D^i$ per unit of time measured in the corresponding domestic or foreign numéraire are defined by:

$$D_t^i = \gamma_t^i L_t^i + \delta_t^i.$$

$\gamma^i$ is the productivity of labor and $\delta^i$ should be interpreted as capital gains or simply real endowments. Note that the term 'dividends' is used here not in the usual sense as the payment to stockholders already net of factor costs. The dividends split up in a wage and a capital interest share. The two produced goods are offered in both countries for consumption. This happens without transaction costs. The involved coefficients follow the technological change:

$$\begin{aligned} d\gamma_t^i/\gamma_t^i &= \mu_t^{\gamma i} dt + \sigma_t^{\gamma i} dW_t, \\ d\delta_t^i/\delta_t^i &= \mu_t^{\delta i} dt + \sigma_t^{\delta i} dW_t \\ \text{with } 0 &\leq \gamma_0^i, \delta_0^i, \end{aligned}$$

where not both $\gamma_0^i$ and $\delta_0^i$ are equal to zero. The production is said to be restricted if the labor demand is not allowed to take negative values. This is quite reasonable for the factor labor, but one could re–interpret $L$ as the use of environmental resources and a negative value as an improvement of the environment. In the following both cases are allowed simultaneously. Unfortunately the restriction prevents the equilibrium from having a closed from solution. That the firm contains production capabilities as well as resources is not a limitation because profits from a linear technology in a competitive market are zero. Furthermore, complete markets are assumed. This ensures that all risks are traded also separately.

**Markets:** The real locally riskless good–account, $1/b^i$, is available for both consumption goods. The stocks of productive technologies are traded with prices $S^i$ measured in the corresponding consumption good. Furthermore, there exist at least $m-3$ financial assets, such that markets are complete. Financial assets are in zero net supply. Because the productive assets pay a dividend steam, the gain processes have to be defined. The gains up to period $t$ are the accumulated dividends minus the labor costs plus the price in $t$ :

$$G_t^i = S_t^i + \int_0^t \left(D_s^i - w_s^i l_s^i\right) / b_{s,t}^i ds.$$

The wage per unit of labor, $w^i$, is measured in the corresponding currency.

**Populations:** Agents' preferences of both countries can be aggregated to a social utility function, $U^i : \mathcal{L}^2 (P \times \lambda)^3 \rightarrow \mathbb{R}$, of time–separable, von–Neumann–Morgenstern type.[107] Consumption takes place in the domestic as well as in the foreign good, $c^d$ and $c^f$, and in leisure time, $l$:

$$U^i\left(c^d, c^f, l\right) = E_P\left[\int_0^{\mathcal{T}} K_s^i n_s^i u^i\left(c_s^d, c_s^f, l_s\right) ds\right].$$

Time–preferences, $K_s^i$, are strictly decreasing at a rate $\kappa_t^i$ and not vanishing:

$$\begin{aligned} \forall t &\in \mathbb{T} : \kappa_t^i > 0, \text{ with} \\ K_s^i &= \exp\left(-\int_0^s \kappa_u^i du\right). \end{aligned}$$

Moreover, the population has a starting size of $n_0^i > 0$ and grows at a deterministic rate

$$dn_t^i / n_t^i = \mu_t^{ni} dt.$$

[107]The aggregated utility function is in the kind of Bentham–utilitarian. This choice is made for simplicity only. And it is ethical acceptable only if all agents under control of the planner are identical and the population growth is 'optimal'. Since this issue is not the objective of the chapter the reader is referred to (Roemer, 1996). A dynamic stochastic consumption model with heterogenous agents is exemplary studied in (Karatzas *et al.*, 1990).

The utility per unit of time is assumed to be logarithmic[108]:

$$\begin{aligned} u^i\left(c_s^d, c_s^f, l_s\right) &= \alpha^{di}\ln c_s^d + \alpha^{fi}\ln c_s^{fi} + \left(1-\alpha^{di}-\alpha^{fi}\right)\ln\left(\Phi^i - l_s\right), \\ & \quad \alpha^{fi}, \alpha^{di}, \alpha^{di}+\alpha^{fi} \in (0,1), \end{aligned}$$

with total individual free–time in an amount of $\Phi^i > 0$.
**Budget sets:** Each country has an initial endowment $\theta^i = \left(\theta^{di}, \theta^{fi}\right) \in \mathbb{R}_+^2$ of productive assets, which are in positive net supply, $\theta^d + \theta^f \in \mathbb{R}_{++}^2$. Let the income in local currency of country $i$ in period zero be defined as

$$\mathfrak{b}_0^i \equiv \mathfrak{b}^i\left(S_0^d, S_0^f, X_0\right) := \theta^{ii}S_0^i + \theta^{\neg ii}S_0^{\neg i}X_0^i.$$

Then the budget set of each country sounds

$$\begin{aligned} \mathfrak{B}^i\left(X, w^i, S_0^d, S_0^f, r^i, Q^i\right) = \Big\{ \left(c^d, c^f, l\right) \in \mathcal{L}^2\left(P\times\lambda\right)^3 \Big| \\ \mathfrak{b}_0^i \geq E_{Q^i}\left[\int_0^{\mathcal{T}} b_{0,s}^i n_s^i\left(c_s^{ii} + c_s^{\neg ii}X_s^i - w_s^i l_s\right) ds\right]\Big\}. \end{aligned}$$

This completes the description of the economy.

### 3.4.2 The Equilibrium Concept

It is now intended to define an equilibrium for this economy. The equilibrium in this example will not be a financial market equilibrium but a no–arbitrage equilibrium, which is also called the martingale approach. This means that at equilibrium state prices, which correspond in a continuous state space to the pricing density times the deflator $z_t^i b_{0,t}^i$, all future markets are open. The future consumption and labor supply is determined in period zero for all measurable future events. This allows one to abstract from modelling portfolio strategies which actually generate the optimal decisions. But it is possible to support the consumption and labor plans by a suitable portfolio

[108] A logarithmic utility function is justified by that it maximizes the expected growth rate of investors' wealth. This fact implies the long–run survival of such investors with logarithmic utility as it has been shown in evolutionary finance models, see (Blume & Easley, 1992).

strategy since they are square integrable and markets are complete (see for instance (Karatzas & Shreve, 1998) Ch. 3).

**Definition 96 (No-arbitrage-Equilibrium)** *A no–arbitrage–equilibrium is a tuple of processes:*

$$\left(\left(\tilde{c}^{di}, \tilde{c}^{fi}, \tilde{l}^{i}, \tilde{w}^{i}, \tilde{L}^{i}, \tilde{S}^{i}, \tilde{r}^{i}, \tilde{\xi}^{i}\right)_{i\in I}, \tilde{X}\right) \in \mathcal{L}^{2}\left(P \times \lambda\right)^{2(7+m)+1},$$

*such that for $i \in I$ :*

1. *Agents maximize utility*

$$\left(\tilde{c}^{di}, \tilde{c}^{fi}, \tilde{l}^{i}\right) \in \underset{\left(c^{di}, c^{fi}, l^{i}\right)\in\mathfrak{B}^{i}\left(\tilde{X}, \tilde{w}^{i}, \tilde{S}_{0}^{d}, \tilde{S}_{0}^{f}, \tilde{r}^{i}, Q^{i}\right)}{\arg\max} U^{i}\left(c^{di}, c^{fi}, l^{i}\right).$$

2. *Firms maximize share holder value*

$$\begin{aligned} \tilde{L}^{i} &\in \arg \max_{L^{i}\in\mathcal{L}^{2}(P\times\lambda)} S_{0}^{i}\left(L^{i}\right), \\ \tilde{S}^{i} &= S^{i}\left(\tilde{L}^{i}\right). \end{aligned}$$

3. *The discounted gain processes of productive assets as well as discounted financial asset prices are martingales in the domestic and the foreign currency with respect to the respective martingale measure $Q^{d}$ or $Q^{f}$, which are defined in the following way:*

$$\left.\frac{dQ^{i}}{dP}\right|_{\mathcal{F}_{t}} = z_{t}^{i}, \quad \textit{where } z_{t}^{i} = \mathcal{E}\left(\int_{0}^{t} \tilde{\xi}_{u}^{i} dW_{u}\right) \quad P \times \lambda - a.s.$$

4. *Good markets clear, that is:*

$$n_{t}^{d}\tilde{c}_{t}^{id} + n_{t}^{f}\tilde{c}_{t}^{if} = D_{t}^{i}\left(\tilde{L}_{t}^{i}\right) \quad P \times \lambda - a.s.$$

5. *Labor markets clear, i.e.:*

$$n_{t}^{i}\tilde{l}_{t}^{i} = \tilde{L}_{t}^{i} \quad P \times \lambda - a.s.$$

*According to the case considered the employment of labor is either restricted to be non–negative, i.e.* $\tilde{l}_t^i \geq 0$*, or not.*

Before the equilibrium is solved, three abbreviations are introduced to shorten the notation thereafter:

**Definition 97** *1. The 'discounted' population growth is defined as:*

$$K^i = \exp\left(\int_0^{\mathcal{T}} \mu_s^{ni} - \kappa_s^i ds\right).$$

*2. Furthermore, the present value of free time in local currency is denoted by:*

$$\mathcal{W}_0^i \equiv \mathcal{W}_0^i\left(\tilde{w}^i, \tilde{r}^i, \tilde{\xi}^i\right) = \Phi^i E_{Q^i}\left[\int_0^{\mathcal{T}} \tilde{w}_s^i n_s^i b_{0,s}^i ds\right].$$

Moreover, ruled out are any technical complications by assuming bounded coefficients:

**Assumption:** The coefficients $\mu^{\gamma i}, \sigma^{\gamma i}, \mu^{\delta i}, \sigma^{\delta i}, \kappa^i, \mu^{ni}$ are non–stochastic and bounded functions of time on the domain $\mathbb{T}$, whereas the coefficients $\alpha^{ji}$ are constant.

Obviously, this assumption can be relaxed considerably, but this is by means not the intention of this example.

### 3.4.3 Equilibrium Prices and Allocation

The whole solution is placed in a single proposition. An explanation goes through each part of the equilibrium afterwards. The proof is given in the Appendix 3.8. Between the two cases – restricted and unrestricted labor supply and demand, respectively – is differentiated by an indicator function to consider both economies in a single proposition.

**Proposition 98** *An unique equilibrium exists. Let un–/restricted labor supply be indicated by* $c = -\infty/0$, *and by the simple functions* $1_{l_s^i} \equiv 1_{\{l_s^i>0\}}$, *and* $\bar{\gamma}_s^i = 1_{l_s^i}\gamma_s^i$.

1. *The martingale measures are determined by the two prices of risk:*

$$\tilde{\xi}_t^i = -\left[\frac{n_t^i\bar{\gamma}_t^i\Phi^i}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i}\sigma_t^{\gamma i} + \frac{\delta_t^i}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i}\sigma_t^{\delta i}\right]^T.$$

2. *The instantaneous compound interest rates are:*

$$\begin{aligned}\tilde{r}_t^i &= \left(\kappa_t^i - \mu_t^{ni}\right)\left(1 - \frac{n_t^{\neg i}\tilde{c}_t^{\neg i}}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i}\right) + \left(\kappa_t^{\neg i} - \mu_t^{n\neg i}\right)\frac{n_t^{\neg i}\tilde{c}_t^{\neg i}}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i} \\ &\quad + \frac{n_t^i\bar{\gamma}_t^i\Phi^i}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i}\left(\mu_t^{\gamma i} + \mu_t^{ni}\right) + \frac{\delta_t^i}{n_t^i\bar{\gamma}_t^i\Phi^i + \delta_t^i}\mu_t^{\delta i} - \left\|\tilde{\xi}^i\right\|^2.\end{aligned}$$

3. *The exchange rate follows the following stochastic process:*

$$\begin{aligned}\tilde{X}_t &= \tilde{X}_0\frac{b_{0,t}^f z_t^f}{b_{0,t}^d z_t^d}, \text{ with the starting value} \\ \tilde{X}_0 &= \frac{\alpha^{fd}\left(\theta^{dd}S_0^d + \tilde{b}^d\mathcal{W}_0^d\right)/\tilde{a}^d + \left(1 - \alpha^{df}/\tilde{a}^f\right)\theta^{df}S_0^d}{\alpha^{df}\left(\theta^{ff}S_0^f + \tilde{b}^f\mathcal{W}_0^f\right)/\tilde{a}^f + \left(1 - \alpha^{fd}/\tilde{a}^d\right)\theta^{fd}S_0^f}.\end{aligned}$$

4. *Labor markets clear at wages* $\tilde{w}_t^i = \gamma_t^i$.

5. *Prices of productive assets in local currency follow the process:*

$$\tilde{S}_t^i = \delta_t^i E_P^t\left[\int_t^{\mathcal{T}} \exp\left(\int_t^s \mu_u^{\delta i} - \tilde{r}_u^i + \sigma_u^{\delta i}\tilde{\xi}_u^i du\right)\mathcal{E}\left(\int_t^s \sigma_u^{\delta i}dW_u\right)ds\right].$$

6. *Utility marginally increases in expenditure by*

$$1/\tilde{\eta}^i = \frac{\mathfrak{b}_0^i + \tilde{b}^i\mathcal{W}_0^i}{n_0^iK^i\tilde{a}^i},$$

*with* $\left(1 - \alpha^{ii} - \alpha^{\neg ii}\right) < \tilde{a}^i \leq 1, 0 < \tilde{b}^i \leq 1$. *The constants* $\tilde{a}^i$ *and* $\tilde{b}^i$ *are* *endogenous and are equal to one if labor is not restricted. The total expenditure of country i is* $\mathfrak{b}_0^i + \tilde{b}^i \mathcal{W}_0^i$. *The share spent on per capita per period per state consumption and leisure time measured in terms of the local good is* $e_s^i = \frac{K_s^i}{z_s^i b_{0,s}^i \tilde{\eta}^i}$. *Therewith the consumption and leisure time reads:*

$$\begin{pmatrix} \tilde{c}_s^{ii} \\ \tilde{c}_s^{\neg ii} \end{pmatrix} = e_s^i \begin{pmatrix} \alpha^{ii} \\ \alpha^{\neg ii} X_s^{\neg i} \end{pmatrix},$$
$$\tilde{l}_s^i = \max\left[\Phi^i - e_s^i \left(1 - \alpha^{ii} - \alpha^{\neg ii}\right) / \tilde{w}_s^i \, ; c\right].$$

**Proof.** See Appendix 3.8. ■

Following is an interpretation of the results. For this reason suppose that labor is not restricted in equilibrium, i.e. let $\tilde{a}^i = \tilde{b}^i = 1$ and $c = -\infty$ with positive labor supply. That does not change the qualitative results but makes them better understandable.

The prices of risk are stochastic convex combinations of the diffusions of technology and resources. The convex combination weights the diffusions according to the share on the dividends, which either the production or the resources have. The risk premia of financial assets in currency $i$ increase if their diffusion points 'more' in the direction of the price of risk, which is the convex combination of $\sigma_t^{\gamma i}$ and $\sigma_t^{\delta i}$. If the diffusion of a security were pointing in the opposite direction, it would offer an insurance to the risk of production and resources, which would imply a relative high price and a low return. The converse is probably true for the productive assets (see 5.), namely if the share of production is not dominating that of dividends, while its diffusion is pointing in 'some completely different direction' than the diffusion of the resources. In this case $\sigma_t^{\delta i}$ is acute–angled to $-\tilde{\xi}^i$, which implies a positive risk premium for the gain process. That means high returns at a lower starting price $\tilde{S}_0^i$ since the risk premium $-\sigma_t^{\delta i} \tilde{\xi}_t^i$ increases the discount factor of resources under $P$, which is $\tilde{r}_t^i - \sigma_t^{\delta i} \tilde{\xi}_t^i$.

The short rates are the sum of five expressions. The first two as well as the second two expressions belong together. The first pair is a convex combination of the discount factors minus the growth rates of the countries' population. Consider the short rate of the domestic country, i.e. $i = d$. The weight of the domestic country is one minus the export share, which is then consequently the weight of the foreign country. The export share is the share of the foreign consumption in terms of the total expenditure in the domestic good, which is the value of domestic free time plus resources. The interest rate compensates for forgone consumption mirrored in the discount rate. Moreover, a higher population growth increases scarcity, i.e. future consumption gets relatively more valuable. This implies a decreasing interest rate or, in other words, an increasing price of future consumption. Note that one has only a closed form solution if discounted population growth, $-\left(\kappa_t^i - \mu_t^{ni}\right)$, is the same for both countries. Otherwise the consumption of the foreign population does not cancel down. But the solution for the optimal consumption (see 6.) contains the discount factor, which itself depends on the short rate process.

The next two expressions are a stochastic convex combination of two expected growth rates. On the one hand this is the growth rate of production, which is itself the sum of the growth rates of technology and labor force, and, on the other hand, the growth rate of resources. If dividends grow fast the relative abundance between consumption in future periods and today is high, which makes it less attractive to save at low interest rates. Therefore in equilibrium the short rate is increasing in the growth of production. The weights of the convex combination are the same like in the prices of risk, i.e. the respective shares on the dividends.

The last expression is the price of the price of risk. Higher risk in production, which means an increasing probability of scarcity and abundance in the future, is reflected in lower interest rates. For risk averse agents savings are the more desirable to insure against scarcity the riskier the environment is. To balance out agents' risk aversion leading to a higher demand for savings the short rate has to decrease in the risk of production.

It should be remarked that real interest rates might even be negative if the populations grow fast and productivity of labor is small. Then the increase in production by the growing labor supply does not outweigh the scarcity caused by the fast increase of the population. For the planner it is then better to postpone consumption into the future for increasing the happiness of more people.[109]

The law of motion of the exchange rate is of course the same as in the arbitrage pricing model. But now the level of the exchange rate in period zero can be determined. It is the ratio of the expenditure received in domestic but spent in foreign currency and the expenditure received in foreign but spent in domestic currency. The nominator is the sum of the income (asset shares and present value of free time) of the domestic population in the domestic good, which is spent on foreign consumption, and the income of the foreign population in the domestic good (only asset shares), which is spent on consumption in the foreign good and foreign free time. The denominator is the reversed expression measured in the foreign good. This corresponds to an equilibrium on the currency market, since this is the ratio of the present values of exports and imports over the whole lifetime settled in period zero. The proof reveals that Walras' law enforces this – so called – central bank equilibrium.

Obviously, competitive labor markets clear at a wage level equal to the productivity. Then profits from the linear technology are zero. This implies that the prices of productive assets coincide with the valuation of resources.

Consumers' optimal decisions are typical for log–utilities. Expenditure for consumption and free time share the budget provided for the particular period and state corresponding to their weights in the utility function. The total expenditure is the sum of the present values of productive assets and free time the population is endowed with. It is split up for per capita, per period, and per state expenditure, which decreases with countries' individual discount factor, but increases with the short rate and the submartingale $1/z_t^i$. This submartingale is responsible for the insurance of scarce states, since it moves in the opposite direction than dividends. Altogether, $e_t^i$ is a contin-

[109] Actually, there seems to be empirically a correlation between high population growth, poorness and high time preferences, which prevents real interest rates from becoming negative.

uous submartingale. The marginal increase of utility in total expenditure decreases with the discounted population size, because dividends must feed more people.

