

Breuer, Wolfgang; Gürtler, Marc

Working Paper

Two-Fund separation and positive marginal utility

Working Paper Series, No. FW11V3

Provided in Cooperation with:

Technische Universität Braunschweig, Institute of Finance

Suggested Citation: Breuer, Wolfgang; Gürtler, Marc (2004) : Two-Fund separation and positive marginal utility, Working Paper Series, No. FW11V3, Technische Universität Braunschweig, Institut für Finanzwirtschaft, Braunschweig

This Version is available at:

<https://hdl.handle.net/10419/55243>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Working Paper Series



Two-Fund Separation and Positive Marginal Utility

by Wolfgang Breuer and Marc Gürtler

No.: FW11V3/04

First Draft: 2004-12-21

This Version: 2006-05-10

Braunschweig University of Technology
Institute for Economics and Business Administration
Department of Finance
Abt-Jerusalem-Str. 7
D-38106 Braunschweig

Two-Fund Separation and Positive Marginal Utility

by Wolfgang Breuer* and Marc Gürtler**

Abstract. The requirement of positive marginal utility only makes it possible to derive a restricted two-fund separation theorem for portfolio selection problems replacing the original separation theorem of Cass and Stiglitz (1970). We use our findings for a re-examination of the bias-in-beta problem in mutual funds performance evaluation and of the relevance of the standard CAPM without borrowing restrictions. We also present empirical evidence for the only limited validity of the separation theorem when explicitly recognizing positive marginal utility. Moreover, quadratic utility functions are not apt to approximate the admissible range of risk preferences in the case of higher-order utility functions.

Keywords: two-fund separation, HARA utility, positive marginal utility, borrowing restrictions, Capital Asset Pricing Model, bias in beta, performance evaluation

JEL-Classification: G11

* **Professor Dr. Wolfgang Breuer**
RWTH Aachen University
Department of Finance
Templergraben 64, 52056 Aachen, Germany
Phone: +49 241 8093539 - Fax: 8092163
eMail: wolfgang.breuer@rwth-aachen.de

** **Professor Dr. Marc Gürtler**
Braunschweig University of Technology
Department of Finance
Abt-Jerusalem-Str. 7, 38106 Braunschweig, Germany
Phone: +49 531 3912895 - Fax: 3912899
eMail: marc.guertler@tu-bs.de

1 Introduction

The two-fund separation theorem originally developed by Tobin (1958) for mean-variance preferences and later on generalized by Hakansson (1969) and – in particular – Cass and Stiglitz (1970) to the consideration of utility functions with hyperbolic absolute risk aversion (HARA, henceforth) is one of the basic ingredients of modern portfolio and capital market theory. For individual portfolio optimization, the two-fund separation theorem states the optimality of a certain (sub-) portfolio of risky assets regardless of the investor's degree of risk aversion and his initial wealth, as long as the investor's utility function belongs to a certain class. For capital market theory, the two-fund separation theorem enables us to derive very straightforward valuation functions for risky payoffs in equilibrium with the most important special case of the Capital Asset Pricing Model according to Sharpe (1964), Lintner (1965), and Mossin (1966) (SLM-CAPM, henceforth).

One of the key prerequisites of the two-fund separation theorem is the possibility to buy or sell arbitrary amounts of a riskless asset with a certain interest rate r_0 . Moreover, an investor's portfolio selection problem is certainly *well defined*, only when – over the whole domain of accessible return realizations – utility is defined and marginal utility is positive. Except for the case of negative exponential utility, the consideration of such a well defined portfolio selection problem for HARA preferences requires a restriction of the admissible amount of riskless lending and borrowing. But then the two-fund separation theorem by Cass and Stiglitz (1970) cannot be applied in its original form any longer. In what follows, we want to examine under which conditions a restricted two-fund separation theorem may hold in the case of borrowing and lending constraints with respect to the riskless asset.

In the following section 2, we describe our formal background and present the basic problem graphically for the special case of quadratic utility. Section 3 then characterizes formally for the general case of HARA utility which restrictions on riskless lending and borrowing, risky investments, investors' preferences and initial wealth as well as return

distributions are necessary and sufficient in order to guarantee that the investor's portfolio selection problem is well defined and exhibits the separation property. In section 4, central implications of our findings are presented to emphasize the practical importance of the problem under consideration. First of all, we derive under which conditions the so-called "bias-in-beta" problem in mutual fund performance evaluation does not occur. Second, for the case of quadratic utility, we are able to state conditions so that the SLM-CAPM remains valid even in the case of only restricted borrowing and lending opportunities. Otherwise, the Black-CAPM (according to Black, 1972) becomes relevant. Section 5 presents an additional empirical application of our theoretical results by examining a set of 45 different mutual funds investing in German stocks under the (alternative) assumptions of quadratic, cubic or bi-quadratic HARA utility. It will be shown that the admissible variations of an investor's risk aversion and initial endowment supporting two-fund separation are quite small, in particular for quadratic utility, casting some doubts upon the practical relevance of the two-fund separation theorem. Moreover, the range of admissible parameter values becomes considerably greater when switching from quadratic to cubic or bi-quadratic utility. In this respect, quadratic utility functions do not approximate results for higher-order utility functions quite well – a result that is in contrast to the well-documented good quality of quadratic utility when approximating optimal portfolio structures for higher-order utility functions. Section 6 summarizes our findings and gives a brief outlook onto potential future research.

2 The basic problem

2.1 The assumptions

In what follows, we consider an individual who wants to invest his initial wealth W_0 from $t = 0$ to $t = 1$ in three different securities $i = 0, 1, 2$. Let x_i be the fraction of W_0 which is invested in security i . Negative values for x_i stand for short sales. Securities 1 and 2 are risky

with \tilde{r}_i ($i = 1, 2$) being their uncertain rate of return from $t = 0$ to $t = 1$. Security 0 is riskless and its interest rate is denoted as r_0 . Furthermore, we define $\tilde{u}_i := \tilde{r}_i - r_0$ as the excess return of security i and assume $E(\tilde{u}_i) > 0$ ($i = 1, 2$) as well as $E(\tilde{u}_1) \neq E(\tilde{u}_2)$.

Let \tilde{W}_1 be the investor's uncertain terminal wealth at time $t = 1$. Assume further that the investor's utility function $U(\tilde{W}_1)$ exhibits hyperbolic absolute risk aversion (HARA), i.e. we have $-U''(\tilde{W}_1)/U'(\tilde{W}_1) = 1/(a + b \cdot \tilde{W}_1)$ ($a, b \in \mathfrak{R}$). In what follows, a *class* of HARA utility functions is defined as the subset of all HARA functions with the same b and arbitrary a . To abbreviate notation, we define the fraction $\tau := a / W_0$ as some kind of standardized risk tolerance, because (for given b) the higher a , the lower is an individual's risk aversion. τ is the most important determinant of an investor's behavior for any given class of HARA utility functions and given return distributions of all securities $i = 0, 1, 2$.

With $y := x_1 / (x_1 + x_2)$ we can characterize the structure of the risky part of the investor's overall portfolio, since y describes the fraction of the investor's risky holding which is invested in security 1. Under the assumption $x \in \mathfrak{R}$ and neglecting special problems resulting from the requirement of positive marginal utility, Cass and Stiglitz (1970) showed that (only) for a given class of HARA utility functions the resulting optimal structure $y^{(\mathfrak{R})^*}$ of the investor's risky engagement is independent of parameter value a and initial wealth W_0 and thus τ for any given return distributions of securities 1 and 2. W_0 and a only determine the *absolute* amount of the individual's risky investment in relation to his riskless borrowing and lending, but they do not influence the *structure* of an investor's risky portfolio. This means, while $y^{(\mathfrak{R})^*}$ is not depending on W_0 and a , the optimal values for x_0, x_1 , and x_2 do.

However, as suggested in the introduction, only the consideration of well defined portfolio selection problems in the following way seems to make sense.

Definition 1. Consider an investor's portfolio selection problem. We call it well defined if violations of the domain of the investor's utility function and negative marginal utility are effectively avoided for all admissible portfolios and all possible return realizations.

Furthermore, we state

Definition 2. Consider an investor with a HARA utility function belonging to a certain class b who faces a well defined portfolio selection problem. Let X , Y , and T be sets of \mathfrak{R} . We assume $x_0 \in X \supset (\underline{x}_0; \bar{x}_0)$, $\underline{x}_0 = \inf X$, $\bar{x}_0 = \sup X$ with $-\infty \leq \underline{x}_0 \leq 0$ and $\infty \geq \bar{x}_0 \geq 1$ as well as $y \in Y \supset (\underline{y}; \bar{y})$, $\underline{y} = \inf Y$, $\bar{y} = \sup Y$ with $-\infty \leq \underline{y} \leq 0$ and $\infty \geq \bar{y} \geq 1$. Let $\underline{u} \geq -\infty$ be the infimum of possible realizations of the investor's overall excess return \tilde{u} for any arbitrary admissible securities portfolio and $\bar{u} \leq \infty$ shall describe the supremum. Let the investor's standardized risk tolerance τ be any of a certain set $T \supset (\underline{\tau}; \bar{\tau})$, $\underline{\tau} = \inf T$, $\bar{\tau} = \sup T$ with $-\infty \leq \underline{\tau} < \bar{\tau} \leq \infty$. We say that the separation theorem holds for this special portfolio selection problem if the optimal structure y^* is the same for all $\tau \in T$. Obviously, this definition only makes sense in the case of $x_0 \neq 1$.

According to Definition 2, the separation property has to be examined for a certain (well defined) given portfolio selection problem which is characterized by intervals X , Y , and T and given excess return distributions of securities 1 and 2. The restrictions with respect to X and Y in Definition 2 assure that any positive investment in risky or riskless securities shall be admissible. There may only be some kind of short sales restrictions. In the same way we will adhere to situations with a given set $X \supset [0, 1]$ for the whole paper. The main goal of our analysis now is to find out the relevance of the two-fund separation theorem of Cass and Stiglitz (1970), when the requirement of positive marginal utility is explicitly allowed for. We do this by addressing the following two questions:

- 1) When it is possible to get the optimal value $y^{(R)*}$ (according to Cass and Stiglitz, 1970) indeed as the solution of a well defined portfolio selection problem with separation, and which values of τ can then be part of T?
- 2) How must X be defined in order to maximize the difference $\bar{\tau} - \underline{\tau}$ (i.e. the possible scope of the separation theorem) under the restriction that a given value τ is part of T?

2.2 A graphical representation

It is easy to present graphically our basic idea for the special case of quadratic utility functions. Figure 1 displays a well defined portfolio selection problem for a German investor at the end of the 1990ies¹ with a quadratic utility function and a time horizon of one month who is searching for the optimal combination of the mutual fund “INVESCO GT German Growth C” and the DAX 100 index as risky securities 1 and 2 as well as riskless lending or borrowing. The DAX 100 (listed until 03/21/2003) consisted of 100 continuously traded shares of German companies including the 30 blue chips of the DAX 30 and the 70 midcap-stocks of the MDAX. Based on historical return realizations, we assume $E(\tilde{u}_1) = 1.71649\%$, $E(\tilde{u}_2) = 1.77189\%$, $\text{Var}(\tilde{u}_1) = 0.32236\%$, $\text{Var}(\tilde{u}_2) = 0.39055\%$, $\text{Cov}(\tilde{u}_1, \tilde{u}_2) = 0.24657\%$, $\bar{u} = 44.90\%$, and $r_0 = 0.22297\%$. P^* denotes the optimal risky subportfolio as implied by the two-fund separation theorem. However, as mentioned previously, we have to restrict the investor’s access to riskless lending and borrowing in order to avoid possibilities of negative marginal utility. As a consequence, mean-variance efficient portfolios do not lie on a straight line any more, but this (bold) line is eventually curved, when the lower bound \underline{x}_0 for riskless lending becomes binding. In what follows, we assume $Y = [0, 1]$ and $\underline{x}_0 = -200\%$. The two-fund separation theorem can only hold for the investor’s risk tolerance τ not being too great so that the optimal portfolio does not lie on the curved section of the set of μ - σ -efficient

¹ For more details see the empirical section 5 below.

combinations. In fact, in Figure 1 we must have $\tau \leq \bar{\tau} = 1.5664$. Obviously, for any other lower bound \underline{x}_0 , it is possible to compute a corresponding upper bound $\bar{\tau}$ as well. However, in addition, the investor's risk tolerance must be greater than a lower bound $\underline{\tau}$, because otherwise non-positive marginal utility would result at least for high portfolio return realizations. In Figure 1, we have $\underline{\tau} = 1.4512$, so that marginal utility (as well as the slope of all corresponding indifference curves) would just be zero for the (assumed) highest possible portfolio return realization $\bar{u} = 44.90\%$. Summarizing, for a given lower bound $\underline{x}_0 = -200\%$, the separation property holds for $T = (1.4512, 1.5664]$. Moreover, we can set the upper bound for x_0 equal to one without causing any additional problems, because riskless lending beyond an upper bound $\bar{x}_0 = -138,7665\%$ will not be optimal for $\tau \in T$. However, the concept of standardized risk tolerance – though useful for analytical derivations – seems somewhat too abstract. We therefore introduce an alternative measure for an individual's risk aversion which we call the relative risk discount d . This measure is defined as the difference between 100 % and the quotient of the certainty equivalent assigned by the investor to his risky portfolio excess return and the expected excess return. For our example and $T = (1.4512, 1.5664]$, a given value of $x_0 = -200\%$ (together with the realization of the risky portfolio P^*) coincides with relative risk discounts in the – rather narrow – range $[48.79\%, 61.98\%)$. Moreover, variations of the lower bound \underline{x}_0 may lead to variations of the admissible set T of standardized risk tolerances τ and thus relative risk discounts. However, lower standardized risk tolerances than $\hat{\tau} = 1.1519$ coinciding with $x_0^* = 20.41\%$ are not possible, as for the minimal required set $X = [0, 1]$ we have $\bar{u} = 14.97\%$ so that (under consideration of $r_0 = 0.22297\%$) values $\tau < 1.1519$ would imply negative marginal utility for $u = \bar{u}$.

>>> **Insert Figure 1 about here** <<<

All these results will be shown analytically in Proposition 3 of section 3. Based on our theoretical considerations, the empirical analysis in section 5 verifies as our first practically

important result our finding of the present example that only small ranges of relative risk discounts are generally in accordance with the two-fund separation theorem and that these ranges are smaller for quadratic utility than for cubic or bi-quadratic one.

Secondly, we are able to identify situations for given return distributions as well as given riskless interest rate so that there are no sets X and T at all which support the two-fund separation theorem. Apparently, if the dashed indifference curve in Figure 1 exhibited negative marginal utility for excess return realization $\bar{u} = 45.12\%$, the separation property according to Definition 2 could not hold for $\underline{x}_0 = -200\%$. Rather interestingly, from Proposition 2 of the formal part of our analysis we will learn that in such a situation the separation property cannot hold for *any* other lower bound \underline{x}_0 , either. It is only such a portfolio selection situation with necessarily negative marginal utility where new information may reduce an investor's maximum expected utility, as is the case for the well-known bias-in-beta problem of fund management. As a second practically quite important result of our formal exposition we thus will be able to state conditions under which bias-in-beta problems will not be possible.

Thirdly, as mentioned at the beginning of this paper, the problem under consideration is not only relevant for individual portfolio optimization, but also for asset pricing theory, since the SLM-CAPM is based on the validity of the two-fund separation theorem. For illustrative purposes, assume P^* in Figure 1 to be identical to the market portfolio of the SLM-CAPM. For the SLM-CAPM to be valid in spite of riskless lending and borrowing constraint X , investors' risk tolerances must not be greater than $\bar{\tau} = 1.5664$. This is equivalent to the requirement that investors' relative risk discounts for the hypothetical case $x_0 = -200\%$ must not be smaller than 48.79%. Otherwise, the riskless borrowing constraint would be binding and the Black-CAPM would be in effect. We are able to show the remarkable fact that this critical relative risk discount is the same for *any* possible lower bound \underline{x}_0 . Summarizing, we

believe that the explicit recognition of the requirement of positive marginal utility is of interest both from an individual's point of view and from a market perspective.

3 A characterization of well defined portfolio selection problems exhibiting the separation property

3.1 Necessary and sufficient restrictions for well defined portfolio selection problems

As already mentioned, well defined portfolio selection problems are only possible if we impose additional restrictions on the set of admissible portfolios or parameter values τ . This is made more precise by

Proposition 1. *Consider an investor with a HARA utility function belonging to a certain class b . Let $\underline{u}(x_0)$ be the infimum of realizations of excess returns $\tilde{u}(x_0, y) := (1 - x_0) \cdot [y \cdot \tilde{u}_1 + (1 - y) \cdot \tilde{u}_2]$ for all admissible values of y and \underline{u} be the infimum of all $\underline{u}(x_0)$. In the same way define supremum values $\bar{u}(x_0)$ and \bar{u} . In this context, we make the plausible assumptions $\underline{u}(0) < 0$ and $\bar{u}(0) > 0$. For given restrictions $x_0 \in X$ and $y \in Y$ the investor's portfolio selection problem is well defined if and only if $\tau > -(\underline{u} + 1 + r_0) \cdot b$ for $b \geq 0$ or $\tau > -(\bar{u} + 1 + r_0) \cdot b$ for $b < 0$.*

Thereby, we have

$$\underline{u} := \min\{(1 - \bar{x}_0) \cdot \bar{u}(0), (1 - \underline{x}_0) \cdot \underline{u}(0)\} \text{ and } \bar{u} := \max\{(1 - \bar{x}_0) \cdot \underline{u}(0), (1 - \underline{x}_0) \cdot \bar{u}(0)\}.$$

Proof. See the Appendix. \square

Proposition 1 gives us the minimum required risk tolerance so that marginal utility is positive even for the highest and the smallest possible return realization for any portfolio under consideration. Not very surprisingly, in the case of constant absolute risk aversion, i.e. $b = 0$, all values $\tau > 0$ lead to well defined portfolio selection problems. In what follows we therefore will only focus on the more interesting case $b \neq 0$. For our example of Figure 1, we

have $b = -1$, $\bar{u} = \bar{u}(-200\%) = 44.90\%$, and thus $\tau > \bar{u} + 1 + r_0 = 1.4512$. As a direct consequence of Proposition 1 we may state the following

Corollary 1. *Consider an investor with a HARA utility function belonging to a certain class $b \neq 0$. For given standardized risk tolerance τ with $-\infty < \tau < \infty$ the investor's portfolio selection problem may be well defined only if $X \neq \mathfrak{R}$ and the domain of \tilde{u} is not unbounded from above and from below for any admissible portfolio. Moreover, we must have $\tau > -b \cdot (1 + r_0)$.*

Proof. If we assume $X = \mathfrak{R}$ or an unbounded domain of \tilde{u} for a certain portfolio, we obviously violate the conditions $\tau > -(\underline{u} + 1 + r_0) \cdot b$ for $b \geq 0$ or $\tau > -(\bar{u} + 1 + r_0) \cdot b$ for $b < 0$, respectively, because we then have $\underline{u} = -\infty$ and $\bar{u} = \infty$. Moreover, these conditions particularly imply $\tau > -b \cdot (1 + r_0)$, since $-b \cdot \underline{u} > 0$ for $b > 0$ and $-b \cdot \bar{u} > 0$ for $b < 0$. \square

The last part of Corollary 1 states that marginal utility must be positive at least in situations where the overall rate of portfolio return equals the riskless interest rate. In fact, for our example of Figure 1 this condition is fulfilled, since we have $\tau > 1 + r_0 = 1.0022297$. Because of the first part of Corollary 1 we must focus on situations with $\underline{x}_0 > -\infty$ and/or $\bar{x}_0 < \infty$, i.e. unlimited riskless lending and borrowing apparently is not admissible.

3.2 Optimal portfolio structure $y^{(\mathfrak{R})^*}$ in spite of restricted riskless lending and borrowing

Against the background of section 3.1, we now analyze how X and T must be designed in order to reconstitute the optimal portfolio structure $y^{(\mathfrak{R})^*}$ evolving for $X = \mathfrak{R}$ and neglecting problems of negative marginal utility. First of all, we have to examine the question under which conditions it will be possible to find non-empty sets X and T which support $y^{(\mathfrak{R})^*}$ as an optimal solution. To do so we need

Lemma 1. Consider a well defined portfolio selection problem and an investor with a HARA utility function belonging to a certain class $b \neq 0$. Let $x_0^*(\tau)$ be the optimal solution for x_0 for given restrictions X and Y as well as given $\tau \in T$ and define

$\xi_0 := (1 - x_0) / (\tau + b \cdot (1 + r_0))$ with $\tau > -b \cdot (1 + r_0) > 0$ as a substitute for x_0 . Let $\xi_0^{(\mathfrak{R})^*}$, $x_0^{(\mathfrak{R})^*}(\tau)$ and $y^{(\mathfrak{R})^*}$ be the optimal values of ξ_0 , x_0 and y , respectively, for given restrictions T and Y and $X = \mathfrak{R}$. In addition, we assume an (inner) optimal solution $y^{(\mathfrak{R})^*} \in Y \setminus \{\underline{y}, \bar{y}\}$.

1) A necessary and sufficient condition for separation to hold is that there is no subset

$[\underline{\tau}^+, \bar{\tau}^+] \subset T$ with $\bar{\tau}^+ - \underline{\tau}^+ > 0$ and $x_0^*(\tau) \in \{\underline{x}_0, \bar{x}_0\}$ for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ ($\Leftrightarrow x_0^*(\tau) \neq x_0^{(\mathfrak{R})^*}$ for all but at most one $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$). Then we have $y^* = y^{(\mathfrak{R})^*}$.

2) Recall $\xi_0^*(\tau) := (1 - x_0^*(\tau)) / (\tau + b \cdot (1 + r_0))$. Another necessary and sufficient condition

for separation to hold is $\xi_0^*(\tau) = \xi_0^{(\mathfrak{R})^*} = \text{const.}$ for all $\tau \in T$.

3) Let $\tau^{(max)} := \inf\{\tau \in \mathfrak{R} \mid x_0^*(\tau) = \underline{x}_0\}$ and $\tau^{(min)} := \sup\{\tau \in \mathfrak{R} \mid x_0^*(\tau) = \bar{x}_0\}$. In the case of

separation we have $\bar{\tau} \leq \tau^{(max)}$, if $\xi_0^* > 0$. For $\xi_0^* < 0$ we have $\underline{\tau} \geq \tau^{(min)}$.

Proof. See the Appendix. \square

Lemma 1 introduces an auxiliary decision variable ξ_0 as a function of an investor's riskless lending x_0 with the interesting feature of its optimal value being independent of standardized risk tolerance τ and corresponding optimal value x_0^* , when the two-fund separation theorem applies. For the example of section 2.2 we have $\xi_0^{(\mathfrak{R})^*} = 5.3176$. Moreover, Lemma 1 in particular states the necessity of the absence of binding restrictions with respect to riskless borrowing and lending for the separation theorem to hold in the case of inner solutions regarding y and thus enables us to answer the questions mentioned above with

Proposition 2. Consider an investor with a HARA utility function belonging to a certain class $b \neq 0$. If and only if $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*}) \leq 0$ for at least one possible realization of $\tilde{u}(0, y^{(\mathfrak{R})^*})$, then for given Y and inner solution $y^* \in Y \setminus \{\underline{y}, \bar{y}\}$ there exist no sets $T, X \subset \mathfrak{R}$ so that the separation property with $y^* = y^{(\mathfrak{R})^*}$ can be re-established. In the case of a border solution $y^* \in \{\underline{y}, \bar{y}\}$ the condition $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*}) \leq 0$ for at least one possible realization of $\tilde{u}(0, y^{(\mathfrak{R})^*})$ implies that the separation property can only be reconstituted for sets T so that the restriction $x_0 \in X$ is binding for all $\tau \in T$.