### 3.4.4 Further Conclusions

Now the remaining two questions asked at the beginning of this section are answered: Which assets play an important role in the projection of Corollary 95 and how could the coefficients $\beta$ look like? As long as assets form a complete market the choice is irrelevant. But if one were taking an arbitrary financial asset, it would either not exist as such in real financial markets or a hypothesis about $\beta$ might be very difficult. Two kinds of securities usually exist, productive assets and bonds. Stocks represent the risk of production and resources, while bonds reflect scarcity. $\beta$ is the unique solution of the following linear equation:

$$\begin{aligned}\sigma_t^X &= \left(\xi_t^f - \xi_t^d\right)^T \\ &= \left[\upsilon_t^d \sigma_t^{\gamma d} + \left(1 - \upsilon_t^d\right) \sigma_t^{\delta d}\right] - \left[\upsilon_t^f \sigma_t^{\gamma f} + \left(1 - \upsilon_t^f\right) \sigma_t^{\delta f}\right] \\ &= \beta_t \begin{pmatrix} \sigma_t^d \\ \sigma_t^f \end{pmatrix}, \quad \text{where } \upsilon_t^i \in (0,1)\,.\end{aligned}$$

To learn $\beta_t$ the diffusions of the assets $\sigma_t^i$ have to be determined. Unfortunately this is not possible in a closed form solution for the productive assets and the bonds, because it incorporates expected values about future volatilities. With the lack of a proof the next assumption needs to be justified:

**Assumption** The sign of the coefficients in $\beta$ are positive for domestic and negative for foreign securities, if these incorporate the country specific risks and do not insure these risks. Bonds and productive assets are of this kind.

This assertion is quite reasonable if the diffusion of assets reflect only the countries' specific risks and do not insure these. This holds for instance if both countries' risk factors are separated, which is the case when the diffusion matrix is a block matrix:

$$\begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} = \begin{pmatrix} \tilde{\sigma}^d & \mathbf{0} \\ \mathbf{0} & \tilde{\sigma}^f \end{pmatrix},$$

and countries' specific risks lie in the positive cone of $\tilde{\sigma}^i$, i.e.

$$\exists \theta^{ji} \in \mathbb{R}^{n^i}_{++} : \sigma^{ji} = \left(\theta^{ji}\right)^T \tilde{\sigma}^i, \ j = \gamma, \delta.$$

This implies the assertion for $\beta = \left(\left(\beta^d\right)^T, \left(\beta^f\right)^T\right)^T$ stated in the conjecture:

$$\begin{aligned} \beta^d &= \left(\upsilon^d \theta^{\gamma d} + \left(1 - \upsilon^d\right) \theta^{\delta d}\right)^T \quad \text{and} \\ \beta^f &= -\left(\upsilon^f \theta^{\gamma f} + \left(1 - \upsilon^f\right) \theta^{\delta f}\right)^T \\ &\Rightarrow \ \beta^d > \mathbf{0} \ \text{ and } \beta^f < \mathbf{0}. \end{aligned}$$

Bonds and productive assets are usually not insurances against market risks; quite the reverse is true, since those assets mirror the country specific risks. Arguing with the $\beta-$determining equation this means that $\sigma_t^i$ points more into the direction of $\xi_t^i$ than of $\xi_t^{\neg i}$. This is even more the case if nominal prices are considered. Then the diffusion vector of the price index adds to that of the prices of risk as well as to the diffusions of assets:

$$\begin{aligned} \hat{\sigma}_t^X &= \left(\hat{\xi}_t^f - \hat{\xi}_t^d\right)^T \\ &= \left[\upsilon_t^d \sigma_t^{\gamma d} + \left(1 - \upsilon_t^d\right) \sigma_t^{\delta d} + \sigma_t^{qd}\right] - \left[\upsilon_t^f \sigma_t^{\gamma f} + \left(1 - \upsilon_t^f\right) \sigma_t^{\delta f} + \sigma_t^{qf}\right] \\ &= \hat{\beta}_t \begin{pmatrix} \sigma_t^d + \mathbf{1}_{n^d} \sigma_t^{qd} \\ \sigma_t^f + \mathbf{1}_{n^f} \sigma_t^{qf} \end{pmatrix}, \text{ where } \upsilon_t^i \in (0,1). \end{aligned}$$

In nominal quantities the diffusion of assets are even more pointing into the direction of the home countries' nominal prices of risk, so that the conjecture should hold in the nominal economy as well.

## 3.5 Empirical Evidence

The following two subsections show an estimation of the projection of the exchange rate changes from Corollary 95. The third subsection focuses on an error correction model implied by Corollary 91. The first estimation is done in two steps. First, an econometric equation is derived suitable for testing the hypothesis of a linear projection. And in the next subsection it is applied to the data set.

Before turning to the question of the empirical validity of the linear projection, one should stress the compromises made in this estimation, in which the practical problems come into play. One cannot seriously maintain to find a manageable set of assets which span the whole uncertainty of financial markets despite the fact that markets might be incomplete. Though the analysis is restricted to a relatively small set of assets complete markets are still assumed. But probably a large set of assets is left out, which could be very valuable in explaining the exchange rate changes. Unfortunately the coefficients almost surely change by altering the set of assets taken into the estimation. But even if markets were incomplete, it would be interesting to analyze whether some kind of orthogonal projection is significant.
An even worse problem is that of a stochastic or time dependent vector of coefficients, $\beta$. In the equilibrium model it has been shown that the prices of risk are stochastic. The diffusions of the securities are generally also stochastic or at least in some complex manner time dependent. Also assets' diffusion and prices of risk determine the vector of coefficients. This makes a regression of the exact relationship almost impossible. This difficulty is not taken into account. A justification for doing so is that the parameters might be stable on a relatively short time horizon considered here. This is of course not theoretically founded and turns out to be not validated on the entire data set.

One might ask the question why the relations stated in the equilibrium model are not estimated directly. There are at least two reasons at hand. First, the equilibrium model is fairly special, so that one can hardly expect to regain it in the data. Indeed there are many unsuccessful empirical investigations

of the influence of fundamental values on the exchange rate changes. But the most convincing reason is that an arbitrage relation should also hold in disequilibrium. It is a much weaker presumption.
The estimation will stand on its own, which means that it is less a test of the model but more that of a linear relationship between logarithmic asset and exchange rate changes. To understand the estimation as a validation of the arbitrage pricing model would be very demanding, since it implies a joint test of all assumptions and simplifications made so far (especially that of time invariant coefficients). It should be seen the other way around. The model has helped to derive an econometric equation, which is not a priori inconsistent with the theory, like the UIP.

### 3.5.1 Econometric Specification

The investigation concentrates on a selection of $n$ securities, $n = n^i$, for each country which are supposed to have a significant effect on the foreign exchange rate: The stock–market indices, $S^i$, and a collection, $j \in \{2, ..., n\}$, of zero bonds, $B^i_j$, with different durations, $\delta \in \{\delta_j | j = 2, ..., n\}$, where $\delta_j = T_j - t > 0$ is increasing in $j$. The bonds represent the term structure of interest rates, probably the most important determinant, and the stock market indices show in contrast the returns on production.[110] Geometric Brownian motions are assumed for the stock market indices:[111]

$$dS^i/S^i = \mu^{Si} dt + \sigma^{Si} dW.$$

The arbitrage pricing model allowed also for bond prices. In the case of zero bonds it is easier to observe the movement of interest rates instead that of bond prices themselves. But obviously, there is a functional relationship between the annual effective interest rate and the bond price of the corre-

[110] It might appear more meaningful to take especially the export and import industry into consideration, but the necessary data was not available.
[111] This is a working hypothesis generally made in applied finance since it simplifies many technicalities in pricing as well as in the estimation. This assumption is criticized to be wrong by (Lo & MacKinlay, 1988). Non–Brownian processes would alter the F–statistics of coefficients and the efficiency of the least squares estimation.

sponding duration. Therefore some arbitrary Itô–processes are specified for the interest rates to show how bond prices emerge. Let the annual effective interest rate $r_j^i$ for the duration $\delta_j = T_j - t > 0$ be specified by the following Itô–process:

$$dr_j^i = \mu_j^{ri} dt + \sigma_j^{ri} dW_t.$$

Only zero rates are considered, so that the zero bond price is defined by:

$$B_j^i(t) = \left(1 + r_j^i(t)\right)^{-\delta_j}$$

Applying Itô's rule to $\ln B_j^i(t)$ shows that the bonds are not geometric Brownian motions as well:[112]

$$\begin{aligned} d\ln B_j^i(t) &= \ln\left(1 + r_j^i(t)\right) dt - \delta_j d\ln\left(1 + r_j^i(t)\right) \\ &= \left[\ln\left(1 + r_j^i(t)\right) - \frac{\delta_j \mu_j^{ri}(t)}{1 + r_j^i(t)} - \frac{\delta_j \left\|\sigma_j^{ri}(t)\right\|^2}{2\left(1 + r_j^i(t)\right)^2}\right] dt \\ &\qquad - \underbrace{\frac{\delta_j \sigma_j^{ri}(t)}{1 + r_j^i(t)}}_{:=\sigma^{B_j i}} dW_t. \end{aligned} \tag{DLogBond}$$

Interest rates for certain times to maturity are considered. This means that instantaneously at every moment another set of bonds as the hedge–instruments is employed. These bonds have exactly the preselected set of durations as their times to maturity. The reason for this detour is the better availability over a long time horizon of interest rates for conventional times to maturity, say for example 3 month, 1 year, 5 and 10 years. This is indeed not the case for bond prices, whose times to maturity are instantaneously decreasing.[113] Moreover, the $\beta-$factor, which has to be estimated in the se-

[112] A continuous semi–martingale is a general geometric Brownian motion, if the trend and the diffusion of its log–SDE are non–stochastic. This is not the case here because the interest rates are assumed to be stochastic.

[113] The interest rates of arbitrary durations could be gained from an interpolation of the discrete yield curve.

quel, is supposed to be more stable over time if the durations are kept fixed. It depends on the diffusion of the assets, which in the case of bonds depends on the duration (see $\sigma^{B_j i}$ above). A diminishing duration would imply from the beginning a changing $\beta-$factor. With fixed durations the diffusions still depend on the changing interest rates and their diffusions, which is not perfect but a problem of minor size. The estimation *presumes* time independent coefficients of the Brownian motion for the bonds as well.

The estimation of general Itô–processes is a rather difficult task with not always a satisfying convergence of estimators. This is due to a possibly non–linear form of the stochastic differential equation, which has to be suitably approximated in discrete time, see (Lo, 1988) for this problem. The advantage of the geometric Brownian motions assumed here is that the log–differences of the variables are linear in the coefficients in discrete as well as in continuous time.

For the estimation one has to discretize the no–arbitrage representation of the exchange rate to fit in the daily observations. 'Daily' means that there is no more than one working day between two observations. The time difference $T-t$ is measured in days and is for the weekdays equal to one and over the weekends and off–days up to 3, seldom 4 days. In the arbitrage pricing model time was measured on an annual basis. This adjustment is obtained by dividing through 365 days per year.[114] The exchange rate of one day is taken into account only if the other variables are also observable. The representation in Corollary 95 suggests the following linear econometric equation in discrete time:

$$\Delta_{t,T} \ln X = \alpha_0 + \underbrace{\frac{\alpha_1}{365}(T-t-1)}_{=:\alpha_1 DDays_{t,T}} + \underbrace{\frac{\alpha_2}{365}\sum_{s=t}^{T}\left(r_s^d - r_s^f\right)}_{=:\alpha_2 DIR_{t,T}} + \beta^T Y_{t,T} + u_{t,T},$$

[114]The actual/actual–valuation is convenient in calculations of yields up to one year.

with $E_P^t(u_{t,T}) = 0$ and in which the explaining factors $Y_{t,T}$ are

$$Y_{t,T} = \Delta_{t,T} \ln \begin{pmatrix} S^d \\ B^d \\ S^f \\ B^f \end{pmatrix} - \frac{1}{365} \sum_{s=t}^{T} \left[ \begin{pmatrix} r_s^d \cdot \mathbf{1}_n \\ r_s^f \cdot \mathbf{1}_n \end{pmatrix} + \right.$$

$$\left. \begin{pmatrix} -Cov_{s,\pm w}\left(\Delta \ln \left(S^d, B^d\right), \Delta \ln X\right) \\ Cov_{s,\pm w}\left(\Delta \ln \left(S^f, B^f\right), \Delta \ln X\right) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} Var_{s,\pm w}\left(\Delta \ln \left(S^d, B^d\right)\right) \\ Var_{s,\pm w}\left(\Delta \ln \left(S^f, B^f\right)\right) \end{pmatrix} \right].$$

The log–differences of the stock market indices are observed directly, whereas those for bonds are calculated due to Formula (DLogBond) by:

$$\Delta_{t,T} \ln B_j^i = \frac{T-t}{365} \ln\left(1 + r_j^i(t)\right) - \delta_j \Delta_{t,T} \ln\left(1 + r_j^i\right). \qquad \text{(Bond)}$$

The durations $\delta_j$ have been chosen such that the interest rates $r_j^i(t)$ are provided by the financial institutions. Whereas the continuously compound interest rates, $r^i$, are only measurable by their corresponding annual effective rates, $\widetilde{r}^i$. For this reason the relationship $r^i = \ln(1 + \widetilde{r}^i)$ covers the short rate differential. For short time intervals $[t, T]$ this implies the approximation:

$$\begin{aligned} \frac{1}{365} \int_t^T r^i(s)\, ds &= -\ln b_{t,T}^i \\ &\simeq \ln \prod\nolimits_{s=t}^{T} \left(1 + \tilde{r}^i(s)\right)^{1/365} \\ &\simeq \frac{(T-t)}{365} \ln\left(1 + \tilde{r}^i(t)\right). \end{aligned}$$

There is, for example, no approximation error if the short rate is constant on short time intervals.

The vectors in the second row of the explaining variables are the covariances of assets with the exchange rate ($Cov_{s,\pm w}$) and the variances of assets ($Var_{s,\pm w}$). Incorporating unobservable moments in the estimation leads to the serious problem of how to measure those time series. The co–/variances are estimated by the usual non–parametric estimator with an arbitrary window, $w$, of 20 observations around the current date $s$. This generates the

time series of co–/variances finally used in the regression. There are other ways to achieve a variance–estimator: filtering rules, variance–models like GARCH, and implicit volatilities. The first alternative is not appropriate in continuous time because the filtering rule involves moments of higher order also unknown. The second one is an extended model with further problems of misspecification. The third method incorporates the problems with an option pricing model. For example, often implicit volatilities do not coincide for different strikes although they should. Since forecasting is not the primary interest future data can be used to generate the time series for the moments. Otherwise one would be restricted to historical volatilities, which perform not so well. The moving window pays tribute to the fact of time–varying co–/variances. The window size has been found by improving the regression via trial and error.
In many preliminary regressions it turned out that the adjustment of log–changes of assets by the co–/variances essentially worsens the regression. This might be due to a possibly bad quality of the unobservable co–/variances, for which a time series had to be generated first. For this reason the co–/variances are separated from the log–changes of assets by taking those as an extra explaining variable. This is done by the following definition of two new time series:

$$\begin{aligned} AssetVar &:= \frac{1}{2}\sigma^X \begin{pmatrix} \sigma^{Sd} \\ \vdots \\ \sigma^{B_n f} \end{pmatrix}^{-1} \left( \left\|\sigma^{Sd}\right\|^2, \ldots, \left\|\sigma^{B_n f}\right\|^2 \right)^T, \\ AssetCov &:= \sigma^X \begin{pmatrix} \sigma^{Sd} \\ \vdots \\ \sigma^{B_n f} \end{pmatrix}^{-1} \begin{pmatrix} -\sigma^{Sd} \\ \vdots \\ \sigma^{B_n f} \end{pmatrix} \left(\sigma^X\right)^T. \end{aligned}$$

In discrete time the unknown series is generated by:

$$\begin{aligned}
\sigma_s^X \begin{pmatrix} \sigma_s^{Sd} \\ \vdots \\ \sigma_s^{B_n f} \end{pmatrix}^{-1} &\simeq Cov_{s,\pm w}\left(\Delta \ln X, \Delta \ln\left(S^d, B^d, S^f, B^f\right)\right)^T \\
&\quad \cdot Cov_{s,\pm w}\left(\Delta \ln\left(S^d, B^d, S^f, B^f\right)\right)^{-1}, \\
\begin{pmatrix} -\sigma_s^{Sd} \\ \vdots \\ \sigma_s^{B_n f} \end{pmatrix} \left(\sigma_s^X\right)^T &\simeq Cov_{s,\pm w}\left(\Delta \ln X, \Delta \ln\left(-S^d, -B^d, S^f, B^f\right)\right), \\
\left(\left\|\sigma_s^{Sd}\right\|^2, \ldots, \left\|\sigma_s^{B_n f}\right\|^2\right) &\simeq Var_{s,\pm w}\left(\Delta \ln\left(S^d, B^d, S^f, B^f\right)\right).
\end{aligned}$$

$Cov_{s,\pm w}$ stands for the empirical covariance vector or matrix respectively and $Var_{s,\pm w}$ for the empirical variances both having a moving window of $w$ observations around date $s$.

The last problem addressed here is collinearity between the bonds. Obviously, if the five years interest rate is rising, the price of the corresponding bond is decreasing and very likely the ten years bond prices are also decreasing (if there is not an adverse movement in the 5 to 10 years forward interest rate). For sure the bond prices admit some collinearity. But collinearity of regressors makes it hard to interpret the estimated coefficients (see (Judge *et al.*, 1985) Ch. 22 for a discussion). To exclude this kind of collinearity (there might be other reasons) the influence of interest rates is separated with respect to different durations. This is done by subtracting from the row (in $Y$) belonging to a bond with the longer the row belonging to the bond with the next shorter time to maturity, i.e. $\Delta_{t,T} \ln B_j^i - \Delta_{t,T} \ln B_{j-1}^i$ $\forall j = 3, \ldots, n$.[115] This corresponds to the changes of forward prices for the period between the maturity of the shorter bond up to the maturity of the longer bond. There might still be collinearities, but with lower evidence as the tests point out in the next subsection.

[115] Analytically speaking, the rows corresponding to bonds of the vector $Y_{t,T}$ are multiplied with $-J_{n+n^*}(-1)^T$, where $J$ denotes a Jordan–block.

Finally the econometric equation reads:

$$\Delta_{t,T} \ln X = \alpha_0 + \alpha_1 DDays_{t,T} + \alpha_2 DIR_{t,T} + \alpha_3 (T-t) \, AssetVar_{t,\pm w} + \alpha_4 (T-t) \, AssetCov_{t,\pm w} + \beta^T \hat{Y}_{t,T} + u_{t,T}, \quad \text{(EcEqu)}$$

where

$$\begin{aligned}\hat{Y}_{t,T} = \Big( & \Delta_{t,T} \ln S^d - \frac{T-t}{365} r_t^d, \ \Delta_{t,T} \ln B_1^d - \frac{T-t}{365} r_t^d, \\ & \Delta_{t,T} \ln \left( B_2^d / B_1^d \right), \dots, \ \Delta_{t,T} \ln \left( B_n^d / B_{n-1}^d \right), \\ & \Delta_{t,T} \ln S^f - \frac{T-t}{365} r_t^f, \ \Delta_{t,T} \ln B_1^f - \frac{T-t}{365} r_t^d, \\ & \Delta_{t,T} \ln \left( B_2^f / B_1^f \right), \dots, \ \Delta_{t,T} \ln \left( B_n^f / B_{n-1}^f \right) \Big)^T .\end{aligned}$$

### 3.5.2 Estimation

What can be expected from the data? Indeed, two ambitious hypothesis are imaginable although one ought not to be convinced of their validity because of the many compromises that have been made:

**Unbiased Expectation Hypothesis:** The foreign exchange market is informationally efficient with respect to the arbitrage pricing model if the coefficients satisfy the following theoretical restrictions:

$$\begin{aligned}\alpha_0 &= \alpha_1 = 0, \\ \alpha_2 &= \alpha_3 = \alpha_4 = 1, \\ \beta_j &> 0 \text{ for } j = 1, \dots, n \text{ and} \\ \beta_j &< 0 \text{ for } j = n+1, \dots, 2n,\end{aligned}$$

provided that the generated time series, AssetVar and AssetCov, coincide with investors' expectations.