Proof. See the Appendix. \square

According to Proposition 2 there may be situations where an inner solution $y^* = y^{(\mathfrak{R})^*}$ cannot be supported by a well defined portfolio selection problem exhibiting the separation property, even if X and T are restricted in any arbitrary manner. Only if $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*}) > 0$ for all possible realizations of $\tilde{u}(0, y^{(\mathfrak{R})^*})$, separation with an inner solution $y^* = y^{(\mathfrak{R})^*}$ can be derived for adequately defined sets X and T . Indeed, in our example of section 2 we have $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot \tilde{u}(0, y^{(\mathfrak{R})^*}) \geq 49.51 \% > 0$.

Things look somewhat different in the case of border solutions $y^* = y^{(\mathfrak{R})^*} \in \{\underline{y}, \bar{y}\}$, because in such a situation binding restrictions $x_0 \in X$ may be consistent with $y^* = y^{(\mathfrak{R})^*}$ for a set $T = [\underline{\tau}^+, \bar{\tau}^+]$ with $\bar{\tau}^+ - \underline{\tau}^+ > 0$. For example, for $Y = \{y^{(\mathfrak{R})^*}\}$ it will in general be possible to find sets $X = \{x_0\}$ and $T = [\underline{\tau}^+, \bar{\tau}^+]$ with $\bar{\tau}^+ - \underline{\tau}^+ > 0$ which lead to a well defined portfolio selection problem and exhibit the separation property even if $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*}) \leq 0$. Certainly, such a situation is not very interesting. This assessment is confirmed if we take into account the requirement $[0, 1] \subset X$ stated previously. For X satisfying this condition, a nonpositive sign of $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*})$ in connection with a border solution

$y^* = y^{(\mathfrak{R})} \in \{\underline{y}, \bar{y}\}$ immediately implies according to Proposition 2 that well defined portfolio selection problems exhibiting the separation property are characterized by $x_0^* \leq 0$ for all $\tau \in T$ or $x_0^* \geq 1$ for all $\tau \in T$. Obviously, such scenarios are of only minor practical importance. Moreover, in capital market equilibrium with two-fund separation and homogenous expectations as well as $[0, 1] \subset Y$, market clearing conditions will require an inner solution for y which offers additional support for the assessment of only minor relevance of border solutions with respect to the holding of risky assets. As a consequence of Proposition 2, we therefore restrict our analysis for the rest of our paper to situations with inner² solutions for y and thus assume the condition $1 + b \cdot \xi_0^{(\mathfrak{R})*} \cdot u(0, y^{(\mathfrak{R})*}) > 0$ to be fulfilled for all possible realizations of $\tilde{u}(0, y^{(\mathfrak{R})*})$. With this in mind, we now want to analyze further which values of τ can be part of T and how X must be defined in order to maximize the difference $\bar{\tau} - \underline{\tau}$ for given $\tau \in T$.

Because of Lemma 1 binding restrictions for borrowing and lending will circumvent the separation theorem from being valid as long as we restrict our analysis to inner solutions for the optimal structure of the risky part of an investor's overall portfolio. Moreover, for given restrictions Y and X , Proposition 1 and Lemma 1 offer us lower and upper bounds for admissible values τ leading to well defined portfolio selection problems that guarantee the separation property for all $\tau \in T$ and arbitrary (not unbounded) return distributions according to Definition 2. To be precise, part 1) or 2) of Lemma 1 in connection with Proposition 1 describe necessary and sufficient conditions for a portfolio selection problem satisfying Definition 2 with inner solution $y^* = y^{(\mathfrak{R})*}$.

Note that $\xi_0^{(\mathfrak{R})*} > 0$ means that for given securities 1 and 2 as well as riskless interest rate r_0 and restrictions X and Y , all investors with a HARA utility function of the same class

² However, separation results for the borderline case of $y^{(\mathfrak{R})} \in \{\underline{y}, \bar{y}\}$ are available from the authors upon request.

will choose $x_0^{(\mathfrak{R})^*}(\tau) < 1$, whereas $\xi_0^{(\mathfrak{R})^*} < 0$ corresponds to the fact that all investors arrive at $x_0^{(\mathfrak{R})^*}(\tau) > 1$, thus selling short the optimal portfolio of risky securities 1 and 2. Obviously, a situation with $\xi_0^{(\mathfrak{R})^*} > 0$ seems to be much more important and interesting. We therefore focus on such situations:³

Proposition 3. *Assume $b \neq 0$. For given Y , the inequalities $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot \bar{u}(0) > 0$ as well as $1 + b \cdot \xi_0^{(\mathfrak{R})^*} \cdot \underline{u}(0) > 0$ shall be true. Moreover, we restrict ourselves to situations with $y^{(\mathfrak{R})^*} \in Y \setminus \{y, \bar{y}\}$ and $\xi_0^{(\mathfrak{R})^*} > 0$. Define $\underline{x}_0^*(\tau) := \min\{1 - \xi_0^{(\mathfrak{R})^*} \cdot (\tau + b \cdot (1 + r_0)), 0\}$, $\bar{x}_0^*(\tau) := \max\{1 - \xi_0^{(\mathfrak{R})^*} \cdot (\tau + b \cdot (1 + r_0)), 1\}$, and a function g with*

$$g(x_0, \tau) := \begin{cases} \tau + b \cdot ((1 - x_0) \cdot \underline{u}(0) + 1 + r_0), & \text{if } (1 - x_0) \cdot b > 0, \\ \tau + b \cdot ((1 - x_0) \cdot \bar{u}(0) + 1 + r_0), & \text{if } (1 - x_0) \cdot b < 0. \end{cases}$$

Then a critical value $\hat{\tau}$ with $g(\underline{x}_0^(\hat{\tau}), \hat{\tau}) = 0$ exists and (only) for any $\tau^+ > \hat{\tau}$ there is a set T with $\tau^+ \in T$ so that the separation property holds for given restrictions Y and $X = [\underline{x}_0^*(\bar{\tau}), 1]$ with $\underline{x}_0^*(\bar{\tau}) \leq 0$. The difference $\bar{\tau} - \underline{\tau}$ with $\tau^+ \in [\underline{\tau}, \bar{\tau}]$ is maximized by setting $\underline{\tau} = \tau^+$ and $\bar{\tau}$ so as to guarantee $g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0$. The portion $\underline{x}_0^*(\bar{\tau})$ converges to $-\infty$ for $\tau^+ \rightarrow \infty$. In addition, only for borrowing restrictions $\underline{x}_0 < \underline{x}_0^*(\hat{\tau})$ the separation theorem may hold and the set T which maximizes $\bar{\tau} - \underline{\tau}$ for given value \underline{x}_0 is (implicitly) defined by $\underline{x}_0^*(\bar{\tau}) = \underline{x}_0$ and $g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0$. Since $\xi_0^{(\mathfrak{R})^*} > 0$ corresponds with $x_0^{(\mathfrak{R})^*}(\tau) < 1$, ceteris paribus deviations from the minimum setting $\bar{x}_0 = 1$ might only reduce the admissible set T for given lower bound \underline{x}_0 .*

Proof. See the Appendix. \square

³ However, separating results are similar for $\xi_0^{(\mathfrak{R})^*} < 0$ as well. Details are available from the authors upon request.

Proposition 3 tells us the range of possible well defined portfolio selection problems for which the separation theorem may hold, if we assume inner optimal solutions for y . In particular, in any case an investor's standardized risk tolerance must exceed a certain threshold value $\hat{\tau}$ and is generally restricted to values of a bounded set T . Rather interestingly, according to Proposition 3 there is a limiting case where (approximately) unrestricted borrowing is allowed and the range of admissible values for τ is (approximately) infinite. Such a situation resembles the original separation theorem by Cass and Stiglitz (1970). Unfortunately, in contrast to Cass and Stiglitz (1970), all those admissible values for τ are (nearly) infinite. Especially this result may highlight the necessity for an explicit notion of positive marginal utility. A graphical interpretation of the main results of Proposition 3 has already been given in section 2. There we have $\underline{\tau} = 1.4512$ and the upper bound $\bar{\tau} = 1.5664$ maximizes the difference $\bar{\tau} - \underline{\tau}$ for given lower bound $\underline{x}_0 = -200\%$.

Moreover, it is straightforward to extend our analysis in Proposition 3 to more than just two risky securities, since we only consider situations with a fixed optimal structure $y^{(R)*}$ of risky portfolios. More interestingly, Proposition 3 enables us to derive conclusions regarding the practical relevance of the bias-in-beta problem and of the SLM-CAPM versus the Black-CAPM.

4 Practical Implications of missing positive marginal utility

4.1 Bias in beta in mutual fund performance evaluation

It is well-known from the theory of mutual fund performance evaluation that better information might deteriorate the assessment of a portfolio manager according to Jensen's (1968) alpha. This phenomenon is called the bias-in-beta problem and in what follows we want to analyze a generalized version of this phenomenon. To this end, we look at the situation of section 2.2 and consider an investor with HARA utility who wants to choose exactly one out of F different funds $f = 1, \dots, F$ in order to combine it optimally with a

reference portfolio P of direct stockholding (the DAX 100, for example) and riskless lending and borrowing. Propositions 2 and 3 now refer to each of the F different portfolio selection problems of this section separately, and it is straightforward to draw conclusions for the overall selection problem. In particular, choosing T and X in such a way so that any of the F sub-problems exhibits the separation property also guarantees the same characteristic for the overall portfolio problem. Though without a discussion of the problem of negative marginal utility, in Breuer and Gürtler (2006) (in particular Propositions 2 and 3) it has been shown that investors' expected (HARA) utility maximization and the two-fund separation theorem lead to a fund ranking according to some kind of generalized and optimized Sharpe measure that comprises the classical performance measures of Treynor (1965), Sharpe (1966), Jensen (1968) and Treynor and Black (1973) as special cases. Against this background, we now introduce

***Definition 3.** A situation where ceteris paribus better information in the sense of at least second-order stochastic dominance for a fund manager f may deteriorate his ranking, although – given this information – he is maximizing investors' expected utility according to the two-fund separation theorem, is called a generalized bias-in-beta problem.*

Grinblatt and Titman (1989) showed that the "conventional" bias-in-beta problem is a consequence of implicitly assuming negative marginal utility. In fact, this cause carries over to the generalized bias-in-beta problem of Definition 3. Only if marginal utility can become negative, first- or second-order-stochastic dominance may not be favorable so that "better" information might indeed not be better. This immediately gives us

***Corollary 2.** The generalized bias-in-beta problem does not occur for the ranking of mutual funds according to the optimized and generalized Sharpe measure of Breuer and Gürtler (2006), as long as there are non-empty sets X and T so that for all F sub-portfolio selection problems under consideration are well defined and exhibit the separation property.*

In particular, such non-empty sets of X and T do exist if all funds fulfil the condition $1 + b \cdot \xi_0^{(R)} \cdot u(0, y^{(R)*}) > 0$ of Proposition 2 for all possible realizations of $\tilde{u}(0, y^{(R)*})$. In case there are such non-empty sets X and T , a fund g with a return distribution of its optimal combination with the reference portfolio P that dominates the corresponding "optimized" return distribution of a fund h according to first- or second-order stochastic dominance will attain a better ranking position.*

Proof. See derivation above: Since the generalized and optimized Sharpe measure leads to the choice of that fund that maximizes investors' expected utility and all fund managers' portfolio selection problems are well defined, better information cannot deteriorate a fund's ranking position. Thereby, it is not necessary to restrict the analysis to inner solutions for y , as in the case of border solutions the only difference lies in the possibility of well defined portfolio selection problems exhibiting the separation property even if $1 + b \cdot \xi_0^{(R)*} \cdot u(0, y^{(R)*}) > 0$ is not true for all of $\tilde{u}(0, y^{(R)*})$. \square

According to Corollary 2, the bias-in-beta-problem is not a weakness of a certain performance measure like Jensen's alpha. Instead, it is the result of implicitly considering portfolio selection problems which cannot fulfill the requirement of positive marginal utility so that the whole approach of performance evaluation based on the two-fund separation theorem becomes invalid. Moreover, Corollary 2 offers an alternative to the suggestion of Grinblatt and Titman (1989) to define so-called positive period weighting measures in order to circumvent the bias-in-beta problem. Instead of this, one could restrict oneself to investors' preferences (like always the negative exponential one) which enable the derivation of non-empty sets X and T leading to a well defined portfolio selection problem that exhibits the separation property. Then, funds can be ranked according to the approach presented by Breuer and Gurtler (2006) thus giving one's ranking a clear portfolio theoretic foundation – something which in general cannot be assured by positive period weighting measures.

4.2 SLM-CAPM versus Black-CAPM

In the SLM-CAPM, the optimal portfolio P^* of risky securities is an inner solution to the investors' portfolio selection problem because of market clearing conditions and the assumption of homogeneous expectations. Moreover, as is known from Black (1972), in the absence of a riskless asset, all mean-variance efficient portfolios can be interpreted as the linear combination of just two given risky portfolios, i.e. return distributions, so that the formal analysis of the previous section 3 directly applies to the equilibrium situation in the SLM-CAPM if we denote these two risky portfolios as securities 1 and 2. However, the SLM-CAPM is only valid if there are no binding constraints for riskless borrowing and lending. From Lemma 1 and the graphical analysis of section 2 it is known that investors' standardized risk tolerance τ therefore must not exceed a certain threshold $\tau^{(\max)}(\underline{x}_0)$. Otherwise, restriction X is binding and thus the Black-CAPM applies. It is therefore easy to compute restrictions on τ which are necessary for the validity of the SLM-CAPM. However, while analytically helpful, the variable τ lacks some vividness so that we focus instead on the examination of minimum required relative risk discounts d . In fact, for any arbitrary given holding x_0 of the riskless asset and maximum standardized risk tolerance, the corresponding (minimum) relative risk discount can be computed. To be precise, we have

Proposition 4. *Assume a well defined portfolio selection problem with $b \neq 0$.⁴ An investor's relative risk discount d_b for given preference parameter b does not depend on x_0 and standardized risk tolerance τ , as long as $\xi_0 := (1 - x_0) / (\tau + b \cdot (1 + r_0))$ and y are unchanged.*

Proof. See the Appendix. \square

Among other things, Proposition 4 implies that in the case of inner solutions with respect to riskless lending and borrowing relative risk discounts are independent of

⁴ Though not of interest here, it should be noted that Proposition 4 holds for $b = 0$ as well.

standardized risk tolerance τ , because optimal values $\xi_0^{(\mathfrak{R})^*}$ and $y^{(\mathfrak{R})^*}$ do not depend on τ . This gives

Corollary 3. *Assume a situation with all investors exhibiting quadratic utility functions and homogenous expectations. Moreover, let the portfolio selection problem of any investor be well defined in the sense of Definition 1 with $\underline{x}_0 = 0$. Then for given values $E[\tilde{u}]$ and $Var[\tilde{u}]$ of expected market portfolio return and corresponding return variance, compute the corresponding certainty equivalent u_{CE} for quadratic utility and $x_0 = 0$ as*

$$\begin{aligned} u_{CE}(\tau) &= -\left\{E[\tau - (1 + r_0 + \tilde{u})]^2\right\}_{b=-1}^{1/2} + \tau - 1 - r_0 \\ &= -\left\{\tau^2 - 2 \cdot \tau \cdot (1 + r_0 + E[\tilde{u}]) + Var[\tilde{u}] + (1 + r_0 + E[\tilde{u}])^2\right\}^{1/2} + \tau - 1 - r_0. \end{aligned}$$

The SLM-CAPM is only valid if and only if there is no investor with a relative risk discount that is smaller than $1 - u_{CE}(\tau^{(max)}(\underline{x}_0 = 0)) / E[\tilde{u}]$ in the (hypothetical) case of solely holding the market portfolio (and with $\tau^{(max)}$ as defined in Lemma 1) 3) for $\underline{x}_0 = 0$). Otherwise, $x_0 = 0$ is binding and only the Black-CAPM holds true. For other lower bounds of x_0 , the same critical relative risk discount results.

Proof. Follows immediately from Lemma 1 3) and Proposition 4. \square

For example, for the time period from 1951 to 2000 in the U.S., Fama and French (2002) offer three estimators of (real) equity risk premia on the market portfolio, i.e. an estimator based on historically realized stock returns amounting to 7.43 % p.a., while the other two estimators are based on the historical development of fundamentals (dividends or earnings) and amount to 2.55 % or 4.32 % p.a., respectively. Moreover, corresponding Sharpe ratios of the market portfolio for these three approaches are presented as 44 %, 15 %, and 25 %, respectively, thus showing a strong difference between the first approach and the two others. From given equity risk premia and corresponding Sharpe ratios we are able to calculate variances of excess returns of 2.8514 %, 2.89 %, and 2.986 %, respectively. With

these few data and an average (real) riskless interest rate of 2.19 % p.a., also according to Fama and French (2002), it is already possible to calculate minimum required relative risk discounts which are necessary to avoid a binding arbitrary lower bound \underline{x}_0 . In fact, we arrive at minimum required relative risk discounts of 47.79 %, 49.24 %, and 49.72 %, respectively. If there is only one individual (with quadratic utility function) whose relative risk discount when (fictitiously) realizing $x_0 = \underline{x}_0$ (and the market portfolio as the risky investment) is less than 47 %, then the valuation function derived by Black (1972) becomes relevant, as long as we restrict ourselves to the consideration of well defined portfolio selection problems according to Definition 1.

5 Mutual funds investing in German stocks and the separation theorem

As an additional illustration we analyze the special mutual fund portfolio selection problem of an investor outlined at the beginning of section 4.1 and already examined in Breuer and Görtler (2005, 2006) (though without explicit recognition of the requirement of positive marginal utility) with $Y = [0, 1]$ and $F = 45$ different mutual funds that are investing on the German stock market.

We start by assuming the investor's utility function to be quadratic. With the two-fund separation being valid, all funds then are ranked according to the respective Sharpe ratios of their best combination with the DAX 100 reference portfolio of direct stock holding, i.e. according to the optimized Sharpe measure as defined in Breuer and Görtler (2006). Table 1 presents excess return moments according to Breuer and Görtler (2005) and (in addition) minimum and maximum realized monthly fund returns for the time period from July 1996 to August 1999 under consideration. Moreover, we follow Breuer and Görtler (2005) in assuming a riskless interest rate r_0 of 0.22297 %.

>>> Insert Table 1 about here <<<

Let $\xi_{0,f}^{(\mathfrak{R})^*}$ be the optimal value of ξ_0 as defined in Lemma 1 for a certain fund f in the case of quadratic utility. In what follows, variables that are depending on characteristics of fund f are indexed by f . We calculate $\xi_{0,f}^{(\mathfrak{R})^*}$ for all funds f under consideration. Two of them do not satisfy the condition $1 - \xi_0^{(\mathfrak{R})^*} \cdot u(0, y^{(\mathfrak{R})^*}) > 0$ for all possible excess return realizations $u(0, y^{(\mathfrak{R})^*}) > 0$ and thus have to be excluded in order to guarantee a well defined portfolio selection problem exhibiting the separation property. For all other 43 funds, values for $\xi_{0,f}^{(\mathfrak{R})^*}$ are positive, so that situations with short selling of risky portfolios do not appear for our example. As a consequence, according to section 3.2 we can immediately conclude that an upper bound $\bar{x}_0 = 1$ is admissible. For 24 funds of the remaining 43 ones we get border solutions $y_f^{(\mathfrak{R})^*} = 0$, i.e. these funds will not be combined positively with the DAX 100 and thus shall be omitted. 12 more funds are characterized by a binding restriction $y_f^{(\mathfrak{R})^*} = 1$, so that these funds will be held without any additional investment in the DAX 100. As we restrict ourselves in this paper to situations with inner solutions with respect to y , these funds are omitted as well. We will return to this issue later on. Summarizing, there are only seven funds left that are considered in more detail and thus are shaded in gray in Table 1.

Now, we assume different (alternative) lower bounds $\underline{x}_0 \in \{0, -1, -2\}$. For each of these three lower bounds and for all of the remaining seven funds, we compute the interval T_f of risk tolerances according to Proposition 3, which support the optimal risky holding $y_f^{(\mathfrak{R})^*}$. The first part of Table 2 displays all resulting intervals T_f as well as their intersections $(\underline{\tau}^{(\max)}, \bar{\tau}^{(\min)})$ for the case of quadratic utility ($b = -1$) and $\underline{x}_0 \in \{0, -1, -2\}$. According to Table 2, resulting intersections become greater with decreasing lower bound \underline{x}_0 , but generally seem to be quite small. Nevertheless, it remains difficult to assess the relevance of the range of possible degrees of risk aversion that are consistent with the existence of a well defined portfolio selection problem according to Definition 1. In addition, we therefore

calculate relative risk discounts for $x_{0,f} = \underline{x}_0$, $y_f = y_f^{(\mathfrak{R})^*}$ and $\tau = \underline{\tau}^{(\max)}$ or $\tau = \overline{\tau}^{(\min)}$, respectively. Maybe somewhat surprisingly, we find that all intervals of resulting relative risk discounts are independent of the value \underline{x}_0 under consideration. This is caused by the fact, that the same fund ($f = 12$: DekaFonds) is responsible for the determination of the lower bound $\underline{\tau}^{(\max)}$ of the resulting intersection of all intervals T_f regardless of \underline{x}_0 . In the same way, the fund $f = 29$: “INVESCO GT German Growth C” always determines the upper bound $\overline{\tau}^{(\min)}$ of the intersection of all T_f . As a consequence, for any τ under consideration we have $x_{0,f=12}^*(\tau) = \underline{x}_0$ and thus constant values $\xi_{0,f=12} = \xi_{0,f=12}^{(\mathfrak{R})^*}$ and $y_{f=12} = y_{f=12}^{(\mathfrak{R})^*}$, so that Proposition 4 directly applies with respect to the relative risk discounts for $f = 7$ and varying values of $x_{0,f}$. Moreover, as we use the same values for $\underline{\tau}^{(\max)}$ and $x_{0,f=12}^* = \underline{x}_0$ for all funds under consideration, we have, by construction, $\xi_{0,f} = \xi_{0,f=12}^{(\mathfrak{R})^*} = \text{const.}$ for any fund f under consideration. As $y_f = y_f^{(\mathfrak{R})^*}$ is constant as well for any fund under consideration, upper bounds for relative risk discounts do not change with varying \underline{x}_0 . The same holds true regarding lower bounds of relative risk discounts caused by the relevance of fund #29.

>>> **Insert Table 2 about here** <<<

Summarizing, for any fund f the ultimate column of Table 2 offers the maximum as well as the minimum relative risk discounts for $x_{0,f} = \underline{x}_0$ and $y_f = y_f^{(\mathfrak{R})^*}$ which do not depend on the choice $\underline{x}_0 \in \{0, -1, -2\}$. Moreover, resulting intervals of relative risk discounts are quite “narrow”, as they amount at most to twelve percentage points. After all, for this empirical example and the assumption of quadratic utility the consequences of the two-fund separation as indicated by the resulting ranking positions of the funds in the third column are only valid for rather a small range of admissible standardized risk tolerances.