**Complete Market Hypothesis:** The chosen assets form complete financial markets if $u_{T,t} \equiv 0$.

The two hypotheses seem to be too demanding. One could be satisfied to find in the data some evidence of the linear representation, i.e. $\beta$ significantly different from zero with the supposed signs, and a significant influence of the short rate differential, i.e. $\alpha_2 \approx 1$.
The data set consists of about 2000 daily observations from January 1990 to November 1997 of the Deutsche Mark/US–Dollar exchange rate[116], interest rates for fixed durations as well as the over night short rates and two stock price indices, DAX and DOW.[117] Figure (5) shows the time series of the daily Deutsche Mark/US–Dollar exchange rate as well as its log–changes.

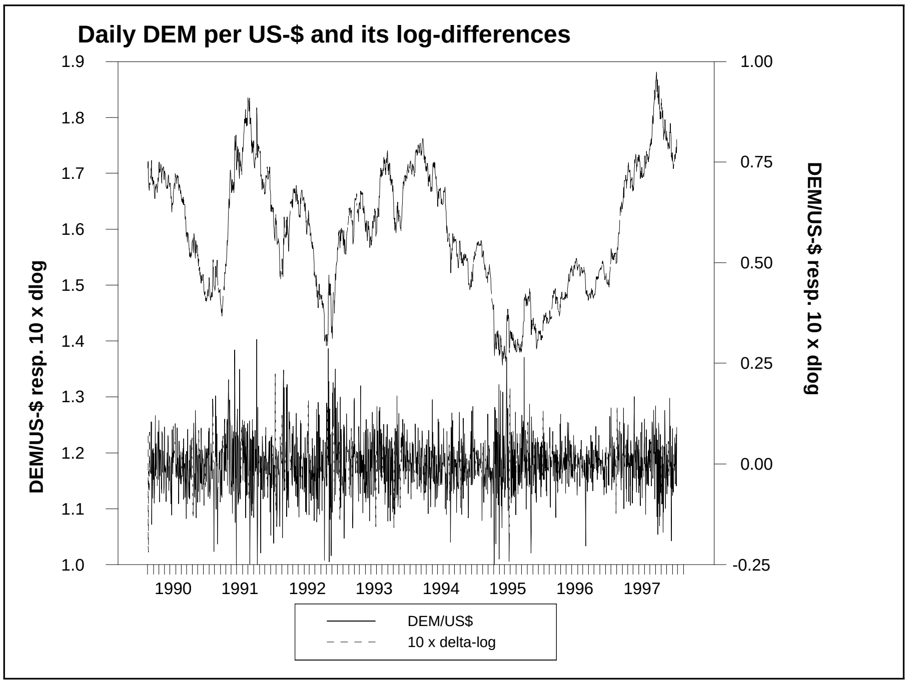


Figure 5: Deutsche Mark per US Dollar

Synthetic bond prices are calculated with the average interest rates on government debt. The conventional durations 10, 5, 1, and 1/4 years have been

[116]The German currency has been chosen instead of the €, because the time series is much longer. The EMU hurdles are avoided by considering the exchange rate only before 1998.
[117]The data sets have been provided unofficially by a London investment bank.

employed. The forwards used in the regression are denoted – for example – by 'D DEM 1Y 3M', which means: log–differences of a German government bond with a duration of 1 year minus its 3 months equivalent, i.e. the 3 months up to 1 year forward rate movement. Or 'DUS$10Y5Y' stands for the log–differences of the 5 to 10 years US–$ forward. In case of the 3 months government bond (DDEMBond3M) only the excess return over the one–day interest rate is considered.

To find out how the coefficients of Equation (EcEqu) behave over the entire time horizon, a rolling linear regression with a moving data–window of 500 observations has been carried out. The graphs of the coefficients indicate that they are more or less not constant over time, what has been expected due to the non–constant (co)variances. Figure (6) shows the coefficient of the DAX in this rolling regression.

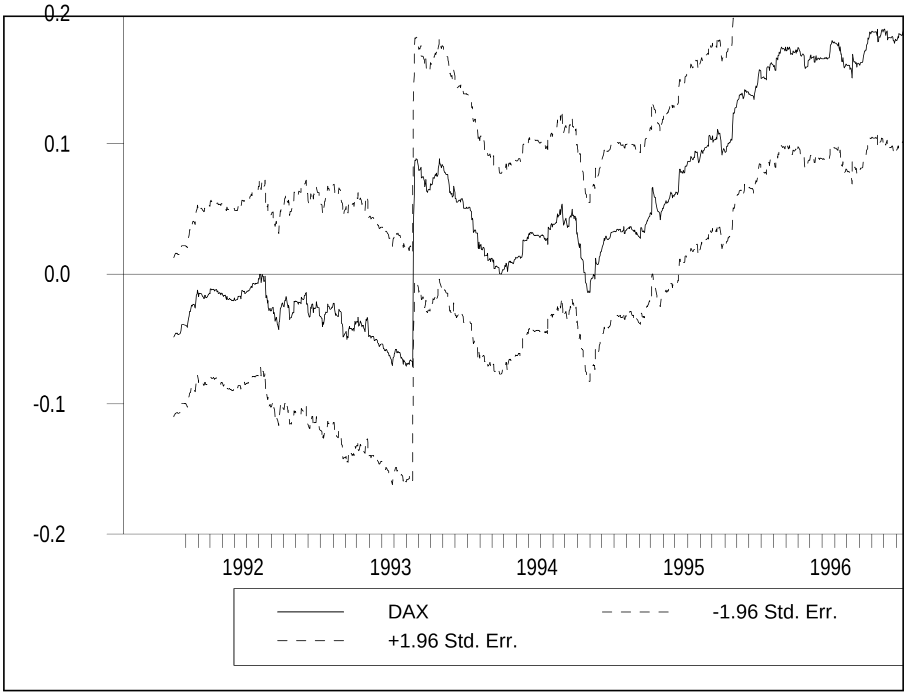


Figure 6: Recursive Regression, Coefficient of the DAX

The true value is most probably not constant because there is no constant laying in the ±1.96 standard error bands of the estimator. Moreover, the coefficient is probably negative until the mid of 1993 violating the conjecture of positive coefficients for domestic assets. In 1993 the estimator jumps to (mostly) positive values.

As this graph indicates, the linear regression is not very robust over the entire horizon. Therefore the more stable sub–period from the 1.1.1993 to the 27.11.1997 is considered with 1085 observations. The next table contains the results of the linear regression of Equation (EcEqu), i.e. the estimated coefficient of the named variable, its standard error and the significancy of having explanatory power for the variance of the log–differences of the exchange rate.[118]

| Variable | Coefficient | Std Error | Significancy |
|---|---:|---|---:|
| Constant | 0.000892 | 0.000399 | 2.6% |
| DDays | −0.138880 | 0.085225 | 1.1% |
| DIR | −3.135795 | 2.885749 | 27.8% |
| AssetVar | −0.919794 | 0.412230 | 2.6% |
| AssetCov | −0.492522 | 0.594426 | 40.8% |
| DDEMBond 3M | 0.113980 | 1.192965 | 92.4% |
| DDEM 1Y3M | −1.281451 | 0.550582 | 2.1% |
| DDEM 5Y1Y | 0.182546 | 0.079947 | 2.3% |
| DDem 10Y5Y | 0.309300 | 0.062536 | 0.0% |
| DDAX | 0.133817 | 0.020296 | 0.0% |
| DUS$Bond 3M | −1.917974 | 1.021684 | 6.1% |
| DUS$ 1Y3M | −0.641852 | 0.399051 | 10.9% |
| DUS$ 5Y1Y | −0.273538 | 0.113226 | 1.6% |
| DUS$ 10Y5Y | 0.243020 | 0.119502 | 4.3% |
| DDOW | −0.038089 | 0.026701 | 15.5% |

[118] The significancy indicates the (two sided) probability $sig$ that the true coefficient $\gamma_i$ has at least a distance of the absolute size of $\hat{\gamma}_i$ from the estimator: $\Pr(|\hat{\gamma}_i - \gamma_i| \geq |\hat{\gamma}_i|) \leq sig$. This event implies that the true coefficient could have the opposite sign only with probability less than $sig$. The t–test presumes normally distributed residuals, which is not exactly the case for the residuals of the following regression. An upper bound for the significancy is determined by Chebyshev's inequality: $sig \leq \left(\frac{StdErr}{Coeff}\right)^2$, which is e.g. about 2% for the DAX.

The variables DIR, AssetCov, DDemBond3M, DUS$ 1Y3M, DDOW are not significant to the 10%–level.[119] The centered $R^2 \simeq 11\%$ is low, but this is common to many empirical results concerning fx–markets. Moreover, the estimation becomes worse by aggregating the data up to one week. This might be a hint that fx–markets realize very fast new information and that high–frequency data could improve the estimation.


Figure 7: Recursive Regression, Coefficient of the Short Rate Differential

The coefficient of the interest rate differential is negative in contrast to the hypothesis. The rolling linear regression over the entire horizon shows, that

[119]This regression applied the "LINREG"–procedure from the statistical software "RATS". More results: Since just differences of logs are involved, one can be secure of cointegration. Further statistics exclude autocorrelation of the residuals ($DW = 2.03$, $Q(36-0) = 30$ (75%)) as well as correlation between residuals and regressors. The residuals are slight positively skewed and fat tailed relatively to a normal distribution. The variance–decomposition shows possible collinearity between the pairs: Constant/DDay, DUS$ 5Y1Y/DUS$ 10Y5Y, and weaker: DDAX/DDOW, AssetVar/AssetCov, AssetVar/DIntr. The regression without one variable of each pair does not change the remaining coefficients a lot, but reduces the $R^2$.

the coefficient of DIR takes negative values in some time intervals, is of relatively small absolute size, and seldom significantly different from zero, see Figure (7) above.

Moreover, Constant and DDays have coefficients significantly different from zero and the coefficient of DDEM 1Y3M and DUS$ 10Y5Y do not have the supposed sign, both in contrast to the conjecture.

The next estimation takes some freedom in the variables. The explanatory power is improved by using lagged and squared variables. The lagged variables are indicated by ($\tau$), in which case the variable is $\tau$ days behind the exchange rate. Squared variables could cover some degrees of non–linearity in the regression, which is denoted by an attached 'Q'. The lagged dependent variable is denoted by DFX DEMUS$. In the next regression those time series are additionally taken into account, which improve the $R^2$ and whose contribution is significantly different from zero. The following table on the next page shows the results.

The $R^2$ increased to 15.5%.[120] An astonishing result is that the US–$/DEM exchange rate shows to overshoot systematically.[121] This effect is decomposable into five factors corresponding to the variables with a significant lag structure. The coefficient of the more lagged variable has always the opposite sign of the more recent variable. This is also valid for the dependent variable itself, since the 3–day lag, DFX DEMUS$(3), has a negative contribution. The dependence on the short rate differential, DIR(1) + DIR(3), stayed negative.

Summarizing, the empirical results are not exciting, neither to sustain the conjectures made above nor to reject those. A linear projection of the log–fx–changes onto the excess returns appears successful, especially for the domestic assets, but the short rate differential has the opposite effect than expected.

[120] The results of the accompanied statistics remain qualitatively almost the same. Except, some more and some stronger collinearities appear, e.g. between DIR(1) and DIR(3), which is due to a non–zero trend in DIR.

[121] This empirically supports the macro–economic fx–model by (Dornbusch, 1976).

| Variable | Coefficient | Std Error | Significancy |
|---|---|---|---|
| Constant | 0.00025 | 0.00051 | 63.0% |
| DDays | −0.12170 | 0.09284 | 19.0% |
| DIR(1) | 56.29542 | 30.62790 | 6.6% |
| DIR(3) | −64.48941 | 30.29015 | 3.6% |
| AssetVar | −0.39110 | 0.42087 | 35.3% |
| DFX DEMUS$(3) | −0.05105 | 0.03079 | 9.8% |
| DDEMBond 3M(1) | −2.86279 | 1.29723 | 2.8% |
| DDEMBond 3M(4) | 1.45515 | 1.15009 | 20.6% |
| DDEM 1Y3M | −0.80551 | 0.61301 | 18.9% |
| DDEM 1Y3M(2) | 2.02343 | 0.55839 | ˜0% |
| DDEM 5Y1Y | 0.35408 | 0.08860 | ˜0% |
| DDEM 10Y5Y | 0.26394 | 0.06896 | ˜0% |
| DAX | 0.11891 | 0.02105 | ˜0% |
| DUS$Bond 3M | −1.83435 | 1.12394 | 10.3% |
| DUS$ 1Y3M | −0.20784 | 0.51815 | 68.8% |
| DUS$ 5Y1Y | −0.09853 | 0.07034 | 16.2% |
| DUS$ 10Y5Y(1) | −0.17189 | 0.08926 | 5.4% |
| DUS$ 10Y5Y(4) | 0.11333 | 0.07014 | 10.7% |
| DDOW | −0.01966 | 0.02808 | 48.4% |
| DDOWQ | −2.26843 | 0.90193 | 1.2% |
| DIRQ | 0.76284 | 0.42145 | 7.1% |

### 3.5.3 Error Correction by Risk Premia

As mentioned in the introduction an important part of the econometric literature on fx markets has analyzed the information efficiency. The semi–strong and strong information efficiency, i.e. the reflection in prices of public and public as well as private information beside the history of prices, is difficult to verify – if not even impossible. A simpler expectation hypothesis is usually supposed, mainly the UIP. Therewith weak information efficiency is tested, e.g. whether the expectation error could have been partly explained by lagged prices and lagged expectation errors. If this were successful, so the conjecture, historic information (lagged prices) had been used inefficiently in the formation of expectations about future spot prices. For example (Hansen & Hodrick, 1980) proposed the following econometric equation to test information efficiency:

$$\begin{aligned} \ln X_t - \ln F_{t-1,t} &= \alpha + \beta \left( \ln X_{t-1} - \ln F_{t-2,t-1} \right) + u_t, \\ \text{with } E_P^{t-1} \left( u_t \right) &= 0. \end{aligned}$$

If the UIP had to hold, i.e. if $E_P^{t-1} \left( \ln X_t - \ln F_{t-1,t} \right) = 0$, the estimated coefficients should be significantly close to zero. Some empirical studies found this in the data, some others claimed evidence for the contrary as well. However in other contributions neither the hypothesis nor the contrary appeared to be significantly sustained. The interpretation of coefficients $\alpha$, $\beta$ being different from zero has been twofold, either due to market inefficiency or due to a risk premium in $\ln X_t$. It has been shown in Proposition 93 why the presumed hypotheses might have been failed. Therefrom to conclude that markets operate inefficiently is not justified. Since even for risk neutral investors the UIP holds in a nominal economy only in special cases (see Lemma 82), it is wrong to identify $E_P^{t-1} \left( \ln X_t \right) - \ln F_{t-1,t}$ with a risk premium.[122] It is a premium caused by random inflation if agents were risk neutral.[123] For

[122] By definition the 'risk premium' measures the difference between the expected equilibrium returns with risk neutral and risk averse agents, respectively.

[123] The term 'risk premium' is correct either in a real economy or if agents perceive money as if it were a consumption good.

this 'risk premium' Hansen and Hodrick assumed an AR(1) process. This procedure is similar to an error correction model, in which the expectation error depends linear on past expectation errors, just that the term 'error' has another meaning. In an economy with money and risk averse investors this 'error' is a mixture of the real risk premium, the inflation risk premium, and a factor due to the change of numéraires from the real to the nominal currency. From Corollary 91 and Lemma 94 it follows

$$\begin{aligned} & E_P^{t-1}\left(\ln \hat{X}_t\right) - \ln \hat{X}_{t-1} - \frac{1}{365}\left(\hat{r}_{t-1}^d - \hat{r}_{t-1}^f\right) \\ \simeq\ & \frac{1}{365 \cdot 2}\left(\left\|\hat{\xi}_{t-1}^d\right\|^2 - \left\|\hat{\xi}_{t-1}^f\right\|^2\right). \end{aligned}$$

Let the last line define this collection of risk premia:

$$\begin{aligned} Y_{t-1,t} :\ =\ & \frac{1}{365 \cdot 2}\left(\underbrace{\left\|\xi_{t-1}^d\right\|^2 - \left\|\xi_{t-1}^f\right\|^2}_{\text{real risk premia}}\right. \\ & \left.\underbrace{+\left\|\sigma_{t-1}^{qd}\right\|^2 - \left\|\sigma_{t-1}^{qf}\right\|^2}_{\text{numéraire change to nominal prices}} \quad \underbrace{-\,\sigma_{t-1}^{qd}\xi_{t-1}^d + \sigma_{t-1}^{qf}\xi_{t-1}^f}_{\text{inflation risk primia}}\right). \end{aligned}$$

In the two previous sections the premia in $Y_{t-1,t}$ has been explicitly spelled out. It is called the fx–premium in the sequel. Although it contains unobservable quantities suitable substitutions were found. This subsection applies the fx–premium as an 'error' correction, i.e. the coefficients $\alpha$ and $\beta$ of an equation are estimated, which is similar to the one formulated by Hansen and Hodrick:

$$\ln \hat{X}_t - \ln \hat{X}_{t-1} = \alpha_0 + \frac{\alpha_2}{365}\left(\hat{r}_{t-1}^d - \hat{r}_{t-1}^f\right) + \beta Y_{t-2,t-1} + u_t.$$

From the theory one is able to formulate a conjecture about $\beta$. The equilibrium model has shown that prices of risk depend on real quantities. And the neutrality of money implies that inflation is caused by the excess growth of

money stocks over the GNP growth. Relatively to financial quantities those variables are inert. Therefore the conjecture is that parts of the fx–premium do not change too frequently compared to the exchange rate, because the fx–premium is determined by prices of risk and inflation. This would lead to a positive $\beta$ significantly different from zero.
This equation cannot be interpreted in a way of information efficiency. This estimation is interesting 'only' for reasons of comparison to previous empirical results and to see whether the fx–premium $Y$ could have explanatory power in this moving average process.
The constant $\alpha_0$ was nearly zero, not at all significant, and therefore excluded from further regressions. The first table displays the result of the regression based on daily observations over the entire horizon.[124]

| Variable | Coefficient | Std Error | Significancy |
|---|---|---|---|
| $Y_{t-2,t-1}$ | 0.0757 | 0.0221 | $\sim 0.0\%$ |
| $\hat{r}^d_{t-1} - \hat{r}^f_{t-1}$ | 0.3183 | 0.15 | 3.4% |
| $\hat{r}^d_{t-8} - \hat{r}^f_{t-8}$ | −0.3084 | 0.15 | 4.0% |