Correspondingly, the underlying “optimized” Sharpe ratios (that – for quadratic utility – lead to the same fund ranking as the Treynor and Black, 1973, appraisal ratio, as shown by Jobson

and Korkie, 1984) are far from generally applicable even under the assumption of quadratic utility. Certainly, intervals of admissible risk tolerances would even be (considerably) smaller if we accounted for all 43 funds of Table 1 that satisfy the condition $1 - \xi_0^{(3)*} \cdot u(0, y^{(3)*}) > 0$ for all possible excess return realizations $u(0, y^{(3)*}) > 0$ (thus including funds with border solutions $y^{(3)*} \in \{0, 1\}$).⁵

The second and the third part of Table 2 present corresponding results for cubic ($b = -0.5$) as well as bi-quadratic ($b = -1/3$) HARA utility. While minimum relative risk discounts are nearly the same as for quadratic utility, absolutely lower values for b increase significantly maximum relative risk discounts which are in line with a well defined portfolio selection problem exhibiting the separation property. Even cubic utility instead of quadratic one makes relative risk discounts beyond 100 % possible which – in the context of Table 2 – means that the investor would be willing to pay for not being obliged to realize a riskless holding of (only) $x_0 = \underline{x}_0$ together with $y_f = y_f^{(3)*}$. The reason for this extension of ranges of admissible relative risk discounts is that $\tau + b \cdot (1 + r_0 + \tilde{u})$ determines the sign of an investor's marginal utility. Absolutely smaller values of b make smaller values of τ compatible with the requirement $\tau + b \cdot (1 + r_0 + \tilde{u}) > 0$ and thus imply that higher relative risk discounts become admissible.

Although admissible domains of τ are quite different for different utility functions, one may use quadratic utility functions as a second-order approximation of higher-order utility functions in order to identify more easily reasonable portfolio structures (see, for example, Samuelson, 1970). In fact, one can determine optimal values $\xi_{0,f,b}^*$ and $y_{f,b}^*$ for $b = -1$ and use this as an approximation for the “true” optimal values of these variables in the case of $b = -0.5$ or $b = -1/3$. Because of the cardinality of utility functions it is then possible to compute

⁵ As available upon request, intervals of admissible relative risk discounts shrink to about 3 percentage points.

the resulting investor's certainty equivalent for "approximative" portfolio composition $(\xi_{0,f,-1}^*, y_{f,-1}^*)$ as a fraction of the investor's certainty equivalent when realizing the "actually" optimal portfolio structure $(\xi_{0,f,b}^*, y_{f,b}^*)$ for $b = -0.5$ or $b = -1/3$. This fraction does not depend on the standardized risk tolerance τ . The value of τ only is relevant in order to determine the corresponding riskless holding x_0 as a function of $\xi_{0,f,b}^*$ according to

$$x_0(\xi_{0,f,b}^*) = 1 - \xi_{0,f,b}^* \cdot (\tau + b \cdot (1 + r_0)).$$

For the seven funds of Table 2 all those fractions of certainty equivalents are almost identical to 100 %. In fact, the lowest value amounts to 99.84 %. Regardless of possible variations in fund rankings for different values of b , this finding indicates the very good approximative quality of optimal portfolios in the case of quadratic utility when preferences are indeed non-quadratic. That (numerical) property of quadratic utility functions is well documented in the literature (see, e.g., Levy and Markowitz, 1979, and Kroll et al., 1984). However, the present paper now suggests the – at first glance somewhat surprising result – that non-quadratic utility functions may be important, because they enlarge the set of possible standardized risk tolerances which are compatible with the separation requirement. As a practically remarkable consequence, the range of possible applications of optimal portfolio structures from quadratic utility goes beyond the intervals of relative risk discounts computed for $b = -1$. That means that in general the Sharpe ratio of optimal risky portfolios and the (equivalent) Treynor and Black appraisal ratio underlying the ranking of funds according to Table 2, case $b = -1$, can be justified quite well as a sorting means for funds in the case of high relative risk discounts (and higher-order HARA utility functions). However, lower relative risk discounts (for assumed riskless lending $x_0 = \underline{x}_0$) than about 50 % to 60 % are not in line with the postulation of a well defined portfolio selection problem exhibiting the separation property. At least up to our best knowledge, such questions do not seem to have been extensively discussed by the existing literature.

6 Summary and outlook

This paper was mainly motivated by the problem that the requirement of positive marginal utility restricts the possibility to derive a two-fund separation theorem as presented by Cass and Stiglitz (1970). We showed that the explicit recognition of positive marginal utility only makes it possible to derive restricted separation properties for portfolio selection problems. We characterized in detail for which portfolio selection problems separation holds and presented two practical applications regarding an analysis of the bias-in-beta problem of mutual fund performance evaluation and the validity of the SLM-CAPM versus the Black-CAPM. Moreover, we gave a simple empirical application of our results that emphasizes our assessment of the only limited validity of the separation theorem and its implications for mutual fund ranking in the light of the requirement of positive marginal utility. Moreover, despite the general good approximative quality of quadratic utility for portfolio optimization in the case of higher-order utility functions, we found the interesting result that higher-order HARA utility functions may lead to a greater range of standardized risk tolerances that imply well defined portfolio selection problems exhibiting the separation property.

As an interesting task for further research one could try to implement constraints of positive marginal utility in the capital asset pricing model of Vanden (2004) that is based on HARA utility as well and assumes (up to now) only nonnegativity constraints for an investor's terminal wealth. Maybe such an extended asset pricing model would do empirically even better than the original approach of Vanden (2004).

Appendix

Proof of Proposition 1:

Since the postulated statement is obvious in the case $b = 0$, we only deal with the case $b \neq 0$.

First of all we can show that for an arbitrary utility function of the HARA type and any wealth level W_1 the investor's utility is defined and his marginal utility is positive, if and only if we have $a + b \cdot W_1 > 0$.⁶

Case 1: $0 \neq b \neq 1$

$$U'(W_1) = (a + b \cdot W_1)^{-1/b} > 0 \Leftrightarrow a + b \cdot W_1 > 0. \quad (\text{A1})$$

Case 1: $b = 1$

$$U'(W_1) = \frac{1}{a + b \cdot W_1} > 0 \Leftrightarrow a + b \cdot W_1 > 0. \quad (\text{A2})$$

Obviously, the property $a + b \cdot W_1 > 0$ also guarantees W_1 to be in the domain of U .

From this statement we know that for each realization $u(0,y)$ we need

$$\begin{aligned} \tau + b \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y)) &> 0 \\ \Leftrightarrow \tau &> -b \cdot (1 - x_0) \cdot u(0, y) - b \cdot (1 + r_0) \end{aligned} \quad (\text{A3})$$

since $W_1 = W_0 \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y))$. In addition, since we have assumed $\underline{u}(0) < 0$ and $\bar{u}(0) > 0$, we get the following inequalities:

$$\begin{aligned} (1 - x_0) \cdot u(0, y) &\geq \min \left\{ \underbrace{(1 - \bar{x}_0)}_{\leq 0} \cdot \underbrace{\bar{u}(0)}_{\geq 0}, \underbrace{(1 - \underline{x}_0)}_{\geq 0} \cdot \underbrace{\underline{u}(0)}_{\leq 0} \right\} = \underline{u}, \\ (1 - x_0) \cdot u(0, y) &\leq \max \left\{ \underbrace{(1 - \bar{x}_0)}_{\leq 0} \cdot \underbrace{\underline{u}(0)}_{\leq 0}, \underbrace{(1 - \underline{x}_0)}_{\geq 0} \cdot \underbrace{\bar{u}(0)}_{\geq 0} \right\} = \bar{u}. \end{aligned} \quad (\text{A4})$$

Summarizing, it follows from (A3) that positive marginal utility is assured if and only if

⁶ In fact, this condition additionally guarantees that the second derivative of an investor's utility function of the HARA type is negative thus exhibiting diminishing positive marginal utility.

$$\begin{aligned} \tau &> -(\underline{u} + 1 + r_0) \cdot b \text{ for } b > 0, \text{ or} \\ \tau &> -(\bar{u} + 1 + r_0) \cdot b \text{ for } b < 0. \end{aligned} \tag{A5}$$

Proof of Lemma 1:

1) Firstly, we assume the existence of a subinterval $[\underline{\tau}^+, \bar{\tau}^+] \subset T$ with $\bar{\tau}^+ - \underline{\tau}^+ > 0$ and

$x_0^*(\tau) \in \{\underline{x}_0, \bar{x}_0\}$ for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$. For our problem this assumption apparently implies

$x_0^*(\tau) = x_0^{(\text{const})}$ for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ with $x_0^{(\text{const})} \in \{\underline{x}_0, \bar{x}_0\}$. For HARA utility functions with

$b \neq 0$, i.e. non-constant absolute risk aversion, and all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ the inner optimal solution

$y^* = y^*(\tau)$ is determined by solving the following necessary condition:

$$\begin{aligned} E((a + b \cdot W_0 \cdot (1 + r_0 + \tilde{u}(x_0^{(\text{const})}, y^*(\tau))))^{-1/b} \cdot (\tilde{u}_1 - \tilde{u}_2)) &= 0 \\ \Leftrightarrow E((\tau + b \cdot (1 + r_0 + \tilde{u}(x_0^{(\text{const})}, y^*(\tau))))^{-1/b} \cdot (\tilde{u}_1 - \tilde{u}_2)) &= 0. \end{aligned} \tag{A6}$$

Assume $y^*(\tau)$ to be constant for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$. With the abbreviations

$\tilde{u} := \tilde{u}(x_0^{(\text{const})}, y^*)$ and $\hat{h}(\tau) := E((\tau + b \cdot (1 + r_0 + \tilde{u}))^{-1/b} \cdot (\tilde{u}_1 - \tilde{u}_2)) = 0$ for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ we

get that all derivatives $\hat{h}^{(k)}(\tau) = -\frac{1}{b} \cdot (-\frac{1}{b} - 1) \cdot \dots \cdot (-\frac{1}{b} - k + 1) \cdot E((\tau + b \cdot (1 + r_0 + \tilde{u}))^{-1/b-k} \cdot$

$(\tilde{u}_1 - \tilde{u}_2))$ are zero for all $\tau \in (\underline{\tau}^+, \bar{\tau}^+)$. Since \hat{h} is continuous, the set $\hat{h}^{-1}(\{0\}) = \{\tau \mid \hat{h}(\tau) = 0\}$

is closed and thus only two cases are possible. Either there exists a maximum

$\tau_0 = \max(\hat{h}^{-1}(\{0\})) \in \mathfrak{R}$ or $\hat{h}^{-1}(\{0\})$ is unbounded, i.e. $\sup(\hat{h}^{-1}(\{0\})) = \infty$. Firstly, we

consider the case of a real maximum τ_0 . Define $\kappa := \max\{|1/b|, 2\} > 0$, $\varepsilon := \bar{\tau}^+ - \underline{\tau}^+ > 0$ and

$\tau := \tau_0 + (\varepsilon / \kappa) > \tau_0$. Since $\underline{\tau}^+ + b \cdot (1 + r_0 + \tilde{u}) > 0$ and $\tau_0 \geq \bar{\tau}^+$, we immediately obtain

$\hat{\tau} + b \cdot (1 + r_0 + \tilde{u}) > \varepsilon$ for all $\hat{\tau} > \tau_0$. Consequently, we know from Taylor's theorem for

arbitrary $k \in \mathbb{N}$ the existence of $\hat{\tau} \in (\tau_0, \tau)$ with

$$\begin{aligned}
|\hat{h}(\tau)| &= \left| \sum_{i=0}^{k-1} \frac{\hat{h}^{(i)}(\tau_0)}{i!} \cdot (\tau - \tau_0)^i + \frac{\hat{h}^{(k)}(\hat{\tau})}{k!} \cdot (\tau - \tau_0)^k \right| \\
&= \left| \mathbb{E} \left(\frac{-\frac{1}{b} \cdot (-\frac{1}{b} - 1) \cdot \dots \cdot (-\frac{1}{b} - k + 1)}{k!} \cdot (\hat{\tau} + b \cdot (1 + r_0 + \tilde{u}))^{-1/b-k} \cdot (\tilde{u}_1 - \tilde{u}_2) \right) \cdot (\varepsilon / \kappa)^k \right| \\
&\leq \mathbb{E} \left(\left| \frac{-\frac{1}{b} \cdot (-\frac{1}{b} - 1) \cdot \dots \cdot (-\frac{1}{b} - k + 1)}{k!} \right| \cdot \underbrace{\left(\frac{\varepsilon / \kappa}{\hat{\tau} + b \cdot (1 + r_0 + \tilde{u})} \right)^k}_{\in (0, 1/\kappa) \text{ since } \hat{\tau} + b \cdot (1 + r_0 + \tilde{u}) > \varepsilon} \cdot \left| (\hat{\tau} + b \cdot (1 + r_0 + \tilde{u}))^{-1/b} \cdot (\tilde{u}_1 - \tilde{u}_2) \right| \right) \quad (A7) \\
&\stackrel{c := (b-1)/b}{\leq} \mathbb{E} \left(\left| \frac{1-c}{\kappa} \right| \cdot \left| \frac{1-\frac{c}{2}}{\kappa} \right| \cdot \dots \cdot \left| \frac{1-\frac{c}{k}}{\kappa} \right| \cdot \left| (\hat{\tau} + b \cdot (1 + r_0 + \tilde{u}))^{-1/b} \cdot (\tilde{u}_1 - \tilde{u}_2) \right| \right) \xrightarrow{k \rightarrow \infty} 0.
\end{aligned}$$

The asserted convergence is obvious, since

$$\begin{aligned}
\left| \frac{1-c}{\kappa} \right| \cdot \left| \frac{1-\frac{c}{2}}{\kappa} \right| \cdot \dots \cdot \left| \frac{1-\frac{c}{k}}{\kappa} \right| &< \underbrace{\left| \frac{1-\frac{c}{2}}{\kappa} \right|^{k-1}}_{< |1-c|/\kappa \leq 1} \xrightarrow{k \rightarrow \infty} 0, \quad \text{if } c < 0, \\
\left| \frac{1-c}{\kappa} \right| \cdot \left| \frac{1-\frac{c}{2}}{\kappa} \right| \cdot \dots \cdot \left| \frac{1-\frac{c}{k}}{\kappa} \right| &< \underbrace{\left| \frac{1}{\kappa} \right|^k}_{\leq 1/2} \xrightarrow{k \rightarrow \infty} 0, \quad \text{if } c > 0.
\end{aligned} \quad (A8)$$

Summarized, we get $\hat{h}(\tau) = 0$ which contradicts $\tau_0 = \max(\hat{h}^{-1}(\{0\})) \in \mathfrak{R}$. Thus, $\sup(\hat{h}^{-1}(\{0\})) = \infty$ which implies the existence of a sequence $(\tau_n)_{n \in \mathbb{N}}$ with $\lim_{n \rightarrow \infty} \tau_n = \infty$ and $\hat{h}(\tau_n) = 0$ for all $n \in \mathbb{N}$. This statement contradicts (A6) for all distributions with $E[\tilde{u}_1] \neq E[\tilde{u}_2]$ since $\lim_{\tau \rightarrow \infty} (\hat{h}(\tau) \cdot \tau^{1/b}) = E[\tilde{u}_1 - \tilde{u}_2] \neq 0$. Thus, the assumption of a constant y^* is not true, i.e. a change in $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ implies different values for y^* . This statement corresponds with the fact that separation does not hold.

Secondly, the assumption $x_0^*(\tau) \in (x_0, \bar{x}_0)$ for all $\tau \in T \setminus \{\underline{\tau}, \bar{\tau}\}$ leads to the validity of the separation theorem according to the analysis presented in Cass and Stiglitz (1970).

2) Let U be a HARA utility function with parameters a and b and \hat{U} be the corresponding one with parameters $a = 1$ and b . With the identity

$$\begin{aligned}
& a + b \cdot W_0 \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y)) \\
&= (a + b \cdot W_0 \cdot (1 + r_0)) \cdot \left(1 + \frac{1 - x_0}{\underbrace{\tau + b \cdot (1 + r_0)}_{=\xi_0}} \cdot b \cdot u(0, y) \right) \tag{A9} \\
&= \underbrace{(a + b \cdot W_0 \cdot (1 + r_0))}_{>0} \cdot (1 + b \cdot \xi_0 \cdot u(0, y)),
\end{aligned}$$

we get $U(W_0 \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y))) = (a + b \cdot W_0 \cdot (1 + r_0))^{1-1/b} \cdot \hat{U}(\xi_0 \cdot u(0, y))$ if $0 \neq b \neq 1$, and $U(W_0 \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y))) = \ln(a + b \cdot W_0 \cdot (1 + r_0)) + \hat{U}(\xi_0 \cdot u(0, y))$ if $b = 1$.

From the cardinality of utility functions we are allowed to maximize $E[\hat{U}(\xi_0 \cdot u(0, y))]$ with respect to ξ_0 and y instead of maximizing $E[U(W_0 \cdot (1 + r_0 + (1 - x_0) \cdot u(0, y)))]$ with respect to x_0 and y .

If there are no binding restrictions for $x_0^*(\tau)$ (i.e. $x_0^*(\tau) = x_0^{(\mathfrak{R})*}(\tau) \in (\underline{x}_0, \bar{x}_0)$) the solution $\xi_0^* = \xi_0^{(\mathfrak{R})*}$ of the optimization problem is obviously independent of τ . But in the case $x_0^*(\tau) = x_0^{(\text{const})} \in \{\underline{x}_0, \bar{x}_0\}$ for $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$ the solution $\xi_0^*(\tau) = (1 - x_0^{(\text{const})}) / (\tau + b \cdot (1 + r_0))$ is strictly decreasing or strictly increasing for all $\tau \in [\underline{\tau}^+, \bar{\tau}^+]$. From part 1) we immediately get the postulated equivalence.

3) Using the identity $x_0^*(\tau) = 1 - \xi_0^* \cdot (\tau + b \cdot (1 + r_0))$ the function $x_0^*(\tau)$ is strictly decreasing in τ if $\xi_0^* = \xi_0^{(\mathfrak{R})*} > 0$ and strictly increasing if $\xi_0^{(\mathfrak{R})*} < 0$. Now assume $\xi_0^{(\mathfrak{R})*} > 0$. For separation to be true we need $x_0^*(\tau) > \underline{x}_0$ for all $\tau \in T \setminus \{\bar{\tau}\}$. This in turn implies the requirement $\tau < \tau^{(\text{max})}$ for all $\tau \in T \setminus \{\bar{\tau}\}$ and consequently $\bar{\tau} \leq \tau^{(\text{max})}$ if $\xi_0^{(\mathfrak{R})*} > 0$. In the same way it follows $\tau > \tau^{(\text{min})}$ for all $\tau \in T \setminus \{\underline{\tau}\}$ and thus $\underline{\tau} \geq \tau^{(\text{min})}$ if $\xi_0^* < 0$.

Proof of Proposition 2:

From Lemma 1, part 2), we know that the condition $\xi_0^*(\tau) = \xi_0^{(\mathfrak{R})*}(\tau) = \text{const.}$ is necessary and

sufficient for the separation theorem to hold if the portfolio problem is well defined. Thus, we only have to show that the fact of a well defined portfolio problem corresponds with

$1 + b \cdot \xi_0^{(\mathfrak{R})*} \cdot u(0, y) > 0$ for all $y \in Y$. By multiplying (A3) with W_0 which is equivalent with the fact of a well defined portfolio selection problem we get from (A9) for all $y \in Y$

$a + b \cdot W_0 \cdot (1 + r_0 + (1 - x_0^{(\mathfrak{R})*}(\tau)) \cdot u(0, y)) > 0 \Leftrightarrow 1 + b \cdot \xi_0^{(\mathfrak{R})*} \cdot u(0, y) > 0$. Consequently, the maintained equivalence is shown.

In addition, let ξ_0^* be the optimal ξ_0 -value in the restricted case $X \neq \mathfrak{R}$. As a result of (A9) it is only possible to reconstitute the separation property if $1 + b \cdot \xi_0^* \cdot u(0, y^{(\mathfrak{R})*}) > 0$.

Thus, the assumption $1 + b \cdot \xi_0^{(\mathfrak{R})*} \cdot u(0, y^{(\mathfrak{R})*}) \leq 0$ leads to $\xi_0^{(\mathfrak{R})*} \neq \xi_0^*$ and $x_0^{(\mathfrak{R})*} \neq x_0^*$,

respectively. The latter statement implies $x_0^{(\mathfrak{R})*} \notin X$, so that the restriction $x_0 \in X$ is binding.

Proof of Proposition 3:

According to Proposition 2 the following condition has to be fulfilled to guarantee separation:

$$\begin{aligned}
 & 1 + b \cdot \xi_0^* \cdot u(0, y) > 0 \\
 \Leftrightarrow & 1 + b \cdot \frac{1 - x_0^*(\tau)}{\tau + b \cdot (1 + r_0)} \cdot u(0, y) > 0 \\
 \Leftrightarrow & (1 - x_0^*(\tau)) \cdot u(0, y) \cdot b > -\tau - b \cdot (1 + r_0) \\
 \Leftrightarrow & \tau > -b \cdot ((1 - x_0^*(\tau)) \cdot u(0, y) + 1 + r_0).
 \end{aligned} \tag{A10}$$

Moreover, the case $\xi_0^* > 0$ obviously corresponds with the case $1 - x_0^*(\tau) > 0$.

Case 1: $b > 0$

$$(A10) \Leftrightarrow \tau > -b \cdot ((1 - \underline{x}_0^*(\tau)) \cdot \underline{u}(0) + 1 + r_0) \Leftrightarrow g(\underline{x}_0^*(\tau), \tau) > 0. \tag{A11}$$

Case 2: $b < 0$

$$(A10) \Leftrightarrow \tau > -b \cdot ((1 - \underline{x}_0^*(\tau)) \cdot \bar{u}(0) + 1 + r_0) \Leftrightarrow g(\underline{x}_0^*(\tau), \tau) > 0. \tag{A12}$$

Since $\underline{x}_0^*(\tau) = \min\{1 - \xi_0^* \cdot (\tau + b \cdot (1 + r_0)), 0\}$ it follows

$$\frac{dg(\underline{x}_0^*(\tau), \tau)}{d\tau} = \begin{cases} 1+b \cdot \xi_0^* \cdot \underline{u}(0) > 0, & \text{if } b > 0 \text{ and } 1 - \xi_0^* \cdot (\tau + b \cdot (1 + r_0)) < 0, \\ 1+b \cdot \xi_0^* \cdot \bar{u}(0) > 0, & \text{if } b < 0 \text{ and } 1 - \xi_0^* \cdot (\tau + b \cdot (1 + r_0)) < 0, \\ 1, & \text{if } 1 - \xi_0^* \cdot (\tau + b \cdot (1 + r_0)) \geq 0. \end{cases} \quad (\text{A13})$$

(A13) implies the existence of a parameter $\delta > 0$ so that $dg(\underline{x}_0^*(\tau), \tau)/d\tau > \delta$ for all τ and thus $\lim_{\tau \rightarrow \infty} g(\underline{x}_0^*(\tau), \tau) = \infty$ and $\lim_{\tau \rightarrow -\infty} g(\underline{x}_0^*(\tau), \tau) = -\infty$, respectively. This leads to the existence of a unique parameter $\hat{\tau}$ with $g(\underline{x}_0^*(\hat{\tau}), \hat{\tau}) = 0$. In addition, we get from (A11) and (A12) $\partial g / \partial x_0 > 0$ if $x_0 < 1$ and $\partial g / \partial x_0 < 0$ if $x_0 > 1$. With the additional obvious property $\partial g / \partial \tau > 0$ we get from (A13) and $g(\underline{x}_0^*(\hat{\tau}), \hat{\tau}) = 0$ the statement

$$g(x_0, \hat{\tau}) > 0 \text{ for all } x_0 \in (\underline{x}_0^*(\hat{\tau}), 1] \text{ and } g(\underline{x}_0^*(\hat{\tau}), \tau) > 0, \quad g(\underline{x}_0^*(\tau), \tau) > 0 \text{ for all } \tau > \hat{\tau}. \quad (\text{A14})$$

Let $\tau^+ > \hat{\tau}$ be a given parameter. With $g(\underline{x}_0^*(\hat{\tau}), \hat{\tau}) = 0$ and $dg(\underline{x}_0^*(\tau), \tau)/d\tau > 0$ we get

$g(\underline{x}_0^*(\tau^+), \tau^+) > 0$. The continuity of g implies the existence of a parameter ε so that

$g(\underline{x}_0^*(\tau), \tau) > 0$ for all $\tau \in (\underline{\tau}, \bar{\tau}) =: T$ with $\underline{\tau} := \tau^+ - \varepsilon > \hat{\tau}$ and $\bar{\tau} := \tau^+ + \varepsilon$. Thus, the maintained existence of T is shown.