The explaining variables are the one day and eight days lagged over night interest rate differentials and the one day lagged daily fx–premium. The coefficients of the interest rate differentials are significant to the 5% level, that of the moving average is in fact highly significant. The centered $R^2$ is at a low of 0.86%. The regression is more successful on an aggregated level. The next estimate is based on weekly log–differences of the exchange rate. The over night interest rates have been substituted by the three months rates[125] and the daily by the weekly fx–premium.[126]

[124] The estimation has been carried out by the "AR!" procedure in RATS with the "HILU" option, which results in a minimal least square estimate by a grid search.
[125] Perhaps the three months rates perform better, because the over night interest rates capture more a narrow liquidity in the money market. This market could be very volatile and is not so relevant for the longer horizon of a week.
[126] The Box–Jenkins procedure in RATS has been applied to this regression, which estimates the coefficients by quasi–maximum–likelihood in conjunction with a Gauss–Newton algorithm.

| Variable | Coefficient | Std Error | Significancy |
|---|---|---|---|
| $Y_{t-8,t-1}$ | 0.6781 | 0.0164 | $\sim 0\%$ |
| $\hat{r}_{t-1}^{3M,d} - \hat{r}_{t-1}^{3M,f}$ | $-0.5634$ | 0.171 | $\sim 0\%$ |
| $\hat{r}_{t-8}^{3M,d} - \hat{r}_{t-8}^{3M,f}$ | 0.5568 | 0.171 | $\sim 0\%$ |

The centered $R^2$ is remarkably 52.5% high. In both regressions one observes an overshooting in the interest rate differential. The influence of the interest rate differentials is now in the sum negative, presumed to be positive like in the last regression. Moreover their common contribution to the explained part of the variance of the endogenous variable is very low. The interest rate differentials are only significant if at least two consecutive differentials are taken into account. Since the interest rate differentials show to have a trend and are therefore collinear, it makes sense to have a look at differences of second order:

| Variable | Coefficient | Std Error | Significancy |
|---|---|---|---|
| $Y_{t-8,t-1}$ | 0.6781 | 0.0164 | $\sim 0\%$ |
| $\left(\hat{r}_{t-1}^{3M,d} - \hat{r}_{t-1}^{3M,f}\right) - \left(\hat{r}_{t-8}^{3M,d} - \hat{r}_{t-8}^{3M,f}\right)$ | $-0.5601$ | 0.171 | $\sim 0\%$ |

The centered $R^2$ stayed almost the same. In the three regressions the lagged fx–premium explains most of the exchange rate changes. Since the fx–premium follows an AR(1) process, the log–exchange rate follows then an ARIMAX(1,1,1) process. The forecasting equation with figures taken from the last regression then sounds:

$$\begin{aligned} \ln \hat{X}_t &= \ln \hat{X}_{t-1} - 0.5601 \left[ \left( \hat{r}_{t-1}^{3M,d} - \hat{r}_{t-1}^{3M,f} \right) - \left( \hat{r}_{t-8}^{3M,d} - \hat{r}_{t-8}^{3M,f} \right) \right] \\ &\quad + 0.678\, u_{t-1} + u_t. \end{aligned}$$

The empirical findings sustain the theoretical results that a fx–premium exists which is relatively sluggish. It is possible to improve the explanatory power of regressions remarkably by considering an AR(1) process for the fx–premium.

## 3.6 Conclusions

The purpose of this chapter was to provide a consistent analysis of the exchange rate from four angles; an educational perspective; from the viewpoint of arbitrage pricing; within an equilibrium framework; and, empirically. The Uncovered Interest Parity (UIP), which has been a very popular hypothesis so far, was the peg to hang on those issues.

In the first part it has been shown by intuitive and mostly verbal arguments that the real exchange rate satisfies the Forward Expectation Hypothesis (FEH) if agents were risk neutral. In this case the exchange rate as well as interest rates follow deterministic processes which trivially implies the FEH. But if this artificial assumption does not hold or nominal numéraires are considered the FEH is satisfied only for purely deterministic economies. In any way, there are no reasons for the UIP to hold in a stochastic environment. In fact, the UIP was proved to be a non–generic property.

The next section embedded the exchange rate into an extended version of the Black–Scholes model. In continuous time finance it is a well known result that the exchange rate is equal to the ratio between the foreign and the domestic price deflators. Each of the deflators is the product of the pricing asset and the savings account belonging to the respective numéraire. That mirrors the function the exchange rate serves: to switch between two different valuation regimes consistently. But even in complete markets not both pricing assets are uniquely determined. Provided the domestic pricing asset is uniquely characterized in complete markets, the exchange rate determines the foreign pricing asset or vice versa. Within the triple of exchange rate and either prices of risk there are degrees of freedom equal to the number of exogenous risks plus one (for the drift and the diffusion of the exchange rate). Thus, only with redundancies independently priced in the two economies, which have without those redundancies already jointly complete asset markets, information can be gained for the trend and the diffusion of the exchange rate (or the missing pricing asset). For only one redundancy a correlation between the trend and the diffusion of the exchange rate has been established. Beside the interest rate differential the correlation involves also the excess returns of assets and their co–variation processes with the exchange rate. Exchange rate's growth rate shows to be linear in the excess returns of assets and the short rate differential, where the latter goes in one–to–one. If more than one redundancy occurs, also the diffusion of the fx–rate will be determined up to a certain number of degrees of freedom, which is equal to the dimension of exogenous risks minus the number of redundancies predominant in the market.

To show the non–viability of the UIP an arbitrary price of risk was presumed for each numéraire. It has been proved that the UIP imposes a non–generic restriction onto the prices of risk in their domain. Thus only special, non–generic economies satisfy the UIP.

The arbitrage pricing model was followed by an equilibrium model with a stochastic environment in continuous time. The arbitrage pricing left open the determination of prices of risk, the interest rates and the starting value of the exchange rate. These four processes and today's exchange rate completely characterize the law of motion of the exchange rate by arbitrage, while they themselves can only be learned in an equilibrium framework. The model is in its nature macro–economic but micro–founded. The investigation was limited to an extensive example economy with two numéraires, two representative agents with logarithmic von–Neumann–Morgenstern utility functions, two linear production technologies, and complete asset markets. In this example it turned out that the prices of risk were a stochastic convex combination of the two diffusions specific for each country, namely the diffusion of production technology and resources. The interest rates have four determining factors. They increase in the foreign and in the domestic deviation of individual discount factors from population growth rates, the growth rates of production, which is itself the sum of the growth rates belonging to the technology and the labour force, the growth rates of resources and the price of the price of risk. The coefficients of the growth rates are stochastic convex combinations, where the weights are either the export share, if foreign variables influence domestic quantities, or the share in supply, which either the production or the resources have. Today's exchange rate is equal to the ratio between total domestic expenditure spent on foreign consumption and total foreign expenditure spent on domestic consumption. Total expenditure contains today's price of future resources as well as the value of populations' total leisure time (priced by risk neutral expected discounted future prices and cash flows). This can be interpreted as an equilibrium on the currency market, in which all future exports and imports are completely hedged today. The results of that example lead also to the conjecture that the growth rate of the exchange rate reacts positively/negatively on domestic/foreign excess returns of assets. This is one of the key hypothesis of the empirical examination.

In the empirical section two reduced form equations were derived to test the conjectures drawn from the two preceding, theoretical models. The sample consists of 1087 daily observations of the DEM per US–$ exchange rate of the years 1993 to 1997, the corresponding interest rates, and stock index prices. The estimation of the first equation mirrored the linearity between the growth rate of the exchange rate on the one hand and the excess returns of assets and the short rate differential on the other hand. The results are only partly mirroring the arbitrage pricing model. Although some linearities were detected between the returns of the exchange rate and the assets, the dependency on the short rate differential is of opposite sign than supposed. Moreover, not all foreign assets and covariance terms between assets and the exchange rate play a significant role. The explanatory power increases but remains low if additional non–linearities were tested. For instance the returns of the Dow Jones Index are only in quadratic form significant. Moreover, some variables, e.g. the short rate differential, cause an overshooting of the exchange rate in that the influence of lagged variables is reversed by more recent variables.

The second equation had been more successful. The expected growth rate of the exchange rate can be split–up into the short rate differential and an unobservable fx–premium (which is not only a risk–premium). If one assumes a persistent fx–premium the equation can be re–written into an error correction model. This approach worked reasonable well for weakly data. The best result reaches an $R^2$ of about 50%, which is high for financial data. Therein only two factors explain the log–differences, namely the weekly differences of the three–month interest rate differential and the lagged forecasting error,

which is the fx–premium. Thus, there is some evidence for an integrated ARIMAX(1,1,1)–process for the exchange rate.
In the arbitrage pricing model there is not much room left for explaining the fx–rate. Of course the assumptions imposed onto the stochastic processes could be relaxed. For example, a simple assumption has been made to put Novikov's condition down to the exogenous drift and diffusion functionals. This assumption, which is not specific to fx–markets, should be weakened.
In the area of equilibrium models the dynamic optimization approach of (Karatzas *et al.*, 1990) in continuous time waits for an extension to several goods. But those general models are often not solvable in closed form. This implies the difficulty of a quantitative interpretation and should be accompanied by developing more refined examples like the one proposed in the text.
In applied econometrics there are two competing methodologies predominant. Either the reduced form equations are derived from theoretical models or one tries best to fit the data by ad–hoc econometric models and – if ever – builds the explaining theoretical model afterwards. Although both ways are justified this essay followed the first approach. The latter bears the risk that one looks for seeming laws in the data which can actually not be valid in any model, for instance the UIP. Especially for fx–markets the ad–hoc econometric method had been very unsuccessful [see (Frankel & Rose, 1995)]. Moreover, a model helps to identify factors, even though in the estimation one does not follow the functional form of the model.
To be more precise, the example economy suggests for the prices of risk a convex combination of the diffusions of resources and of technological change. The fx–premium is equal to one half of the difference between the Euclidean norms of prices of risk. This observation results in the difference of instantaneous volatilities of country specific processes. Therefore, *volatilities of fundamentals*, for example the variance of production, may have an important influence on the fx–premium. Volatilities of fundamentals were not yet regarded in econometric models for fx–markets.

## 3.7 Appendix, Proofs of Section 3.3.2

### 3.7.1 SDE of the exchange rate

**Lemma 99** $\frac{X_t^i}{X_0^i}\frac{b_t^t}{b_t^{\neg i}}$ *starts in* 1 *and is a positive martingale with respect to the measure* $Q^i$.

**Proof.** $X_t^i/b_t^{\neg i}$ is the locally riskless bank account of currency $\neg i$ measured in currency $i$. Thence it is a risky asset, which has to be a martingale under $Q^i$ after discounting it by the locally riskless bank account of country $i$.
By arbitrage the exchange rate must be positive, where $X_0^i$ is the $\mathcal{F}_0$–measurable and normalizing constant. ■
A random process with these properties induce a change of measure, here from $Q^i$ to $Q^{\neg i}$.
In the appendices the formal computations make use of the following symbolic re–definitions: $d$ stands for 1 and $f$ for $-1$ **only** if either $X$ or $F$ are concerned. This is because $X^f = \frac{1}{X}$ and $X^d = X$. These abbreviations imply, for example, that $\ln X^i = i \ln X$. Note, that this applies to $X$ and $F$ only!

**Lemma 100 (SDE of the exchange rate)** *In complete markets the flow of the exchange rate is given by*

$$dX/X = \left(r^d - r^f - \xi^d\left(\xi^f - \xi^d\right)\right)dt + \left(\xi^f - \xi^d\right)dW.$$

**Proof.** $\frac{X_t^i}{X_0^i}\frac{b_t^t}{b_t^{\neg i}}$ is a Martingale with respect to the measure $Q^i$. Moreover the exchange rate is strictly positive by a simple arbitrage argument. If the diffusion of $X$ is $\sigma^X X$, then it is $\sigma^X$ for $\ln X$. The diffusion is not affected by a change of measure. This implies two representation for $X$ :

$$X_t^i \overset{Q^i}{=} X_0^i \exp\left(\int_0^t r_u^i - r_u^{\neg i} du\right)\mathcal{E}_t\left(\left(i\sigma^X\right)\cdot \tilde{W}^i\right),$$

where $\tilde{W}^i$ are the Brownian motions due to the change from $P$ to $Q^i$ :

$$d\tilde{W}_t^i \overset{P}{=} dW_t - \xi_t^i dt$$

and $\mathcal{E}_t$ denotes the exponential up to time $t$:

$$\mathcal{E}_t\left(z \bullet y\right) = \exp\left(\int_0^t z_u dy_u - \frac{1}{2}\int_0^t \left\|z_u\right\|^2 du\right).$$

Now undoing the change of measure by the prices of risk $\xi^i$ yields:

$$d\ln X^i = \left(r^i - r^{\neg i} - \frac{1}{2}\left\|\sigma^X\right\|^2 - i\sigma^X\xi^i\right)dt + i\sigma^X dW.$$

Say $\hat{H}$ are $m$ risky assets with diffusion $\hat{\sigma}$ in the foreign currency, which together with the locally riskless bank account complete the foreign financial market. By the definition of the foreign price of risk the assets solve the SDE:

$$d\hat{H} = \left(r^f\hat{H} - \hat{\sigma}\xi^f\right)dt + \hat{\sigma}dW.$$

Whereas in the domestic currency the assets $\hat{H}X$ have to solve by Itô's rule the SDE:

$$\begin{aligned} d\left(\hat{H}X\right) &= \hat{H}dX + Xd\hat{H} + d\left\langle \hat{H}, X\right\rangle \\ &= X\left(\hat{H}\mu^X + r^f\hat{H} - \hat{\sigma}\xi^f + \hat{\sigma}\left(\sigma^X\right)^T\right)dt + X\left(\hat{H}\sigma^X + \hat{\sigma}\right)dW. \end{aligned}$$

By the definition of the domestic price of risk these SDEs must coincide. Then

$$d\left(\hat{H}X\right) = \left(\hat{H}Xr^d - X\left(\hat{H}\sigma^X + \hat{\sigma}\right)\xi^d\right)dt + X\left(\hat{H}\sigma^X + \hat{\sigma}\right)dW$$

implies the equation

$$\hat{H}r^d - \left(\hat{H}\sigma^X + \hat{\sigma}\right)\xi^d = \hat{H}\mu^X + r^f\hat{H} - \hat{\sigma}\xi^f + \hat{\sigma}\left(\sigma^X\right)^T.$$

Inserting the result for $\mu^X$ in this equation and solving for the diffusion of the exchange rate reads:

$$\begin{aligned}
\hat{H}r^d - \left(\hat{H}\sigma^X + \hat{\sigma}\right)\xi^d &= \hat{H}\left(r^d - r^f - \sigma^X\xi^d\right) + r^f\hat{H} - \hat{\sigma}\xi^f + \hat{\sigma}\left(\sigma^X\right)^T \Leftrightarrow \\
\hat{\sigma}\left(\sigma^X\right)^T &= \hat{\sigma}\left(\xi^f - \xi^d\right).
\end{aligned}$$

And because markets are complete, i.e. $rank\left(\hat{\sigma}\right) = m$, the solution for the diffusion of the exchange rate is $\left(\sigma^X\right)^T = \xi^f - \xi^d$. This implies the following stochastic differential equation for the log–exchange rate:

$$\begin{aligned}
d\ln X &= \left(r^d - r^f - \frac{1}{2}\left\|\xi^f - \xi^d\right\|^2 - \left(\xi^f - \xi^d\right)\xi^d\right)dt + \left(\xi^f - \xi^d\right)dW \\
&= \left(r^d - r^f + \frac{1}{2}\left\|\xi^d\right\|^2 - \frac{1}{2}\left\|\xi^f\right\|^2\right)dt + \left(\xi^f - \xi^d\right)dW.
\end{aligned}$$

Applying Itô's rule to $\ln X$, which reads $d\ln X = dX/X - d\left\langle X, X\right\rangle/2$, one arrives at the stochastic differential equation stated in the lemma. ■

### 3.7.2 Proposition (FX–Flow)

The proof is divided in three steps.

1. It is checked that the diffusion matrix of assets is regular.
2. Here some preparation results for step four are given concerning the inverse of a small matrix adjustment.
3. The prices of risk $\xi^i$ for both currencies are derived.
4. The change of measure in the SDE of the foreign exchange rate is reversed and the equation: $\mu^X = r^d - r^f + \sigma^X\xi^d$, is solved for the trend $\mu^X$.

#### Regularity of the Vola–Matrix

**Lemma 101 (Regular Diffusion-Matrix)** *If domestic financial markets are complete and (FX–Spanning) is assumed, then the diffusion matrix of* *assets which span the risks of domestic financial markets is regular, i.e.*

$$rank \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} = m.$$

*Moreover*

$$-\sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} \begin{pmatrix} 0 \\ H^f \end{pmatrix} \neq 1.$$

**Proof.** Since the scalar $X$ is strictly positive it does not effect the rank of the diffusion matrix of assets constituting complete domestic financial markets, i.e.

$$rank \begin{pmatrix} \sigma^d \\ X\sigma^f + XH^f\sigma^X \end{pmatrix} = rank \begin{pmatrix} \sigma^d \\ \sigma^f + H^f\sigma^X \end{pmatrix}.$$

By complete domestic asset markets the former matrix has rank $m$. Then $rank \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} \geq m-1$ because $\begin{pmatrix} 0 \\ H^f \end{pmatrix} \sigma^X$ is a rank adjustment at most by one degree. Suppose that the rank is $m-1$. Then a portfolio $\theta^0 \neq 0$ exists such that $\theta^0 \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} = 0$. This portfolio divided by a constant would also generate the volatility of the foreign bond in domestic currency:

$$\theta^0 \begin{pmatrix} \sigma^d \\ \sigma^f + H^f\sigma^X \end{pmatrix} \Big/ \theta^0 \begin{pmatrix} 0 \\ H^f \end{pmatrix} = \sigma^X.$$

Note that a zero nominator would be a contradiction to complete domestic asset markets. By (FX–Spanning) there also is a portfolio $\theta^X$ such that:

$$\theta^X \begin{pmatrix} \sigma_t^d \\ \sigma^f \end{pmatrix} = \sigma^X.$$

But this means that there is a portfolio $\hat{\theta}$, defined by

$$\hat{\theta} = \theta^X \left( \theta^0 \begin{pmatrix} 0 \\ H^f \end{pmatrix} \right) - \theta^0 \left( 1 + \theta^X \begin{pmatrix} 0 \\ H^f \end{pmatrix} \right),$$

in the null–space of the diffusion matrix

$$\hat{\theta} \begin{pmatrix} \sigma^d \\ \sigma^f + H^f \sigma^X \end{pmatrix} = 0.$$

And since $\theta^X \nparallel \theta^0$ they cannot sum up to zero, i.e. $\hat{\theta} \neq 0$. This is a contradiction to complete domestic asset markets.
The second claim still needs to be proven. Domestic asset markets are complete if and only if for any vector $\hat{\theta} \neq 0$ it is not true that

$$\begin{aligned}
\begin{pmatrix} \sigma^d \\ \sigma^f + H^f \sigma^X \end{pmatrix} \hat{\theta} &= 0 \Leftrightarrow \\
-\begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ H^f \end{pmatrix} \sigma^X \hat{\theta} &= \hat{\theta} \Rightarrow \\
-\sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ H^f \end{pmatrix} \sigma^X \hat{\theta} &= \sigma^X \hat{\theta} \text{ and } \sigma^X \hat{\theta} \neq 0 \Rightarrow \\
-\sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ H^f \end{pmatrix} &= 1.
\end{aligned}$$

However domestic financial markets are complete and thereby the equation in the last row can never hold. ■

**Some preliminary Linear Algebra**

The next step of the proof needs a transformation of a certain vector. This minor problem of linear algebra is considered now.