We have to maximize the difference $\bar{\tau} - \underline{\tau}$ such that $\tau^+ \in (\underline{\tau}, \bar{\tau})$ and separation holds. For that purpose we show:

a) Given the parameter $\underline{\tau} > \hat{\tau}$ the difference $\bar{\tau} - \underline{\tau}$ is maximized by setting $\bar{\tau}$ in such a way so

as to guarantee $g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0$,

b) Given two parameters $\underline{\tau}_2 > \underline{\tau}_1 > \hat{\tau}$ and corresponding parameters $\bar{\tau}_1, \bar{\tau}_2$ determined

according to a), we get $\bar{\tau}_2 - \underline{\tau}_2 > \bar{\tau}_1 - \underline{\tau}_1$.

Proof of a):

For $\underline{\tau} > \hat{\tau}$ we immediately get from (A14) $g(\underline{x}_0^*(\underline{\tau}), \underline{\tau}) > 0$. Using the definition of g it

results $\lim_{x_0 \rightarrow -\infty} g(x_0, \tau) = -\infty$. Since $\lim_{\bar{\tau} \rightarrow \infty} \underline{x}_0^*(\bar{\tau}) = -\infty$ in the case $\xi_0^* > 0$, this fact implies the existence of a parameter $\bar{\tau} > \underline{\tau}$ with

$$g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0. \quad (\text{A15})$$

Let $\bar{\tau}$ fulfill property (A15). With $\partial \underline{x}_0^*(\tau) / \partial \tau \leq 0$ the assumption $\underline{\tau} \geq \hat{\tau}$ leads to

$$g(\underline{x}_0^*(\tau), \tau) > 0 \text{ for all } \tau \in (\underline{\tau}, \bar{\tau}) \quad (\text{A16})$$

so that separation holds in the interval $(\underline{\tau}, \bar{\tau})$ and $\bar{\tau}$ is maximal with this property for given $\underline{\tau}$ since $g(\underline{x}_0^*(\tau), \tau) < 0$ for all $\tau > \bar{\tau}$.

Proof of b):

From (A15) we are able to determine the following relationship between $\underline{\tau}$ and $\bar{\tau}$:

$$\underline{\tau} = \begin{cases} -b \cdot (1 + r_0 + (1 - \underline{x}_0^*(\bar{\tau})) \cdot \underline{u}(0)), & \text{if } b > 0, \\ -b \cdot (1 + r_0 + (1 - \underline{x}_0^*(\bar{\tau})) \cdot \bar{u}(0)), & \text{if } b < 0. \end{cases} \quad (\text{A17})$$

The assumption $\underline{\tau}_2 > \underline{\tau}_1$ implies $\bar{\tau}_2 > \bar{\tau}_1$ and the difference $\bar{\tau} - \underline{\tau} = g(\underline{x}_0^*(\bar{\tau}), \bar{\tau})$ is monotone increasing in $\bar{\tau}$. Consequently, $\underline{\tau}_2 > \underline{\tau}_1$ leads to $\bar{\tau}_2 - \underline{\tau}_2 > \bar{\tau}_1 - \underline{\tau}_1$.

From a) and b) it follows immediately that the difference $\bar{\tau} - \underline{\tau}$ is maximized by setting $\underline{\tau} = \tau^+$ and $\bar{\tau}$ in such a way so as to guarantee $g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0$. Since $\bar{\tau} \rightarrow \infty$ for $\tau^+ \rightarrow \infty$ we obviously get $\underline{x}_0^*(\bar{\tau}) \xrightarrow{\tau^+ \rightarrow \infty} -\infty$.

We know that $\underline{x}_0^*(\tau)$ is decreasing in τ , since $\xi_0^* > 0$. If we assume $\underline{x}_0 \geq \underline{x}_0^*(\hat{\tau})$ separation has to hold for τ with $\underline{x}_0^*(\tau) \geq \underline{x}_0 \geq \underline{x}_0^*(\hat{\tau})$, i.e. $\tau \leq \hat{\tau}$. From (A13) we know $g(\underline{x}_0^*(\tau), \tau) \leq 0$ for all $\tau \leq \hat{\tau}$ which leads to a contradiction. Thus, we immediately get the necessary condition $\underline{x}_0 < \underline{x}_0^*(\hat{\tau})$ for separation to hold. Given the border \underline{x}_0 it is obvious that $\underline{x}_0^*(\bar{\tau}) = \underline{x}_0$ and $g(\underline{x}_0^*(\bar{\tau}), \underline{\tau}) = 0$.

Proof of Proposition 4:

b ≠ 1:

Using $\xi_0 = (1 - x_0) / (\tau + b \cdot (1 + r_0))$, we have:

$$\begin{aligned}
 u_{CE} &= \frac{E \left\{ \left[\tau + b \cdot (1 + r_0 + (1 - x_0) \cdot \tilde{u}(0, y)) \right]^{\frac{b-1}{b}} \right\}^{\frac{b}{b-1}} - \tau}{b} - 1 - r_0 \\
 &= \frac{E \left\{ \left[(1 - x_0) \cdot \left(\frac{1}{\xi_0} \right) + b \cdot \tilde{u}(0, y) \right]^{\frac{b-1}{b}} \right\}^{\frac{b}{b-1}} - (\tau + b \cdot (1 + r_0))}{b} \\
 &= (1 - x_0) \cdot \frac{E \left\{ \left[\left(\frac{1}{\xi_0} \right) + b \cdot \tilde{u}(0, y) \right]^{\frac{b-1}{b}} \right\}^{\frac{b}{b-1}} - \frac{1}{\xi_0}}{b}.
 \end{aligned} \tag{A18}$$

Since $E[(1 - x_0) \cdot \tilde{u}(0, y)] = (1 - x_0) \cdot E[\tilde{u}(0, y)]$ and ξ_0 as well as y are assumed to be constant, we finally get $d_b = 1 - u_{CE} / E[\tilde{u}]$ to be independent of x_0 .

b = 1:

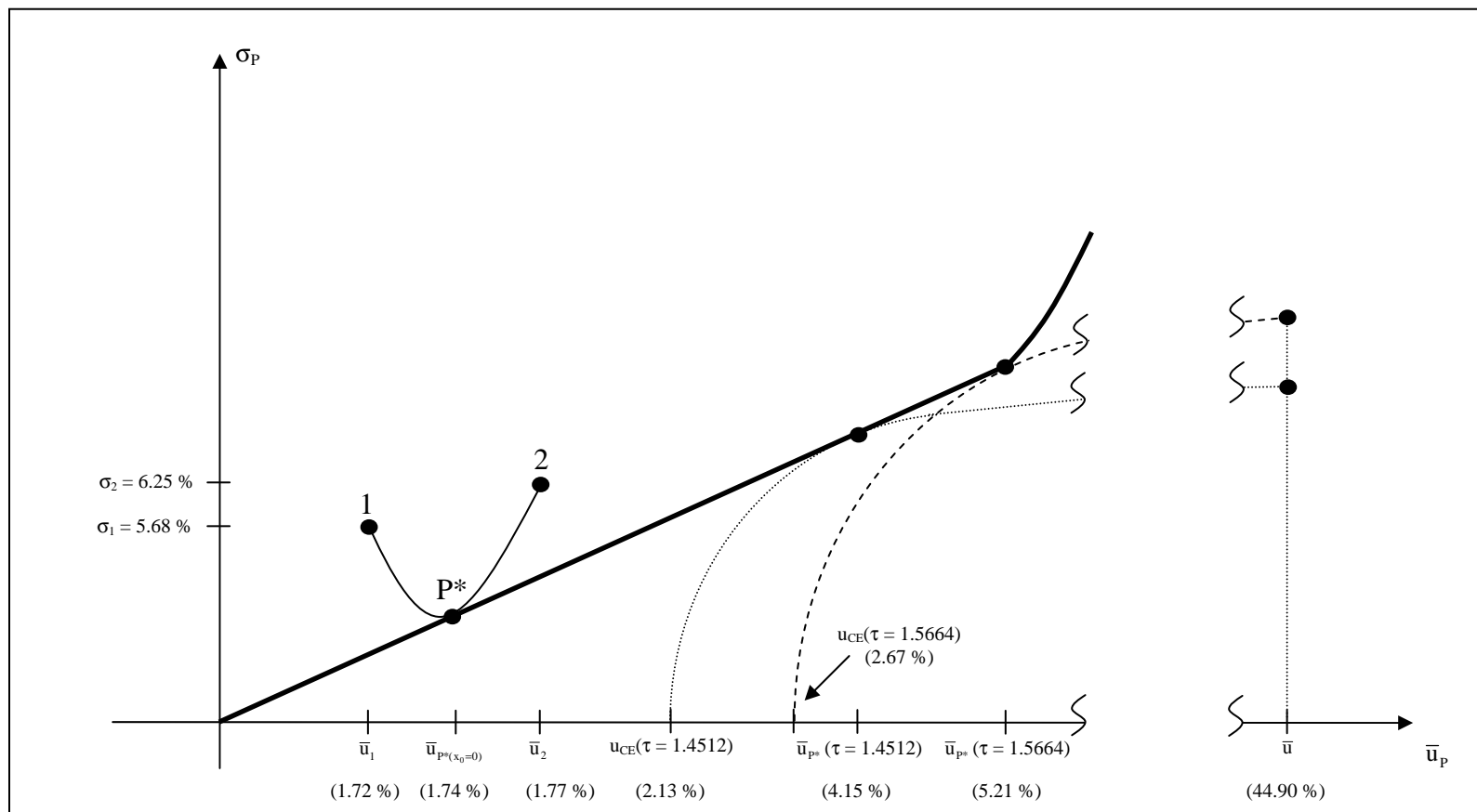
$$\begin{aligned}
 u_{CE} &= \exp(E[\ln(\tau + 1 + r_0 + (1 - x_0) \cdot \tilde{u}(0, y))]) - \tau - 1 - r_0 \\
 &= \exp(E[\ln((1 - x_0) \cdot \left(\frac{1}{\xi_0} \right) + \tilde{u}(0, y))]) - (\tau + 1 + r_0) \\
 &= (1 - x_0) \cdot \{ \exp(E[\ln(\left(\frac{1}{\xi_0} \right) + \tilde{u}(0, y)])]) - \left(\frac{1}{\xi_0} \right) \}.
 \end{aligned} \tag{A19}$$

Again, we get the postulated statement.

References

- Black F. Capital market equilibrium with restricted borrowing. *Journal of Business* 1972; 45; 444-455.
- Breuer W, Gürtler M. Investors' direct stock holdings and performance evaluation for mutual funds. *Kredit und Kapital* 2005; 38; 541-572.
- Breuer W, Gürtler M. Performance evaluation, portfolio selection, and HARA utility. *European Journal of Finance* 2006; 12; coming soon.
Download: <http://www.bfw.rwth-aachen.de/publika/wpaper/download/wp0201.pdf>.
- Cass D, Stiglitz JE. The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *Journal of Economic Theory* 1970; 2; 122-160.
- Fama EF, French KR. The equity premium. *Journal of Finance* 2002; 57; 637-659.
- Grinblatt M, Titman S. Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies* 1989; 2; 393-421.
- Hakansson NH. Risk disposition and the separation property. *Journal of Financial and Quantitative Analysis* 1969; 8; 401-416.
- Jensen MC. The performance of mutual funds in the period 1956-1964. *Journal of Finance* 1968; 23; 389-416.
- Jobson JD, Korkie B. On the Jensen measure and marginal improvements in portfolio performance – A note. *Journal of Finance* 1984; 39; 245-251.
- Kroll Y, Levy H, Markowitz HM. Mean-variance versus direct utility maximization. *Journal of Finance* 1984; 39; 46-61.
- Levy H, Markowitz HM. Approximating expected utility by a function of mean and variance. *American Economic Review* 1979; 69; 308-317.
- Lintner J. The valuation of risk and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 1965; 47; 13-37.
- Mossin J. Equilibrium in a capital asset pricing market. *Econometrica* 1966; 34; 768-783.
- Samuelson PA. The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments. *Review of Economic Studies* 1970; 37; 537-542.
- Sharpe WF. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*; 19; 425-442.
- Sharpe WF. Mutual fund performance. *Journal of Business* 1966; 39; 119-138.
- Tobin J. Liquidity preference as behaviour towards risk. *Review of Economic Studies* 1958; 25; 65-86.
- Treynor JL. How to rate management of investment funds, *Harvard Business Review* 1965; 43 (January/February); 63-75.
- Treynor JL, and Black F. How to use security analysis to improve portfolio selection. *Journal of Business* 1973; 46; 66-86.
- Vanden JM. Options trading and the CAPM. *Review of Financial Studies* 2004; 17; 207-238.