**Lemma 102 (Inverse of small rank adjustment)** *Consider a regular matrix $A \in \mathbb{R}^m \otimes \mathbb{R}^m$ and two vectors $x, y \in \mathbb{R}^m$ such that $-x^T A^{-1} y \neq 1$. Then the inverse of a 'small rank adjustment' $A + yx^T$ can be transformed* *in the following way:*

$$\left(A+yx^T\right)^{-1}=A^{-1}-\frac{1}{1+x^TA^{-1}y}A^{-1}yx^TA^{-1}.$$

The formula can be found in [(Horn & Johnson, 1985) p. 19].[127] A corollary of this lemma is the following:

**Corollary 103** *Let $A, x, y$ satisfy the requirements of the last lemma. Then*

$$\left(1-x^T\left(A+yx^T\right)^{-1}y\right)^{-1}x^T\left(A+yx^T\right)^{-1}=x^TA^{-1}.$$

**Proof.** The following vector of this expression transforms into

$$x^T\left(A+yx^T\right)^{-1}=x^TA^{-1}\left(1-\frac{x^TA^{-1}y}{1+x^TA^{-1}y}\right).$$

Inserting this result into the denominator yields

$$1-x^TA^{-1}y\left(1-\frac{x^TA^{-1}y}{1+x^TA^{-1}y}\right)=\left(1-\frac{x^TA^{-1}y}{1+x^TA^{-1}y}\right),$$

which proves the simplification stated in the corollary. ■

**Corollary 104** *Let $A, x, y$ satisfy the requirements of the last lemma. Then*

$$\left(A+yx^T\right)^{-1}\left(I+yx^TA^{-1}\right)=A^{-1}.$$

**Proof.** By multiplying both sides with $A+yx^T$. ■

**Change of Measure**

**Lemma 105 (Prices of Risk)** *The prices of risk $-\xi^i$, which are minus the*

[127] The proof is simply done by multiplying each side with $\left(A+yx^T\right)$. The more general case, in which $yx^T$ might be a 'true' rank adjustment for an arbitrary square matrix $A$ and two arbitrary vectors $x, y$ of matching length is treated in [(Campbell & Meyer, 1979), Theorem 3.1.3 p. 47].

*diffusions of the pricing density processes, are*

$$\xi^d = -\begin{pmatrix} \sigma^d \\ H^f\sigma^X + \sigma^f \end{pmatrix}^{-1} \left[\begin{pmatrix} \mu^d \\ H^f\mu^X + \mu^f + \sigma^f\sigma^X \end{pmatrix} - r^d \begin{pmatrix} H^d \\ H^f \end{pmatrix}\right]$$

$$\xi^f = -\begin{pmatrix} \sigma^d - H^d\sigma^X \\ \sigma^f \end{pmatrix}^{-1} \cdot \left[\begin{pmatrix} \mu^d - H^d\mu^X - \sigma^d\sigma^X + H^d \left\|\sigma^X\right\|^2 \\ \mu^f \end{pmatrix} - r^f \begin{pmatrix} H^d \\ H^f \end{pmatrix}\right].$$

**Proof.** By the no–arbitrage condition $H^i b^i$, $H^{\neg i} X^i b^i$ and $X^i \frac{b^i}{b^{\neg i}}$ are $Q^i$–martingales. As $H^d$ and $H^f X$ generate a complete asset structure in the domestic financial market they induce the change of measure form $P$ to $Q^d$. First, one computes $H^f X$ by Itô's rule:

$$\begin{aligned} d\left(H^f X\right) &= H^f dX + X dH^f + d\left\langle H^f, X\right\rangle \\ &= X\left(H^f\mu^X + \mu^f + \sigma^f\sigma^X\right) dt \\ &\quad + X\left(H^f\sigma^X + \sigma^f\right) dW. \end{aligned}$$

With the use of Girsanov's Theorem the Brownian motion under $Q$ has to be:

$$d\tilde{W}^d = dW + \begin{pmatrix} \sigma^d \\ XH^f\sigma^X + X\sigma^f \end{pmatrix}^{-1} \cdot \left[\begin{pmatrix} \mu^d \\ XH^f\mu^X + X\mu^f + X\sigma^f\sigma^X \end{pmatrix} - r^d \begin{pmatrix} H^d \\ H^f X \end{pmatrix}\right] dt.$$

Hence the drifts of $H^d$ and $H^f X$ are replaced by $r^d H^d$ and $r^d H^f X$ as required by the no–arbitrage implication with respect to the risk neutral measure $Q^d$. In the proof of Section 3.7.1 it was shown that the expression preceding $dt$ is minus the domestic price of risk. The exchange rate cancels down in the bottom row.

The same calculation can be done for the foreign financial market. The domestic assets in foreign currency follow by Itô's rule the stochastic differential equation:

$$\begin{aligned} d\left(H^d X^{-1}\right) &= -H^d X^{-2} dX + X^{-1} dH^d - X^{-2} d\left\langle H^d, X\right\rangle + H^d X^{-3} d\left\langle X\right\rangle \\ &= X^{-1}\left(\mu^d - H^d \mu^X - \sigma^d \sigma^X + H^d \left\|\sigma^X\right\|^2\right) dt \\ &\quad + X^{-1}\left(\sigma^d - H^d \sigma^X\right) dW. \end{aligned}$$

Again this implies the change of measure:

$$\begin{aligned} d\tilde{W}^f &= dW + \begin{pmatrix} X^{-1}\sigma^d - H^d X^{-1}\left(\sigma^X\right)^T \\ \sigma^f \end{pmatrix}^{-1} \cdot \\ &\quad \left[\begin{pmatrix} X^{-1}\mu^d - H^d X^{-1}\mu^X - X^{-1}\sigma^d\sigma^X + H^d X^{-1}\left\|\sigma^X\right\|^2 \\ \mu^f \end{pmatrix}\right. \\ &\quad \left. - r^f \begin{pmatrix} H^d X^{-1} \\ H^f \end{pmatrix}\right] dt. \end{aligned}$$

$X^{-1}$ cancels down now in the upper row, which results then in the foreign price of risk. ■

**FX–Trend**

**Lemma 106** **($\mu^X/\sigma^X$–Arbitrage Relation)** *If there are equivalent martingale measures $Q^i$ the following relation between trend and diffusion of the exchange rate must hold:*

$$\mu^X = \left(r^d - r^f\right) + \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} \mu^d - r^d H^d \\ \mu^f - r^f H^f + \sigma^f \sigma^X \end{pmatrix}.$$

**Proof.** In the previous Lemma 100 of Section 3.7.1 it has been shown that the trend of the exchange rate satisfies

$$\mu^X = \left(r^d - r^f\right) - \xi^d \sigma^X.$$

In this equation the substitution $\mu^X = \bar{\mu}^X + \left(r^d - r^f\right)$ is employed and the result for $\xi^d$ from Lemma 105 is inserted. Then it reads:

$$\bar{\mu}^X = \sigma^X \begin{pmatrix} \sigma^d \\ H^f \sigma^X + \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} \mu^d - r^d H^d \\ \mu^f - r^f H^f + H^f \bar{\mu}^X + \sigma^f \sigma^X \end{pmatrix}.$$

Solving this equation for $\bar{\mu}^X$ results in

$$\begin{aligned} \bar{\mu}^X &= \left(1 - \sigma^X \begin{pmatrix} \sigma^d \\ H^f \sigma^X + \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ H^f \mu^X \end{pmatrix}\right)^{-1} \cdot \\ &\quad \sigma^X \begin{pmatrix} \sigma^d \\ H^f \sigma^X + \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} \mu^d - r^d H^d \\ \mu^f + \sigma^f \sigma^X - r^f H^f \end{pmatrix}. \end{aligned}$$

By identifying

$$A = \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}, x = \left(\sigma^X\right)^T, y = \begin{pmatrix} \mathbf{0} \\ H^f \mu^X \end{pmatrix}$$

the Corollary 103 can be applied to the scalar and the first vector of the left hand side of the previous equation. In Lemma 101 all the prerequisites of the corollary have been showed . Undoing the change of variable in the equation yields the relation between trend and diffusion of the exchange rate stated in the lemma. ■

**Remark 10** *One might think that only $n^f - 1$ risky foreign assets plus the foreign bond are enough securities to change the measure in the domestic currency, because the exchange rate makes the bond also risky. Then one would arrive at $n^f + n^d + 1 = m + 1$ 'spanning' securities. Indeed this is enough to change the measure but it does not suffice to derive the arbitrage relation for the trend. In this case one risky asset is missing to apply the linear pricing rule. The determining equation for $\mu^X$ would read:*

$$\begin{aligned} \bar{\mu}^X &= \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f_{n^f-1} + H^f_{-1} \sigma^X \\ B^f \sigma^X \end{pmatrix}^{-1} \begin{pmatrix} \mu^d - r^d H^d \\ H^f_{-1} \bar{\mu}^X + \mu^f_{-1} + \sigma^f_{-1} \sigma^X - r^f H^f_{-1} \\ B^f \bar{\mu}^X + r^f B^f - r^f B^f \end{pmatrix} \\ &= \bar{\mu}^X. \end{aligned}$$

*Therefore at least $n^f$ risky foreign assets are necessary to calculate $\mu^X$. All assets together, foreign and domestic, must span the entire $\mathcal{L}^2(P \times \lambda)$.*

### 3.7.3 Proof of Proposition Generalized PPP

**Proposition 107 (Generalized PPP)** *Suppose there are two ways to achieve complete domestic asset markets. Either by $n^d + \bar{n}$ domestic and $n^f$ foreign assets or by $n^d$ domestic and $n^f + \bar{n}$ foreign assets. The two possibilities differ only in the substitution of $\bar{n}$ foreign assets by the same number of domestic assets, whereas all other assets remain the same. The changing assets are denoted by a bar on top. For both completions FX–spanning shall hold.*
*Under these assumptions markets are free of arbitrage opportunities if and only if the diffusion of the exchange rate satisfies the following equation up to $m - \bar{n}$ degrees of freedom:*

$$\begin{aligned}
\left(\sigma^X\right)^T \quad = \quad & \begin{pmatrix} \bar{\sigma}^d \\ \sigma^d \end{pmatrix}^+ \begin{pmatrix} \bar{\mu}^d - \bar{H}^d r^d \\ \mu^d - H^d r^d \end{pmatrix} + y^d \\
& - \begin{pmatrix} \bar{\sigma}^f \\ \sigma^f \end{pmatrix}^+ \begin{pmatrix} \bar{\mu}^f - \bar{H}^f r^f \\ \mu^f - H^f r^f \end{pmatrix} + y^f \quad P \times \lambda - a.s. \\
& \text{for some arbitrary } y^i \in \mathcal{N} \begin{pmatrix} \bar{\sigma}^i \\ \sigma^i \end{pmatrix},
\end{aligned}$$

*where the superscript $+$ stands for the Moore–Penrose inverse.*

This proof extensively uses the Moore–Penrose inverse and orthogonal projections, see Section 3.9.

**Proof.** The replaced and the replacing assets respectively are denoted by a bar on top. Then the change of measure in the domestic currency can be done in two ways:

$$\begin{aligned}\xi_t &= -\begin{pmatrix}\sigma^d\\ \bar\sigma^f+\bar H^f\sigma^X\\ \sigma^f+H^f\sigma^X\end{pmatrix}^{-1}\left[\begin{pmatrix}\mu^d\\ \bar\mu^f+\bar\sigma^f\sigma^X+\mu^X\bar H^f\\ \mu^f+\sigma^f\sigma^X+\mu^X H^f\end{pmatrix}-r^d\begin{pmatrix}H^d\\ \bar H^f\\ H^f\end{pmatrix}\right]\\ &= -\begin{pmatrix}\sigma^d\\ \bar\sigma^d\\ \sigma^f+H^f\sigma^X\end{pmatrix}^{-1}\left[\begin{pmatrix}\mu^d\\ \bar\mu^d\\ \mu^f+\sigma^f\sigma^X+\mu^X H^f\end{pmatrix}-r^d\begin{pmatrix}H^d\\ \bar H^d\\ H^f\end{pmatrix}\right].\end{aligned}$$

By multiplying with the volatility matrix of the second row this equation is equivalent to:

$$\bar\sigma^d\begin{pmatrix}\sigma^d\\ \bar\sigma^f+\bar H^f\sigma^X\\ \sigma^f+H^f\sigma^X\end{pmatrix}^{-1}\left[\begin{pmatrix}\mu^d\\ \bar\mu^f+\bar\sigma^f\sigma^X+\bar H^f\mu^X\\ \mu^f+\sigma^f\sigma^X+H^f\mu^X\end{pmatrix}-r^d\begin{pmatrix}H^d\\ \bar H^f\\ H^f\end{pmatrix}\right]$$
$$=\bar\mu^d-r^d\bar H^d.$$

To shorten the notation the excess return is abbreviated:

$$\begin{pmatrix}x^d\\ \bar x^d\\ \bar x^f\\ x^f\end{pmatrix}=\begin{pmatrix}\mu^d-r^dH^d\\ \bar\mu^d-r^d\bar H^d\\ \bar\mu^f-r^f\bar H^f\\ \mu^f-r^fH^f\end{pmatrix}.$$

Inserting the solution for $\mu^X$ derived in Lemma 106 of the last Section 3.7.2 the expression in []–brackets simplifies by factoring out a vector to:

$$\left[I+\begin{pmatrix}\mathbf{0}\\ \bar H^f\\ H^f\end{pmatrix}\sigma^X\begin{pmatrix}\sigma^d\\ \bar\sigma^f\\ \sigma^f\end{pmatrix}^{-1}\right]\begin{pmatrix}x^d\\ \bar x^d+\bar\sigma^f\sigma^X\\ x^f+\sigma^f\sigma^X\end{pmatrix}.$$

By the last Corollary 104 in Section 3.7.2 the multiplication of the inverse to the right of the matrix in []–brackets reads:

$$\left(\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}+\begin{pmatrix}\mathbf{0}\\ \bar{H}^f\\ H^f\end{pmatrix}\sigma^X\right)^{-1}\left[I+\begin{pmatrix}\mathbf{0}\\ \bar{H}^f\\ H^f\end{pmatrix}\sigma^X\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}^{-1}\right]=\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}^{-1}.$$

By those two steps one arrives at the equation

$$\bar{\sigma}^d\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}^{-1}\left(\begin{pmatrix}x^d\\ \bar{x}^f\\ x^f\end{pmatrix}+\begin{pmatrix}0\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}\left(\sigma^X\right)^T\right)=\bar{x}^d.$$

Now the equation transforms for some $y_1\in\mathcal{N}\left(\bar{\sigma}^d\right)$ into:

$$\begin{pmatrix}0\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}\left(\sigma^X\right)^T=\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}\left[\bar{\sigma}^{d+}\bar{x}^d+y_1\right]-\begin{pmatrix}x^d\\ \bar{x}^f\\ x^f\end{pmatrix}.$$

The first rows restrict the choice for $y_1$ :

$$\begin{aligned}0 &= \sigma^d\left[\bar{\sigma}^{d+}\bar{x}^d+y_1\right]-x^d\Rightarrow\\ y_1 &= \left(I-\bar{\sigma}^{d+}\bar{\sigma}^d\right)\left(-\bar{\sigma}^{d+}\bar{x}^d+\sigma^{d+}x^d+y_2\right),\end{aligned}$$

for a $y_2\in\mathcal{N}\left(\sigma^d\right)$, which implies

$$y_1=\left(I-\bar{\sigma}^{d+}\bar{\sigma}^d\right)\sigma^{d+}x^d+y_3$$

where $y_3=\left(I-\bar{\sigma}^{d+}\bar{\sigma}^d\right)\left(I-\sigma^{d+}\sigma^d\right)\theta_3$ for a $\theta_3\in\mathbb{R}^m$.
One has to argue now that a solution for $y_1$ exists. This claim is proven by contradiction: Supposing a solution does not exist would violate the no–arbitrage requirement. The derived solution for $y_1$ guarantees that the new domestic assets do not introduce arbitrage opportunities relatively to the existing domestic assets. Inserting the solution back into the restriction for

$y_1$ results in:

$$\begin{aligned} x^d - \sigma^d \left[ \left( I - \bar{\sigma}^{d+} \bar{\sigma}^d \right) \sigma^{d+} x^d + y_3 \right] &= \\ \sigma^d \bar{\sigma}^{d+} \bar{\sigma}^d \sigma^{d+} x^d &= \sigma^d \bar{\sigma}^{d+} \bar{x}^d . \end{aligned}$$

It states that the excess returns of old and new domestic assets, $x^d$ and $\bar{x}^d$ respectively, must coincide in the space which both assets span: $\mathcal{R}\left(\left(\sigma^d\right)^T\right)$ and $\mathcal{R}\left(\left(\bar{\sigma}^d\right)^T\right)$, respectively. Multiplying the matrices in the last row onto the right hand side of assets' diffusions $\sigma^d$ and $\bar{\sigma}^d$, which results in $\sigma^d \bar{\sigma}^{d+} \bar{\sigma}^d$ $\sigma^{d+} \sigma^d dW$ and $\sigma^d \bar{\sigma}^{d+} \bar{\sigma}^d dW$, this point gets clear: The terms $\sigma^{d+} \sigma^d$ and $\bar{\sigma}^{d+} \bar{\sigma}^d$ are the orthogonal projections of risks onto the space which old and new assets span respectively. The spanned risks of the old assets are again projected on the space the new assets span. By the last pre–multiplication with $\sigma^d$ both projected risks are compared on the same space while eliminating orthogonal parts of $\mathcal{R}\left(\left(\sigma^d\right)^T\right)$. In this space the remaining risks coincide and so by arbitrage the returns have to be equal if transformed in the same way. Although this is only a orthogonal projection and not a complete spanning of the new domestic assets (complete spanning would lead to incomplete markets), this is indeed a no–arbitrage condition because the projection residuals can be hedged away by the other assets in complete markets. This proves the existence of a solution for $y_1$ by the no–arbitrage condition.
Finally the equation solves for the diffusion of the exchange rate:

$$\begin{aligned} \left(\sigma^X\right)^T &= \bar{\sigma}^{d+} \bar{x}^d + \left( I - \bar{\sigma}^{d+} \bar{\sigma}^d \right) \sigma^{d+} x^d + y_3 \\ &\quad + \begin{pmatrix} \bar{\sigma}^f \\ \sigma^f \end{pmatrix}^+ \begin{pmatrix} \bar{x}^f \\ x^f \end{pmatrix} + y_4 \\ &\quad \text{for some } y_4 \in \mathcal{N} \begin{pmatrix} \bar{\sigma}^f \\ \sigma^f \end{pmatrix} . \end{aligned}$$

It needs to be shown the existence of some vectors $\theta_3, \theta_5 \in \mathbb{R}^m$ such that (while ignoring the domestic country's index):

$$\bar{\sigma}^+\bar{x} + \left(I - \bar{\sigma}^+\bar{\sigma}\right)\sigma^+ x + \left(I - \bar{\sigma}^+\bar{\sigma}\right)\left(I - \sigma^+\sigma\right)\theta_3 = \begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}^+ \begin{pmatrix}\bar{x}\\ x\end{pmatrix} + \left(I - \begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}^+ \begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}\right)\theta_5.$$

By arbitrage there is a price of risk $\xi^d$ which solves:

$$\begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}\xi = \begin{pmatrix}\bar{x}\\ x\end{pmatrix}.$$

This implies

$$\begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}\begin{pmatrix}\bar{\sigma}\\ \sigma\end{pmatrix}^+ \begin{pmatrix}\bar{x}\\ x\end{pmatrix} = \begin{pmatrix}\bar{x}\\ x\end{pmatrix} \text{ and}$$

$$\bar{\sigma}\bar{\sigma}^+\bar{x} = \bar{x} \text{ as well as } \sigma\sigma^+ x = x.$$

Multiplying the considered equation by the diffusion matrix $\left((\bar{\sigma})^T \ (\sigma)^T\right)^T$ from the right yields the following equation for the bottom row:

$$\sigma\bar{\sigma}^+\bar{x} - \sigma\bar{\sigma}^+\bar{\sigma}\sigma^+ x - \sigma\bar{\sigma}^+\bar{\sigma}\left(I - \sigma^+\sigma\right)\theta_3 = 0,$$

whereas in the upper row all terms cancel down. Now the definition of the price of risk implies:

$$\sigma\bar{\sigma}^+\bar{\sigma}\left(I - \sigma^+\sigma\right)\xi - \sigma\bar{\sigma}^+\bar{\sigma}\left(I - \sigma^+\sigma\right)\theta_3 = 0.$$

Thence $\theta_3 = \xi$ is a solution. Shown the equality in the space of $\left((\bar{\sigma})^T \ (\sigma)^T\right)^T$ it must be shown now in the null–space. But this obviously holds for

$$\theta_5 = \bar{\sigma}^+\bar{x} + \left(I - \bar{\sigma}^+\bar{\sigma}\right)\sigma^+ x + \left(I - \bar{\sigma}^+\bar{\sigma}\right)\left(I - \sigma^+\sigma\right)\theta_3.$$

The last step of the proof shows that arbitrary $\theta_3, y_4$ span the same space as arbitrary $\theta_5, y_4$. This holds if the following condition is satisfied

$$\forall\theta_5\ \exists\theta_3,\theta_4 \text{ and } \forall\theta_3\ \exists\theta_4,\theta_5 \text{ in } \mathbb{R}^m:$$

$$\left(I-\bar{\sigma}^{d+}\bar{\sigma}^{d}\right)\left(I-\sigma^{d+}\sigma^{d}\right)\theta_3 =$$

$$\left(I-\begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}^+\begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}\right)\theta_5 - \overbrace{\left(I-\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}^+\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}\right)\theta_4}^{y_4=}.$$

The condition holds in $\mathcal{R}\left(\bar{\sigma}^T\right)$. It remains to prove it in $\mathcal{N}\left(\bar{\sigma}\right)$. Note that $\mathcal{N}\left(\begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}\right) \subset \mathcal{N}\left(\bar{\sigma}^d\right)$ and that $\mathcal{R}\left(\begin{pmatrix}\sigma^d\\ \bar{\sigma}^f\\ \sigma^f\end{pmatrix}^T\right) = \mathbb{R}^m$ implies $\mathcal{N}\left(\sigma^d\right) + \mathcal{N}\left(\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}\right) = \mathbb{R}^m$. Thus, $\forall\theta_5,\theta_3\ \exists\theta_4$ in $\mathbb{R}^m$ :

$$\left(I-\begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}^+\begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}\right)\theta_5 =$$

$$\left(I-\sigma^{d+}\sigma^{d}\right)\theta_3 + \left(I-\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}^+\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}\right)\theta_4.$$

Bringing all equations together yields the assertion of the proposition:

$$\begin{aligned}\left(\sigma^X\right)^T \quad = \quad & \begin{pmatrix}\bar{\sigma}^d\\ \sigma^d\end{pmatrix}^+\begin{pmatrix}\bar{x}^d\\ x^d\end{pmatrix} + y^d \\ & -\begin{pmatrix}\bar{\sigma}^f\\ \sigma^f\end{pmatrix}^+\begin{pmatrix}\bar{x}^f\\ x^f\end{pmatrix} + y^f \\ & \text{for some } y^i \in \mathcal{N}\begin{pmatrix}\bar{\sigma}^i\\ \sigma^i\end{pmatrix}.\end{aligned}$$

■

### 3.7.4 Proof of Non–Viability of the UIP

**Proposition 108 (Non–Viability of the UIP)** *Suppose the short rates have continuous, bounded sample paths. Assume moreover that the two economies are specified by arbitrary prices of risk* $-\xi^i : \mathbb{T}\times\Omega \to \mathbb{R}^m$ *in* $\mathcal{L}^2(P\times\lambda)^m$ *implying the existence of suitable pricing assets* $z^i$ *and martingale measures* $Q^i$*. Then the UIP, which is*

$$\forall t,\tau \in \mathbb{T}, t \le \tau : E_P^t \triangle_{t,\tau} \ln X := E_P^t \int_t^\tau d\ln X_u = \ln B_{t,\tau}^f - \ln B_{t,\tau}^d,$$

*holds only for economies of Lebesgue–measure zero in the range of* $\xi^i$*. By assuming the UIP the dimension of the range shrinks by one.*

**Proof.** With respect to the risk–neutral measures $Q^i$ and the corresponding prices of risk $\xi^i$ the price of a zero coupon bond $B_{t,\tau}^i$ in $t$, which matures in $\tau$ with a secure final payment of one currency unit, sounds:

$$\begin{aligned} B_{t,\tau}^i &= E_{Q^i}^t \exp\left\{-\int_t^\tau r_u^i du\right\} \\ &= E_{Q^i}^t \exp\left\{-(\tau-t)\, r_t^i - \int_t^\tau \int_t^u dr_v^i du\right\} \Leftrightarrow \end{aligned}$$

$$\begin{aligned} & B_{t,\tau}^i \exp\left\{(\tau-t)\, r_t^i\right\} \\ =\; & E_P^t \exp\left\{-\int_t^\tau \int_t^u \mu_v^{ri} + \sigma_v^{ri}\xi_v^i dv du - \int_t^\tau \int_t^u \sigma_v^{ri} dW_v du\right\} \\ =\; & E_P^t \exp\left\{-\int_t^\tau \int_t^u \mu_v^{ri} + \sigma_v^{ri}\xi_v^i dv du - \int_t^\tau \int_u^\tau \sigma_u^{ri} dv dW_u\right\} \\ =\; & E_P^t \exp\left\{-\int_t^\tau (\tau-u)\left(\mu_u^{ir} + \sigma_u^{ir}\xi_u^i\right) du - \int_t^\tau (\tau-u)\, \sigma_u^{ir} dW_u\right\}. \end{aligned}$$

The last but one equation follows from Fubini's Theorem for stochastic integrals [see (Heath *et al.*, 1992), Corollary 2 p. 99]. The expected change in the log–exchange rate is:

$$\begin{aligned} E_P^t \triangle_{t,\tau} \ln X &= E_P^t \left[ \int_t^\tau r_u^d - r_u^f + \frac{1}{2} \left( \left\| \xi_u^d \right\|^2 - \left\| \xi_u^f \right\|^2 \right) du \right] \\ &= \left( r^d - r^f \right) (T - t) + \\ &\quad E_P^t \left[ \int_t^\tau (\tau - u) \left( \mu_u^{rd} - \mu_u^{rf} \right) + \frac{1}{2} \left( \left\| \xi_u^d \right\|^2 - \left\| \xi_u^f \right\|^2 du \right) \right]. \end{aligned}$$

Without loss of generality $\xi^i$ can orthogonally be decomposed in $\alpha^i \sigma^{ir} + \beta^i \rho^i$ for $\rho^i \perp \sigma^{ir}$, $\|\rho^i\| = 1$ and some scalars $\alpha^i, \beta^i$. After the substitution of $\xi^i$ by its decomposition for all $t, \tau \in \mathbb{T}$ with $\tau \geq t$ the following equation must hold for the UIP to be viable:

$$\begin{aligned} &\left( r_t^d - r_t^f \right) (\tau - t) + E_P^t \left[ \int_t^\tau (\tau - u) \left( \mu_u^{rd} - \mu_u^{rf} \right) \right. \\ &\left. + \frac{1}{2} \left( \left( \alpha_u^d \right)^2 \left\| \sigma_u^{rd} \right\|^2 + \left( \beta_u^d \right)^2 \left\| \rho_u^d \right\|^2 - \left( \alpha_u^f \right)^2 \left\| \sigma_u^{rf} \right\|^2 - \left( \beta_u^f \right)^2 \left\| \rho_u^f \right\|^2 \right) du \right] = \end{aligned}$$

$$\begin{aligned} &(\tau - t) r_t^d - (\tau - t) r_t^f \\ &+ \ln E_P^t \exp \left\{ - \int_t^\tau (\tau - u) \left( \mu_u^{rf} + \alpha_u^f \left\| \sigma_u^{rf} \right\|^2 \right) du - \int_t^\tau (\tau - u) \sigma_u^{rf} dW_u \right\} \\ &- \ln E_P^t \exp \left\{ - \int_t^\tau (\tau - u) \left( \mu_u^{rd} + \alpha_u^d \left\| \sigma_u^{rd} \right\|^2 \right) du - \int_t^\tau (\tau - u) \sigma_u^{rd} dW_u \right\}. \end{aligned}$$

By re–arranging the terms of one country to one side of the equation it is equivalent to: $\forall t, \tau \in \mathbb{T}, \tau \geq t$

$$\begin{aligned} &E_P^t \left( \int_t^\tau \left( \beta_u^d \right)^2 du \right) + 2 \ln E_P^t \exp \left\{ \int_t^\tau (\tau - u) \left( E_P^t \left( \mu_u^{rd} \right) - \mu_u^{rd} - \alpha_u^d \left\| \sigma_u^{rd} \right\|^2 \right) \right. \\ &\qquad \left. + \frac{1}{2} E_P^t \left( \left( \alpha_u^d \right)^2 \left\| \sigma_u^{rd} \right\|^2 \right) du - \int_t^\tau (\tau - u) \sigma_u^{rd} dW_u \right\} = \\ &E_P^t \left( \int_t^\tau \left( \beta_u^f \right)^2 du \right) + 2 \ln E_P^t \exp \left\{ \int_t^\tau (\tau - u) \left( E_P^t \left( \mu_u^{rf} \right) - \mu_u^{rf} - \alpha_u^f \left\| \sigma_u^{rf} \right\|^2 \right) \right. \\ &\qquad \left. + \frac{1}{2} E_P^t \left( \left( \alpha_u^f \right)^2 \left\| \sigma_u^{rf} \right\|^2 \right) du - \int_t^\tau (\tau - u) \sigma_u^{rf} dW_u \right\}. \end{aligned}$$

Since this equation holds for all starting periods $t$ and maturities $\tau$ it is still satisfied after differentiating for $t$ at $\tau = t$ :

$$\left(\beta_t^d\right)^2 + \frac{1}{2} E_P^t \left( \left(\alpha_t^d\right)^2 \left\| \sigma_t^{rd} \right\|^2 \right) = \left(\beta_t^f\right)^2 + \frac{1}{2} E_P^t \left( \left(\alpha_t^f\right)^2 \left\| \sigma_t^{rf} \right\|^2 \right).$$

Taking this into account the equation imposes a one to one correspondence onto the square–values of the scalars $\beta^d$ and $\beta^f$ (and possibly some lower bound on one of both expressions). The projections $\left(\beta^d \rho^d, \beta^f \rho^f\right)$ of $\left(\xi^d, \xi^f\right)$ are in a subset of $\mathcal{L}^2 \left(P \times \lambda\right)^{2m}$. The second moment of these parts are controlled by the scalars $\left(\beta^d, \beta^f\right) \in \mathcal{L}^2 \left(P \times \lambda\right)^2$ alone. The last equation establishes a one–to–one correspondence between the square–values, which is a cross in $\mathbb{R}^2$ and thereby of one dimension less than the original space. Therefore this relation reduces the dimension of the range of $\left(\xi^d, \xi^f\right)$ by one. The restricted range is a subset of Lebesgue–measure zero in $\mathbb{R}^{2m}$. ■

### 3.7.5 Introducing Money

**Lemma 109** *The instantaneous nominal interest rates* $\hat{r}^i$*, nominal prices of risk* $\hat{\xi}^i$*, and the nominal exchange rate* $\hat{X}$ *are given by:*

$$\begin{aligned} \hat{r}_t^i &= r_t^i + \mu_t^{qi} + \sigma_t^{qi} \left(\hat{\xi}_t^i\right)^T, \\ \hat{\xi}_t^i &= \xi_t^i - \sigma_t^{qi}, \\ \hat{X}_t &= X_t \frac{q_t^d}{q_t^f} = X_0 \frac{q_t^d}{q_t^f} \frac{b_{0,t}^f z_t^f}{b_{0,t}^d z_t^d} = X_0 \frac{\hat{b}_{0,t}^f \hat{z}_t^f}{\hat{b}_{0,t}^d \hat{z}_t^d}. \end{aligned}$$

**Proof.** The nominal bonds $\hat{B}^i_{t,T}$ with maturity $T$ are in their respective currency nominal risk free, but in real terms risky:

$$\begin{aligned}
\hat{B}^i_{t,T} &= q^i_t E^t_{Q^i}\left[b_{t,T}/q^i_T\right] \\
&= E^t_{Q^i}\left[\exp -\left\{\int_t^T \mu^{qi}_s - \frac{1}{2}\left\|\sigma^{qi}_s\right\|^2 + r_s ds + \int_t^T \sigma^{qi}_s dW_s\right\}\right] \\
&= E^t_P\left[\exp\left\{\int_t^T \frac{1}{2}\left\|\sigma^{qi}_s\right\|^2 - \frac{1}{2}\left\|\xi^i\right\|^2 - \mu^{qi}_s - r_s ds + \int_t^T \xi^i - \sigma^{qi}_s dW_s\right\}\right] \\
&= \exp\left\{\int_t^T \left\|\sigma^{qi}_s\right\|^2 - \mu^{qi}_s ds\right\} E^t_{\tilde{P}}\left[\exp\left\{-\int_t^T \sigma^{qi}_s \xi^i_s + r^i_s ds\right\}\right], \text{ where} \\
d\tilde{P} &= \mathcal{E}\left(\int_t^T \xi^i - \sigma^{qi}_s dW_s\right) dP.
\end{aligned}$$

Then the nominal short rates are:

$$\begin{aligned}
\hat{r}^i_t &= -\lim_{T\to t}\frac{\partial}{\partial T}\ln \hat{B}^i_{t,T} \\
&= r^i_t + \mu^{qi}_t + \sigma^{qi}_t \xi^i_t - \left\|\sigma^{qi}_t\right\|^2.
\end{aligned}$$

The nominal prices of risk $\hat{\xi}^i$ are determined by considering the nominal prices of a set of real assets, which do not abolish the diffusion of the price inflator:

$$\begin{aligned}
d\left(H^i_t q^i_t\right) &= \left(r^i_t H^i_t - \sigma^i_t \xi^i_t + \mu^{qi}_t H^i_t + \sigma^i_t \sigma^{qi}_t\right) q^i_t dt + \left(\sigma^i_t + H^i_t \sigma^{qi}_t\right) q_t dW_t \\
&= \left(\hat{r}^i_t H^i_t - \left(\sigma^i_t + H^i_t \sigma^{qi}_t\right)\hat{\xi}^i_t\right) q^i_t dt + \left(\sigma^i_t + H^i_t \sigma^{qi}_t\right) q_t dW_t.
\end{aligned}$$

This implies

$$\begin{aligned}
&\left(\left(r^i_t + \mu^{qi}_t + \sigma^{qi}_t \xi^i_t - \left\|\sigma^{qi}_t\right\|^2\right) H^i_t - \left(\sigma^i_t + H^i_t \sigma^{qi}_t\right)\hat{\xi}^i_t\right) q^i_t = \\
&\qquad\qquad \left(r^i_t H^i_t - \sigma^i_t \xi^i_t + \mu^{qi}_t H^i_t + \sigma^i_t \sigma^{qi}_t\right) q^i_t,
\end{aligned}$$

which is equivalent to

$$\hat{\xi}^i_t - \xi^i_t = \left(\sigma^{qi}_t\right)^T.$$

The identity

$$\hat{X}_t = X_t \frac{q_t^d}{q_t^f}$$

rules out arbitrage opportunities via good and currency markets. Applying the result to the nominal exchange rate its SDE follows immediately:

$$\begin{aligned}
d\ln\hat{X}_t &= d\ln X_t + d\ln q_t^d - d\ln q_t^f \\
&= \left[r_t^d - r_t^f + \frac{1}{2}\left\|\xi_t^d\right\|^2 - \frac{1}{2}\left\|\xi_t^f\right\|^2\right] dt + \left(\xi_t^f - \xi_t^d\right)^T dW_t \\
&\quad + \left[\left(\mu_t^{qd} - \mu_t^{qf}\right) - \frac{1}{2}\left(\left\|\sigma_t^{qd}\right\|^2 - \left\|\sigma_t^{qf}\right\|^2\right)\right] dt + \left(\sigma_t^{qd} - \sigma_t^{qf}\right) dW_t \\
&= \left[\hat{r}_t^d - \hat{r}_t^f + \frac{1}{2}\left(\left\|\xi_t^d\right\|^2 + \left\|\sigma_t^{qf}\right\|^2 - \left\|\xi_t^f\right\|^2 - \left\|\sigma_t^{qf}\right\|^2\right)\right. \\
&\quad \left. -\sigma_t^{qd}\xi_t^d + \sigma_t^{qf}\xi_t^f\right] dt + \left[\left(\xi_t^f - \xi_t^d\right)^T + \sigma_t^{qd} - \sigma_t^{qf}\right] dW_t \\
&= \left[\hat{r}_t^d - \hat{r}_t^f + \frac{1}{2}\left(\left\|\hat{\xi}_t^d\right\|^2 - \left\|\hat{\xi}_t^f\right\|^2\right)\right] dt + \left(\hat{\xi}_t^f - \hat{\xi}_t^d\right)^T dW_t.
\end{aligned}$$

■

### 3.7.6 Corollary Reduced Form Equation

**Corollary 110 (Reduced Form Equation)** *Suppose that domestic financial markets are complete and that FX–spanning holds. Let* $\beta = \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1}$.