Figure 1: Two-fund separation for restricted borrowing when utility is quadratic (schematic, not true-to-scale)



The curve connecting points 1, P^* , and 2 describes the set of all achievable risky portfolios for $y \in [0,1]$.
 The bold line presents all mean-variance efficient overall portfolios.
 The dashed and the dotted curve represent indifference curves for different standardized risk tolerances τ .

Table 1: Unbiased estimators for expectation values \bar{u}_f , standard deviations σ_f , and covariances σ_{fp} of excess returns of funds (investing in German stocks) and reference portfolio P as well as minimum and maximum realized monthly excess fund returns

No.	name of fund	\bar{u}_f	σ_f	σ_{fp}	$\underline{u}(0,0)$	$\bar{u}(0,0)$
1	Aberdeen Global German Eq	0.46351 %	5.77708 %	0.33096 %	5.77708 %	0.33096 %
2	ABN AMRO Germany Equity	2.42189 %	7.09676 %	0.42209 %	7.09676 %	0.42209 %
3	AC Deutschland	1.86378 %	7.09276 %	0.41137 %	7.09276 %	0.41137 %
4	ADIFONDS	2.16243 %	7.22614 %	0.44304 %	7.22614 %	0.44304 %
5	Baer Multistock German Stk A	1.77270 %	5.48620 %	0.32287 %	5.48620 %	0.32287 %
6	Baring German Growth	2.85000 %	7.05836 %	0.33608 %	7.05836 %	0.33608 %
7	BBV Invest Union	1.90946 %	6.30927 %	0.38537 %	6.30927 %	0.38537 %
8	Berlinwerte Weberbank OP	1.57595 %	5.68085 %	0.33807 %	5.68085 %	0.33807 %
9	CB Lux Portfolio Euro Aktien	1.79676 %	6.77890 %	0.42088 %	6.77890 %	0.42088 %
10	Concentra	1.85919 %	6.71783 %	0.41575 %	6.71783 %	0.41575 %
11	CS EF (Lux) Germany	1.58297 %	6.66003 %	0.40816 %	6.66003 %	0.40816 %
12	DekaFonds	1.91459 %	6.81638 %	0.42138 %	6.81638 %	0.42138 %
13	DELBRÜCK Aktien UNION-Fonds	1.42919 %	6.25222 %	0.38175 %	6.25222 %	0.38175 %
14	Dexia Eq L Allemagne C	1.67865 %	6.23957 %	0.38700 %	6.23957 %	0.38700 %
15	DIT Fonds für Vermögensbildung	1.32405 %	5.79650 %	0.34777 %	5.79650 %	0.34777 %
16	DIT Wachstumsfonds	1.88919 %	6.28905 %	0.37674 %	6.28905 %	0.37674 %
17	DVG Fonds SELECT INVEST	2.07243 %	6.61112 %	0.40792 %	6.61112 %	0.40792 %
18	DWS Deutschland	1.60784 %	6.08441 %	0.36909 %	6.08441 %	0.36909 %
19	EMIF Germany Index plus B	1.57108 %	6.45667 %	0.40139 %	6.45667 %	0.40139 %
20	Fidelity Fds Germany	1.72892 %	6.24931 %	0.37989 %	6.24931 %	0.37989 %
21	Flex Fonds	1.39730 %	5.98888 %	0.36524 %	5.98888 %	0.36524 %
22	Frankfurter Sparinvest Deka	1.81324 %	6.41583 %	0.39600 %	6.41583 %	0.39600 %
23	FT Deutschland Dynamik Fonds	1.79459 %	6.59269 %	0.40786 %	6.59269 %	0.40786 %
24	Gerling Deutschland Fonds	1.41054 %	5.19347 %	0.31236 %	5.19347 %	0.31236 %
25	HANSAeffekt	1.73973 %	6.49867 %	0.40096 %	6.49867 %	0.40096 %
26	Hauck Main I Universal Fonds	1.45865 %	6.58482 %	0.40521 %	6.58482 %	0.40521 %
27	Incofonds	2.13865 %	6.04074 %	0.34912 %	6.04074 %	0.34912 %
28	Interselx Equity Germany B	1.72514 %	6.60614 %	0.40989 %	6.60614 %	0.40989 %
29	INVESCO GT German Growth C	1.71649 %	5.67770 %	0.24657 %	5.67770 %	0.24657 %
30	Investa	2.11541 %	6.92485 %	0.42699 %	6.92485 %	0.42699 %
31	Köln Aktienfonds DEKA	1.83865 %	6.54772 %	0.40355 %	6.54772 %	0.40355 %
32	Lux Linea	1.71378 %	7.60317 %	0.46976 %	7.60317 %	0.46976 %
33	Metallbank Aktienfonds DWS	2.07324 %	5.14655 %	0.26836 %	5.14655 %	0.26836 %
34	MK Alfakapital	1.98243 %	7.41669 %	0.45851 %	7.41669 %	0.45851 %
35	MMWI PROGRESS Fonds	1.76081 %	6.71760 %	0.41379 %	6.71760 %	0.41379 %
36	Oppenheim Select	1.69757 %	6.47148 %	0.39475 %	6.47148 %	0.39475 %
37	Parvest Germany C	1.60108 %	6.31697 %	0.39222 %	6.31697 %	0.39222 %
38	Plusfonds	2.40324 %	6.83304 %	0.40050 %	6.83304 %	0.40050 %
39	Portfolio Partner Universal G	1.09946 %	6.08717 %	0.32420 %	6.08717 %	0.32420 %
40	Ring Aktienfonds DWS	1.86784 %	6.15453 %	0.37430 %	6.15453 %	0.37430 %
41	SMH Special UBS Fonds 1	1.90811 %	6.60503 %	0.40739 %	6.60503 %	0.40739 %
42	Thesaurus	1.72811 %	6.36330 %	0.39459 %	6.36330 %	0.39459 %
43	Trinkaus Capital Fonds INKA	1.71541 %	6.49609 %	0.40013 %	6.49609 %	0.40013 %
44	UniFonds	1.74784 %	6.42735 %	0.39665 %	6.42735 %	0.39665 %
45	Universal Effect Fonds	1.74568 %	6.27421 %	0.38306 %	6.27421 %	0.38306 %
P	DAX 100	1.77189 %	6.24936 %	0.39055 %	6.24936 %	0.39055 %

Table 2: Intervals $T_{f,b}$ of admissible standardized of risk tolerances and relative risk discounts (in the cases $b = -1$, $b = -1/2$, $b = -1/3$)

$b = -1$	Ran-king	$y^{(3y)*}$	$\underline{x}_0 = 0 \%$		$\underline{x}_0 = -100 \%$		$\underline{x}_0 = -200 \%$		relative risk discounts	
			$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$d^*(\bar{\tau}^{(min)})$	$d^*(\underline{\tau}^{(max)})$
Berlinwerte Weberbank OP	2	29.65 %	1.1520	1.2312	1.3016	1.4598	1.4512	1.6885	60.10 %	72.12 %
DekaFonds	7	6.23 %	1.1601	1.2418	1.3176	1.4808	1.4752	1.7201	62.98 %	75.55 %
Fidelity Fds Germany	6	5.61 %	1.1520	1.2402	1.3016	1.4778	1.4512	1.7155	62.56 %	75.05 %
Frankfurter Sparinvest Deka	3	36.43 %	1.1520	1.2419	1.3016	1.4810	1.4512	1.7204	63.00 %	75.58 %
INVESCO GT German Growth C	1	61.21 %	1.1520	1.1905	1.3016	1.3783	1.4512	1.5664	48.79 %	58.70 %
Köln Aktienfonds DEKA	4	14.61 %	1.1520	1.2418	1.3016	1.4808	1.4512	1.7201	62.98 %	75.55 %
Universal Effect Fonds	5	9.48 %	1.1520	1.2402	1.3016	1.4776	1.4512	1.7153	62.54 %	75.03 %
Intersections ($\underline{\tau}^{(max)}, \bar{\tau}^{(min)}$):			1.1601	1.1905	1.3176	1.3783	1.4752	1.5664		
$b = -0.5$	Ran-king	$y^{(3y)*}$	$\underline{x}_0 = 0 \%$		$\underline{x}_0 = -100 \%$		$\underline{x}_0 = -200 \%$		relative risk discounts	
			$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$d^*(\bar{\tau}^{(min)})$	$d^*(\underline{\tau}^{(max)})$
Berlinwerte Weberbank OP	3	29.53%	0.5760	0.7315	0.6508	0.9618	0.7256	1.1922	60.00%	144.82%
DekaFonds	4	23.37%	0.5800	0.7447	0.6588	0.9883	0.7376	1.2319	63.63%	151.92%
Fidelity Fds Germany	7	bor. sol.								
Frankfurter Sparinvest Deka	2	53.75%	0.5760	0.7417	0.6508	0.9823	0.7256	1.2229	62.85%	149.77%
INVESCO GT German Growth C	1	64.33%	0.5760	0.6906	0.6508	0.8800	0.7256	1.0695	48.67%	120.26%
Köln Aktienfonds DEKA	5	17.75%	0.5760	0.7421	0.6508	0.9830	0.7256	1.2240	62.89%	150.76%
Universal Effect Fonds	6	1.63%	0.5760	0.7407	0.6508	0.9802	0.7256	1.2197	62.50%	150.03%
Intersections ($\underline{\tau}^{(max)}, \bar{\tau}^{(min)}$):			0.5800	0.6906	0.6588	0.8800	0.7376	1.0695		
$b = -1/3$	Ran-king	$y^{(3y)*}$	$\underline{x}_0 = 0 \%$		$\underline{x}_0 = -100 \%$		$\underline{x}_0 = -200 \%$		relative risk discounts	
			$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$d^*(\bar{\tau}^{(min)})$	$d^*(\underline{\tau}^{(max)})$
Berlinwerte Weberbank OP	3	29.81%	0.3840	0.5667	0.4339	0.7993	0.4837	1.0319	59.82%	216.62%
DekaFonds	4	30.51%	0.3866	0.5811	0.4392	0.8281	0.4917	1.0751	63.81%	225.66%
Fidelity Fds Germany	6	bor. sol.								
Frankfurter Sparinvest Deka	2	61.08%	0.3840	0.5768	0.4339	0.8195	0.4837	1.0621	62.68%	220.45%
INVESCO GT German Growth C	1	65.19%	0.3840	0.5255	0.4339	0.7169	0.4837	0.9083	48.47%	182.62%
Köln Aktienfonds DEKA	5	19.94%	0.3840	0.5776	0.4339	0.8211	0.4837	1.0646	62.80%	224.35%
Universal Effect Fonds	6	bor. sol.								
Intersections ($\underline{\tau}^{(max)}, \bar{\tau}^{(min)}$):			0.3866	0.5255	0.4392	0.7169	0.4917	0.9083		

Shaded funds imply border solutions (“bor. sol.”) for $b = -0.5$ or $b = -1/3$ and thus are consequently not considered in these cases.