*(1) Asset prices are free of arbitrage if and only if the log–exchange rate is the solution to the SDE*

$$d\ln X = \left(r^d - r^f + \beta \begin{pmatrix} \mu^d - r^d H^d - \frac{1}{2}\sigma^d\left(\sigma^X\right)^T \\ \mu^f - r^f H^f + \frac{1}{2}\sigma^f\left(\sigma^X\right)^T \end{pmatrix}\right) dt + \sigma^X dW,$$

*which implies the orthogonal projection on asset prices*

$$d\ln X = \left(r^d - r^f\right) dt + \beta\left[d\begin{pmatrix} H^d \\ H^f \end{pmatrix} - \begin{pmatrix} r^d H^d - \frac{1}{2}\sigma^d\left(\sigma^X\right)^T \\ r^f H^f + \frac{1}{2}\sigma^f\left(\sigma^X\right)^T \end{pmatrix} dt\right].$$

*(2) Suppose additionally that assets are geometric Brownian:*

$$dH_k^i(t) = \mu_k^i(t)\, H_k^i(t)\, dt + \sigma_k^i(t)\, H_k^i(t)\, dW_t,\ k = 1, \dots, \hat{n}^i,\ i \in I.$$

*Then by arbitrage the log–exchange rate can be projected in the following way:*

$$\begin{aligned} d \ln X &= \left(r^d - r^f\right) dt + \beta \left[ d \ln \begin{pmatrix} H^d \\ H^f \end{pmatrix} - \begin{pmatrix} r^d \mathbf{1} \\ r^f \mathbf{1} \end{pmatrix} dt + \right. \\ & \left. \begin{pmatrix} -\sigma^d \\ \sigma^f \end{pmatrix} \sigma^X dt + \frac{1}{2} \left( \left\| \sigma_1^d \right\|^2, \dots, \left\| \sigma_{n^f}^f \right\|^2 \right)^T dt \right]. \end{aligned}$$

**Proof.** By Itô's rule $d \ln X = dX/X - 1/2d \langle X, X \rangle / X^2$ holds for continuous SDEs. Thereby the expression involving the diffusion matrices in the trend of the log–exchange rate from Proposition 89 transforms to

$$\sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \sigma^f \end{pmatrix} \left(\sigma^X\right)^T - \frac{1}{2} \sigma^X \left(\sigma^X\right)^T = \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \begin{pmatrix} -\sigma^d/2 \\ \sigma^f/2 \end{pmatrix} \left(\sigma^X\right)^T.$$

The implication follows then by the proper substitution:

$$\begin{aligned} \beta \begin{pmatrix} \mu^d \\ \mu^f \end{pmatrix} dt + \sigma^X dW &= \sigma^X \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix}^{-1} \left[ \begin{pmatrix} \mu^d \\ \mu^f \end{pmatrix} dt + \begin{pmatrix} \sigma^d \\ \sigma^f \end{pmatrix} dW \right] \\ &= \beta d \begin{pmatrix} H^d \\ H^f \end{pmatrix}. \end{aligned}$$

For the Brownian motion processes asset prices cancel down between the inverse of the diffusion matrix and the trends. ■

## 3.8 Appendix, Proofs of the Example Economy

**Proof.** The proof is divided into three steps. In the first step consumers' optimal decision is derived, then the labor markets and the valuation of productive assets will be analyzed, and, in a last step, the market clearing conditions are used to calculate the prices of risk, the short rates and the exchange rate.

**Consumers:** The social utility maximization is considered first. Because of complete markets the stochastic program reduces to a deterministic program, in which the optimal consumption and labor decisions are chosen path–wise optimally already in period zero. The 'stochastic' program of country $i$ sounds:

$$\mathfrak{L}^i\left(c^i,\eta^i\right) = E_P\left[\left.\int_0^{\mathcal{T}} n_s^i K_s^i u^i\left(c_s^{ii}, c_s^{\neg ii}, l_s^i\right) ds\right| \mathcal{F}_0\right] - \\ \eta^i\left[E_P\left[\left.\int_0^{\mathcal{T}}\left(c_s^{ii} + X_s^i c_s^{\neg ii} + w_s^i\left(\Phi^i - l_s^i\right)\right) z_s^i n_s^i b_{0,s}^i ds\right| \mathcal{F}_0\right] - \mathfrak{b}_0^i\right]$$

with

$$K_s^i = \exp\left(-\int_0^s \kappa_u^i du\right).$$

Because of the quasi–concavity of the von–Neuman–Morgenstern utility function and the linearity of the budget constraint the first order conditions (FOC) of the Lagrangian function are necessary and sufficient for an unique maximum. The FOCs are:

$$n_s^i K_s^i \nabla u^i\left(c_s^{ii}, c_s^{\neg ii}, l_s^i\right) = \eta^i n_s^i z_s^i b_{0,s}^i \begin{pmatrix} 1 \\ X_s^i \\ -w_s^i \end{pmatrix} \text{ and}$$

$$\mathfrak{b}_0^i = E_P\left[\left.\int_0^{\mathcal{T}}\left(c_s^{ii} + X_s^i c_s^{\neg ii} - w_s^i l_s^i\right) z_s^i n_s^i b_{0,s}^i ds\right| \mathcal{F}_0\right],$$

where the budget reads $\mathfrak{b}_0^i = \theta^{ii} S_0^i + \theta^{\neg ii} S_0^i X_0^i$. The first FOC is equivalent to:

$$\begin{pmatrix} c_s^{ii} \\ c_s^{\neg ii} \\ l_s^i \end{pmatrix} = \left(\nabla u^i \left(c_s^{ii}, c_s^{\neg ii}, l_s^i\right)\right)^{-1} \circ \left[\frac{\eta^i z_s^i b_{0,s}^i}{K_s^i} \begin{pmatrix} 1 \\ X_s^i \\ -w_s^i \end{pmatrix}\right].$$

Logarithmic utility functions have been assumed. Then the utility gradient simplifies in the following way:

$$\begin{aligned} u^i \left(c_s^{ii}, c_s^{\neg ii}, s\right) &= \alpha^{ii} \ln c_s^{ii} + \alpha^{\neg ii} \ln c_s^{\neg ii} + \left(1 - \alpha^{ii} - \alpha^{\neg ii}\right) \ln \left(\Phi^i - l_s^i\right) \\ &\quad \text{for } \alpha^{\neg ii}, \alpha^{ii}, \alpha^{ii} + \alpha^{\neg ii} \in (0,1), \Phi^i > 0, \end{aligned}$$

which implies the gradient

$$\nabla u^i \left(c_s^{ii}, c_s^{\neg ii}, l_s^i\right) = \begin{pmatrix} \alpha^{ii} / c_s^{ii} \\ \alpha^{\neg ii} / c_s^{\neg ii} \\ -\frac{1-\alpha^{ii}-\alpha^{\neg ii}}{\Phi^i - l_s^i} \end{pmatrix}.$$

By solving the FOCs the optimal consumption demand and labor supply are:

$$\begin{pmatrix} c_s^{ii} \\ c_s^{\neg ii} \\ l_s^i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \Phi^i \end{pmatrix} + \frac{K_s^i}{z_s^i b_{0,s}^i \eta^i} \begin{pmatrix} \alpha^{ii} \\ \alpha^{\neg ii} X_s^{\neg i} \\ \frac{-\left(1-\alpha^{ii}-\alpha^{\neg ii}\right)}{w_s^i} \end{pmatrix}.$$

These are the unrestricted solutions. Whereas the consumption is always strictly positive the unrestricted labor supply could be negative. This is, for example, the case if the wage is very low and resources are not scarce. Therefore one has to restrict the labor supply at zero:

$$l_s^i = \left[\Phi^i - \frac{K_s^i \left(1 - \alpha^{ii} - \alpha^{\neg ii}\right)}{z_s^i b_{0,s}^i \eta^i w_s^i}; 0\right]^+,$$

where $[a\,;b]^+$ is the abbreviation for $\max(a,b)$. By substituting the solution in the budget constraint one can solve for the Lagrangian multiplier:

$$\begin{aligned}
\mathfrak{b}_0^i &= E_P\left[\int_0^{\mathcal{T}} n_s^i\left(c_s^{ii}+X_s^i c_s^{\neg ii}-w_s^i l_s^i\right) z_s^i b_{0,s}^i ds\right] \\
&= E_P\left[\int_0^{\mathcal{T}} n_s^i\left(\frac{K_s^i}{z_s^i b_{0,s}^i \eta^i}\alpha^{ii}+X_s^i\frac{K_s^i}{z_s^i b_{0,s}^i \eta^i}\alpha^{\neg ii}X_s^{\neg i}\right.\right. \\
&\quad \left.\left. -w_s^i\left[\Phi^i-\frac{K_s^i\left(1-\alpha^{ii}-\alpha^{\neg ii}\right)}{z_s^i b_{0,s}^i \eta^i w_s^i};0\right]^+\right) z_s^i b_{0,s}^i ds\right],
\end{aligned}$$

which is further simplified to

$$\begin{aligned}
\mathfrak{b}_0^i &= E_P\left[\int_0^{\mathcal{T}} n_s^i\left(\frac{K_s^i}{\eta^i}\left(\alpha^{ii}+\alpha^{\neg ii}\right)\right.\right. \\
&\quad \left.\left. -\left[\Phi^i w_s^i z_s^i b_{0,s}^i-\frac{K_s^i\left(1-\alpha^{ii}-\alpha^{\neg ii}\right)}{\eta^i};0\right]^+\right) ds\right] \\
&= E_P\left[\int_0^{\mathcal{T}} n_s^i\left(\frac{K_s^i}{\eta^i}-\left[\Phi^i w_s^i z_s^i b_{0,s}^i;\frac{K_s^i\left(1-\alpha^{ii}-\alpha^{\neg ii}\right)}{\eta^i}\right]^+\right) ds\right].
\end{aligned}$$

The term $E_P[\cdot\,;\cdot]^+$ on the right hand side of this equation represents a call option, in which the Lagrangian multiplier is in the denominator of the strike price. Since the right hand side is strictly decreasing and continuous in $\eta^i$ (negative for large and plus infinity for a zero $\eta^i$) this equation has a unique positive solution for $\eta^i$. Nevertheless the equation has no closed form solution. At least $1/\eta^i$ is linear in $\mathfrak{b}_0^i$, that is

$$\begin{aligned}
\eta^i &= \frac{n_0^i K^i a^i}{\mathfrak{b}_0^i + b^i \mathcal{W}_0^i}, \text{ where} \\
0 &< a^i, b^i \leq 1, a^i > \left(1-\alpha^{ii}-\alpha^{\neg ii}\right).
\end{aligned}$$

Therefore the solution for an unrestricted labor supply corresponds to $a = b = 1$. For this reason let $\bar{\eta}^i$ with the corresponding $\bar{a}^i, \bar{b}^i$ represent the solution according to both restricted and unrestricted labor supply.

**Firms:** Now the proof turns to the valuation of productive assets. The linear technology forces the equilibrium wages to be:

$$w_t^i = \gamma_t^i.$$

In the case where marginal productivity is at any level of employment equal to the wage, the firms are indifferent to any level of employment. Since the gains from a linear production technology are zero in equilibrium only the resources are left for the firms: $D_t^i\left(L_t^i\right) - w_t^i L_t^i = \delta_t^i$. The gain process accumulates these resources:

$$\begin{aligned} G_t^i &= S_t^i + \int_0^t b_{t,s}^i \left(D_s^i\left(L_s^i\right) - w_s^i l_s^i\right) ds \\ &= S_t^i + \int_0^t b_{t,s}^i \delta_s^i ds. \end{aligned}$$

The discounted gain processes of productive assets $G_t^i b_{0,t}^i$ must be martingales with respect to the measure $Q^i$. This implies that the value of the firm is equal to the expected value with respect to the martingale measure of discounted resources:

$$\begin{aligned} G_t^i &= E_{Q^i}^t\left[b_{t,T}^i\left(S_T^i + \int_0^T b_{T,s}^i \delta_s^i ds\right)\right] \\ &= E_{Q^i}^t\left[b_{t,T}^i S_T^i\right] + E_{Q^i}^t\left[\int_t^T b_{t,s}^i \delta_s^i ds\right] + \int_0^t b_{t,s}^i \delta_s^i ds \Rightarrow \\ S_t^i &= E_{Q^i}^t\left[b_{t,T}^i S_T^i\right] + E_{Q^i}^t\left[\int_t^T b_{t,s}^i \delta_s^i ds\right] \\ &\underset{T\to\mathcal{T}}{\longrightarrow} E_{Q^i}^t\left[\int_t^{\mathcal{T}} b_{t,s}^i \delta_s^i ds\right], \text{ because } S_{\mathcal{T}}^i = 0. \end{aligned}$$

With the definitions of $\delta_s^i$ and $z_t^i$ this results in

$$\begin{aligned} S_t^i &= \delta_t^i E_P^t\left[\int_t^{\mathcal{T}} \exp\left(\int_t^s \mu_u^{\delta i} - r_u^i + \sigma_u^{\delta i}\xi_u^i du\right) \mathcal{E}\left(\int_t^s \sigma_u^{\delta i} dW_u\right) ds\right] \Rightarrow \\ dS_t^i &= S_t^i \frac{d\delta_t^i}{\delta_t^i} - \delta_t^i dt - S_t^i\left(\mu_t^{\delta i} - r_t^i + \sigma_t^{\delta i}\xi_t^i\right) dt \\ &= \left[S_t^i\left(r_t^i - \sigma_t^{\delta i}\xi_t^i\right) - \delta_t^i\right] dt + S_t^i \sigma_t^{\delta i} dW_t. \end{aligned}$$

Note that in finite time a price bubble cannot exist for a productive asset in an equilibrium even if it is in positive net supply, since it bursts in $\mathcal{T}$ so that $S^i_{\mathcal{T}} = 0$ must hold.

**Markets:** The next step is to exploit the market clearing conditions to derive the prices of risk and the interest rates. Labor markets clear at the level of consumers' labor supply. In the following the case of zero labor supply, i.e. $1_{l^i} \equiv 1_{\{l^i_t > 0\}}$, is handled by the definition $\bar{\gamma}^i_t = \gamma^i 1_{l^i}$. This implies that the dividends are the sum of production and resources:

$$D^i_t = \bar{\gamma}^i_t n^i_t \left( \Phi^i - \frac{K^i_t}{b^i_{0,t} z^i_t \bar{\eta}^i} \frac{\left(1 - \alpha^{ii} - \alpha^{\neg ii}\right)}{\gamma^i_t} \right) + \delta^i_t.$$

In equilibrium the supply of consumption goods has to be equal to the demand, i.e. for all $t \in \mathbb{T}$ and $i, j \in I$ :

$$n^i_t c^{ji}_t + n^{\neg i}_t c^{j \neg i}_t = D^i_t,$$

which implies the following equation $\forall t \in \mathbb{T}$ :

$$n^i_t \frac{K^i_t}{b^i_{0,t} z^i_t \bar{\eta}^i} \alpha^{ii} + n^{\neg i}_t \frac{K^{\neg i}_t}{b^{\neg i}_{0,t} z^{\neg i}_t \bar{\eta}^{\neg i}} \alpha^{i \neg i} X^i_t = \bar{\gamma}^i_t n^i_t \left( \Phi^i - \frac{K^i_t}{b^i_{0,t} z^i_t \bar{\eta}^i} \frac{\left(1 - \alpha^{ii} - \alpha^{\neg ii}\right)}{\gamma^i_t} \right) + \delta^i_t.$$

By substituting $X_t = X_0 z^{\neg i}_t b^{\neg i}_{0,t} / \left(z^i_t b^i_{0,t}\right)$ and by multiplying with $z^i_t b^i_{0,t} / n^i_t K^i_t$ this equation transforms into the here called '*G*ood *M*arkets'–equation:

$$\begin{aligned} &\alpha^{ii}/\bar{\eta}^i + n^{\neg i}_t K^{\neg i}_t \alpha^{i \neg i} X^i_0 / \left(\bar{\eta}^{\neg i} n^i_t K^i_t\right) = \\ &\qquad z^i_t b^i_{0,t} \bar{\gamma}^i_t \Phi^i / K^i_t - 1_{l^i} \left(1 - \alpha^{ii} - \alpha^{\neg ii}\right) / \bar{\eta}^i + z^i_t b^i_{0,t} \delta^i_t / \left(n^i_t K^i_t\right). \quad \text{(GM)} \end{aligned}$$

Considering only the stochastic differentials the diffusion part of the GM–equation reads:

$$\begin{aligned} 0 &= z^i_t b^i_{0,t} \left[ n^i_t \bar{\gamma}^i_t \Phi^i \left(\xi^i_t + \sigma^{\gamma i}_t\right) + \delta^i_t \left(\xi^i_t + \sigma^{\delta i}_t\right)\right] / \left(n^i_t K^i_t\right) \Leftrightarrow \\ \xi^i_t &= -\frac{n^i_t \bar{\gamma}^i_t \Phi^i \sigma^{\gamma i}_t + \delta^i_t \sigma^{\delta i}_t}{n^i_t \bar{\gamma}^i_t \Phi^i + \delta^i_t}. \end{aligned}$$

Since $\xi_t^i$ is a convex combination of bounded processes it induces a martingale measure and a pricing asset. Now the locally deterministic trend parts are left over to solve for the interest rates:

$$
\begin{aligned}
&\left[\left(\mu^{n\neg i}-\kappa^{\neg i}\right)-\left(\mu^{ni}-\kappa^{i}\right)\right] n_t^{\neg i} K_t^{\neg i} \alpha^{i\neg i} X_0^i / \left(\bar{\eta}^{\neg i} n_t^i K_t^i\right) = \\
&\qquad \left(\kappa^i+\mu^{\gamma i}-r^i+\left(\xi^i\right)^T \sigma^{\gamma i}\right) z_t^i b_{0,t}^i \bar{\gamma}_t^i \Phi^i / K_t^i \\
&\qquad\quad +\left(\mu^{\delta i}-r^i+\left(\xi^i\right)^T \sigma^{\delta i}-\mu^{ni}+\kappa^i\right) z_t^i b_{0,t}^i \delta_t^i / \left(n_t^i K_t^i\right).
\end{aligned}
$$

Finally, one solves for the interest rates while replacing again $X_0/\left(z_t^i b_{0,t}^i\right) = X_t / z_t^{\neg i} b_{0,t}^{\neg i}$:

$$
\begin{aligned}
r_t^i &= \kappa_t^i+\frac{n_t^i \bar{\gamma}_t^i \Phi^i \mu^{\gamma i}+\delta_t^i\left(\mu_t^{\delta i}-\mu_t^{ni}\right)}{n_t^i \bar{\gamma}_t^i \Phi^i+\delta_t^i}-\left\|\xi_t^i\right\|^2 \\
&\quad +\left[\frac{n_t^{\neg i} K_t^{\neg i} \alpha^{i\neg i}}{\bar{\eta}^{\neg i} z_t^{\neg i} b_{0,t}^{\neg i}} X_t^i \cdot \frac{1}{n_t^i \bar{\gamma}_t^i \Phi^i+\delta_t^i}\right]\left(\left(\kappa_t^{\neg i}-\mu_t^{n\neg i}\right)-\left(\kappa_t^i-\mu_t^{ni}\right)\right).
\end{aligned}
$$

This solution is only closed if $\left(\kappa_t^i-\mu_t^{ni}\right)-\left(\kappa_t^{\neg i}-\mu_t^{n\neg i}\right)=0$ because the denominator of the factor proceeding this expression contains the locally riskless bank account, which itself depends on the short rate process. As the short rates and the prices of risk have been determined for both currencies the SDE of the exchange rate follows by arbitrage:

$$
\begin{aligned}
X_t &= X_0 \frac{b_{0,t}^f z_t^f}{b_{0,t}^d z_t^d} \Rightarrow \\
d \ln X &= \left[r^d-r^f+\frac{1}{2}\left\|\xi^d\right\|^2-\frac{1}{2}\left\|\xi^f\right\|^2\right] dt+\left(\xi^f-\xi^d\right) dW.
\end{aligned}
$$

The substitution of the short rates and the prices of risk by their solutions leads to a SDE in terms of mostly exogenous factors:

$$
\begin{aligned}
d\ln X_t \;=\; & \left[ \frac{n_t^d \bar{\gamma}_t^d F^d \left(\mu^{\gamma d} + \mu^{nd}\right) + \delta_t^i \mu_t^{\delta d}}{n_t^d \bar{\gamma}_t^d F^d + \delta_t^d} - \frac{n_t^f \bar{\gamma}_t^f F^f \left(\mu^{\gamma f} + \mu^{nf}\right) + \delta_t^f \mu_t^{\delta f}}{n_t^f \bar{\gamma}_t^f F^f + \delta_t^f} \right. \\
& + \left( \left(\kappa^d - \mu^d\right) - \left(\kappa^f - \mu^f\right) \right) \cdot \\
& \left( 1 + \frac{n_t^f K_t^f \alpha^{df} X_t}{\bar{\eta}^f z_t^f b_{0,t}^f \left(n_t^d \bar{\gamma}_t^d F^d + \delta_t^d\right)} - \frac{n_t^d K_t^d \alpha^{fd}}{X_t \bar{\eta}^d z_t^d b_{0,t}^d \left(n_t^f \bar{\gamma}_t^f F^f + \delta_t^f\right)} \right) \\
& \left. - \frac{1}{2} \frac{\left\| n_t^d \bar{\gamma}_t^d F^d \sigma_t^{\gamma d} + \delta_t^d \sigma_t^{\delta d} \right\|^2}{\left(n_t^d \bar{\gamma}_t^d F^d + \delta_t^d\right)^2} + \frac{1}{2} \frac{\left\| n_t^f \bar{\gamma}_t^f F^f \sigma_t^{\gamma f} + \delta_t^f \sigma_t^{\delta f} \right\|^2}{\left(n_t^f \bar{\gamma}_t^f F^f + \delta_t^f\right)^2} \right] dt + \\
& \left[ \frac{n_t^f \bar{\gamma}_t^f F^f \sigma_t^{\gamma f} + \delta_t^f \sigma_t^{\delta f}}{n_t^f \bar{\gamma}_t^f F^f + \delta_t^f} - \frac{n_t^d \bar{\gamma}_t^d F^d \sigma_t^{\gamma d} + \delta_t^d \sigma_t^{\delta d}}{n_t^d \bar{\gamma}_t^d F^d + \delta_t^d} \right] dW_t.
\end{aligned}
$$

The last unknown is the exchange rate in period zero. The price of each country's domestic good has been normalized to one. This normalization is possible because the budged constrained is homogenous of degree zero in prices. This implies that the market clearing conditions have one degree of freedom. Since a market clearing condition at any point in time has Lebesgue–measure zero, Walras' law applies only to an integrated market clearing condition. The following transformations show Walras' law for the present value of the excess demand:

$$
\underbrace{E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \left(c_s^{dd} n_s^d + c_s^{df} n^f\right) ds\right)}_{\text{valued consumption}} = \underbrace{E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \left(\bar{\gamma}_s^d n_s^d l_s^d + \delta_s^d\right) ds\right)}_{\text{valued production}} \tag{DCM}
$$

This equation is equivalent to

$$\begin{aligned}
&\mathfrak{b}_0^d - E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \left(c_s^{fd} n_s^d X_s - \bar{\gamma}_s^d n_s^d l_s^d\right) ds\right) + \\
&X_0\left(\mathfrak{b}_0^f - E_P\left(\int_0^{\mathcal{T}} b_{0,s}^f z_s^f \left(c_s^{ff} n_s^f - \bar{\gamma}_s^f n_s^f l_s^f\right) ds\right)\right) = \\
&E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \bar{\gamma}_s^d n_s^d l_s^d ds\right) + S_0^d.
\end{aligned}$$

This equation is again equivalent to the corresponding Walras' identity on foreign good markets:

$$E_P\left(\int_0^{\mathcal{T}} b_{0,s}^f z_s^f \left(c_s^{fd} n_s^d + c_s^{ff} n^f\right) ds\right) = E_P\left(\int_0^{\mathcal{T}} b_{0,s}^f z_s^f \left(\bar{\gamma}_s^f n_s^f l_s^f + \delta_s^f\right) ds\right).$$

By Walras' law $X_0$ is determined by one of these two equivalent integrated market clearing conditions. Here the market for the domestic consumption good (DCM) is chosen. By the substitution of the equilibrium consumption demand and labor supply Equation (DCM) transforms to:

$$\begin{aligned}
&E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \left(\frac{K_s^d}{b_{0,s}^d z_s^d \bar{\eta}^d}\alpha^{dd} n^d + \frac{K_s^f}{b_{0,s}^f z_s^f \bar{\eta}^f}\alpha^{df} X_s n^f\right) ds\right) = \\
&\qquad E_P\left(\int_0^{\mathcal{T}} b_{0,s}^d z_s^d \bar{\gamma}_s^d n_s^d \left(F^d - \frac{K_s^d}{b_{0,s}^d z_s^d \bar{\eta}^d}\frac{\left(1-\alpha^{dd}-\alpha^{fd}\right)}{\gamma_s^d}\right) ds\right) + S_0^d.
\end{aligned}$$

While this equation is equivalent to:

$$\begin{aligned}
\alpha^{dd} n_0^d \frac{K^d}{\bar{\eta}^d} + \alpha^{df} n_0^f X_0 \frac{K^f}{\bar{\eta}^f} &= n_0^d \frac{K^d}{\bar{\eta}^d}\left(\alpha^{dd}+\alpha^{fd}\right) - \mathfrak{b}_0^d + S_0^d \Leftrightarrow \\
\alpha^{df} n_0^f X_0 \frac{K^f}{\bar{\eta}^f} - \alpha^{fd} n_0^d \frac{K^d}{\bar{\eta}^d} &= S_0^d - \mathfrak{b}_0^d.
\end{aligned}$$

With the solution for $\bar{\eta}^i = \frac{n_0^i K^i \bar{a}^i}{\mathfrak{b}_0^i + \bar{b}^i \mathcal{W}_0^i}$ and the definition $\mathfrak{b}_0^i = \theta^{ii} S_0^i + \theta^{\neg ii} S_0^i X_0^i$ this equation sounds

$$X_0 \left( \theta^{ff} S_0^f + \theta^{df} S_0^d / X_0 + \bar{b}^f \mathcal{W}_0^f \right) \alpha^{df} / \bar{a}^f$$

$$- \left( \theta^{dd} S_0^d + \theta^{fd} S_0^f X_0 + \bar{b}^d \mathcal{W}_0^d \right) \alpha^{fd} / \bar{a}^d = S_0^d - \theta^{dd} S_0^d - \theta^{fd} S_0^d X_0 .$$

Finally the solution for the exchange rate follows:

$$X_0 = \frac{\alpha^{fd} \left( \theta^{dd} S_0^d + \bar{b}^d \mathcal{W}_0^d \right) / \bar{a}^d + \left( 1 - \alpha^{df} / \bar{a}^f \right) \theta^{df} S_0^d}{\alpha^{df} \left( \theta^{ff} S_0^f + \bar{b}^f \mathcal{W}_0^f \right) / \bar{a}^f + \left( 1 - \alpha^{fd} / \bar{a}^d \right) \theta^{fd} S_0^f} .$$

This finishes the explicit calculation of all equilibrium values, which shows existence and uniqueness at once. ■

## 3.9 Moore–Penrose Inverse and related Issues

### 3.9.1 Singular Value Decomposition

Any matrix $A \in \mathbb{R}^n \otimes \mathbb{R}^m$ admits a singular value decomposition

$$A = V \Sigma W^T .$$

$\Sigma \in \mathbb{R}^n \otimes \mathbb{R}^m$ contains on its main diagonal all non–negative square–roots of the eigenvalues of $A^T A$ whereas all off–diagonal elements are zero. $V \in \mathbb{R}^n \otimes \mathbb{R}^n$ and $W \in \mathbb{R}^m \otimes \mathbb{R}^m$ consist of the orthonormal eigenvectors of $AA^T$ and $A^T A$, thus $W^T W = I_m$ as well as $VV^T = I_n$. Due to possible singularities in $AA^T$ and $A^T A$, i.e. if $rank\left(AA^T\right) < n - 1$ or $rank\left(A^T A\right) < m - 1$, their eigenvectors are not uniquely determined. The eigenvectors are ordered in the same way as the corresponding singular values.

### 3.9.2 Moore–Penrose Inverse

The Moore–Penrose inverse $A^+$ is uniquely defined by the following four properties:

1. $AA^+$ and $A^+A$ are symmetric,

2. $AA^+A = A$ and $A^+AA^+ = A$.

$A^+$ has the following two representations:

$$\begin{aligned} A^+ &= W\Sigma^+V^T \\ &= \lim_{\varepsilon\to 0} A^T\left(AA^T + \varepsilon I_n\right)^{-1}, \end{aligned}$$

where $\Sigma^+$ is the transpose of $\Sigma$ with the non–zero elements of $\Sigma$ replaced by their reciprocals.

### 3.9.3 Orthogonal Projections

$AA^+, A^+A, I - AA^+$ and $I - A^+A$ are the orthogonal projectors on $\mathcal{R}(A)$, $\mathcal{R}\left(A^T\right), \mathcal{N}\left(A^T\right)$ and $\mathcal{N}(A)$ respectively. $\mathcal{R}(A)$ denotes the span of $A$ and $\mathcal{N}(A)$ the null–space, thus $\mathcal{R}\left(A^T\right)^{\perp} = \mathcal{N}(A)$. Projectors are idempotent and have only eigenvalues equal to zero or one.
For more results related to generalized inverses and projections the reader is referred to (Campbell & Meyer, 1979).

# References


Amin, Kaushik I., & Jarrow, Robert A. 1991. Pricing Foreign Currency Options under Stochastic Interest Rates. *Journal of International Money and Fiance*, **10**, 310 – 329.

Basak, Suleyman, & Gallmeyer, Michael. 1999. Currency Prices, the Nominal Exchange Rate, and Security Prices in a Two–Country Dynamic Monetary Equilibrium. *Mathematical Finance*, **9**(1), 1 – 30.

Blume, L., & Easley, D. 1992. Evolution and Market Behavior. *Journal of Economic Theory*, **58**, 9 – 40.

Breuer, Wolfgang. 1996. Das Siegel–Paradoxon. *Wirtschaftswissenschaftliches Studium*, **25**(8), 413 – 415.

Campbell, S. L., & Meyer, Jr. C. D. 1979. *Generalized Inverses of Linear Transformations*. Surveys and Reference Works in Mathematics. London: Pitman.

Chiang, Thomas C. 1988. The Forward Rate as a Predictor of the Future Spot Rate – A Stochastic Coefficient Approach. *Journal of Money, Credit, and Banking*, **20**(2), 212 – 232.

Chung, K. L., & Williams, R. J. 1990. *Introduction to Stochastic Integration*. 2nd edn. Probability and Its Applications. Boston: Birkhäuser.

Clark, Stephen A. 1993. The Valuation Problem in Arbitrage Price Theory. *Journal of Mathematical Economics*, **22**, 463 – 478.

Dornbusch, Rüdiger. 1976. Expectations and Exchange Rate Dynamics. *Journal of Political Economy*, **84**(6), 1161 – 1176.

Dothan, M. U. 1990. *Prices in financial markets*. New York: Oxford University Press.

Duffie, Darrell. 1996. *Dynamic Asset Pricing Theory*. 2nd edn. Princeton USA: Princeton University Press.

El–Karoui, Nicole, & Quenez, Marie–Claire. 1995. Dynamic Programming and Pricing of Contingent Claims in an Incomplete Market. *SIAM Journal of Control and Optimization*, **33**(1), 29 – 66.

Fama, Eugene F. 1984. Forward and Spot Exchange Rates. *Journal of Monetary Economics*, **14**(3), 319 – 338.

Fehr, Ernst, & Tyran, Jean-Robert. 2000. Does Money Illusion Matter? *American Economic Review*. Forthcoming.

Flesaker, B., & Hughston, L. 1996. *Positive Interest: Foreign Exchange*. Risk Publications London. Chap. 22, pages 352 – 367.

Föllmer, Hans, & Schweizer, Martin. 1991. Hedging of Contingent Claims under Incomplete Information. *Pages 389 – 414 of:* Davis, M. H. A., & Elliott, R. J. (eds), *Applied Stochastic Analysis*. Stochastics Monographs Vol. 5. London/New York: Gordon and Breach.

Frankel, Jeffrey A., & Rose, Andrew K. 1995. Empirical Research on Nominal Exchange Rates. *Chap. 33, pages 1689 – 1729 of:* Grossman, G., & Rogoff, K. (eds), *Handbook of International Economics*, vol. 3. Amsterdam: Elsevier Science B.V.

Geanakoplos, John, & Mas–Colell, Andreu. 1989. Real Indeterminacy with Financial Assets. *Journal of Economic Theory*, **47**, 22 – 38.

Geweke, John, & Feige, Edgar. 1979. Some Joint Tests of the Efficiency of Markets for Forward Foreign Exchange. *Review of Economics and Statistics*, **61**(August), 334 – 341.

Hansen, Lars Peter, & Hodrick, Robert J. 1980. Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis. *Journal of Political Economy*, **88**(5), 829 – 853.

Harrison, J. M., & Kreps, D. M. 1979. Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, **20**, 381 – 408.

Heath, David, Jarrow, Robert, & Morton, Andrew. 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, **60**(1), 77 – 105.

Hens, Th., Jäger, E., Kirman, A., & Philips, L. 1999. Exchange Rates and Oligopoly. *European Economic Review*, **43**, 621 – 648.

Hens, Thorsten. 1998. Elements of General Equilibrium Analysis. *Chap. 5, pages 139 – 210 of:* Kirman, Alan (ed), *Festschrift in Honour of Debreu's 75th Birthday*. Oxford: Blackwell Publishers.

Hodrick, R. J. 1987. *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*. Chur, Switzerland: Harwood Academic Publishers.

Horn, Roger A., & Johnson, Charles R. 1985. *Matrix Analysis*. 1996 edn. Cambridge University Press.

Huang, Chi–Fu, & Pagès, Henri. 1992. Optimal Consumption and Portfolio Policies with an Infinite Horizon: Existence and Convergence. *Annals of Applied Probability*, **2**(1), 36 – 64.

IMF. 1997. *IMF World Economic Outlook*. International Monetary Fund.

Judge, George G., Griffiths, William E., Hill, R. Carter, Lütkepohl, Helmut, & Lee, Tsoung-Chao. 1985. *The Theory and Practice of Econometrics*. 2nd edn. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.

Karatzas, I., Lehoczky, J. P., & Shreve, S. E. 1990. Existence and Uniqueness of Multi–Agent Equilibrium in a Stochastic, Dynamic Consumption / Investment Model. *Mathemathics of Operations Research*, **15**(1), 80 – 128.

Karatzas, Ioannis, & Shreve, Steven E. 1988. *Brownian Motion and Stochastic Calculus*. Berlin: Springer–Verlag.

Karatzas, Ioannis, & Shreve, Steven E. 1998. *Methods of Mathematical Finance*. Applications of Mathematics, Stochastic Modelling and Applied Probability, vol. 39. New York: Springer–Verlag.

Lewis, Karen K. 1995. Puzzels in International Financial Markets. *Chap. 37, pages 1913 – 1971 of:* Grossman, G., & Rogoff, K. (eds), *Handbook of International Economics*, vol. 3. Amsterdam: Elsevier Science B.V.

Lexton, Douglas, Isard, Peter, Faruqee, Hamid, Prasad, Eswar, & Turtelboom, Bart. 1998 (May). *Multimod Mark III – The Core Dynamics and Steady–State Models*. Internatinal Monetary Fund, Washington DC.

Liptser, R. S., & Shiryaev, A. N. 1977. *Statistics of Random Processes. Vol. I: Gerneral Theory*. New York: Springer–Verlag.

Lo, A. W., & MacKinlay, A. C. 1988. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, **1**, 41 – 66.

Lo, Andrew W. 1988. Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data. *Econometric Theory*, **4**, 231 – 247.

McCallum, B. E. 1994. A Reconsideration of the Uncovered Interest Rate Parity Relationship. *Journal of Monetary Economics*, **33**, 105 – 132.

Meese, R. A. 1989. Empirical Assessment of Foreign Currency Risk Premium. *In:* Stone, Courtesan (ed), *Financial Risk: Theory, Evidence and Implications*. Boston: Kluwer Academic Press.

Musiela, Marek, & Rutkowski, Marek. 1997. *Martingale Methods in Financial Modelling*. Corrected 2nd printing 1998 edn. Applications of Mathematics, Stochastic Modelling and Applied Probability 36. Berlin: Springer–Verlag.

Roemer, John E. 1996. *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press.

Saá–Requejo, J. 1994. *The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets*. mimeo INSEAD Université de Paris.

Schmidt, Roland. 1993 (January). *Why the Forward Rate is a Biased Predictor of the Future Spot Rate If Investors are Risk Neutral*. mimeo Universität Bonn.

Shafer, W., & Sonnenschein, H. 1982. Market Demand and Excess Demand Functions. *Chap. 14, pages 71 – 95 of:* Arrow, K., & Intrilligator, M. (eds), *Handbook of Mathematical Economics Vol. II*. Handbook of Mathmetical Economics, vol. II. Amsterdam: North–Holland.

Siebert, Anne. 1989. The Risk Premium in the Foreign Exchange Market. *Journal of Money, Credit, and Banking*, **21**(1), 49 – 64.

Siegel, Jeremy J. 1972. Risk, Interest, and Forward Exchange. *Quarterly Journal of Economics*, **86**(May), 303 – 306.

Taylor, Mark P. 1995. The Economics of Exchange Rates. *Journal of Economic Literature*, **33**(March), 13 – 47.

The Economist. 1995. Big, the Foreign–Exchange Market. *The Economist*, September 23rd – 29th, 77 – 78.

Trojani, Fabio. 1999. *From Uncovered Interest Rates Parity Towards The Identification of Exchange Rates Risk Premia*. Ph.D. thesis, Universität Zürich.

Winters, Bill. 1999. Masterclass – with JP Morgan: Wrong–Way Exposure. *Risk*, July, 52 – 55.

Zapatero, Fernando. 1995. Equilibrium Asset Prices and Exchange Rates. *Journal of Economic Dynamics and Control*, **19**, 787 – 811.

# 4 Curriculum Vitae

| | |
|---|---|
| 06.10.1970 | born in Cologne |
| 05.90 | A–Level Degree (Abitur) |
| 07.90 – 06.91 | Military Service |
| 08.95 | Diploma in Economics (Diplom–Volkswirt), University of Bonn |
| 09.95 – 04.96 | Teaching Assistent in Business Economics & Finance, University of Bonn |
| 05.96 – 07.99 | Teaching Assistent in Micro–Economics & Finance, University of Bielefeld |
| since 08.99 | Consultant, Financial Risk Solutions, Deloitte & Touche, Düsseldorf |