A Concave Security Market Line*

Enrico G. De Giorgi[†] **Thierry Post**[‡] **Atakan Yalçın**[§]

April 26, 2019

Abstract

We provide theoretical and empirical arguments in favor of a diminishing marginal premium for market risk. In capital market equilibrium with binding portfolio restrictions, investors with different risk aversion levels generally hold different sets of risky securities. Whereas the traditional linear relation breaks down, equilibrium can be described or approximated by a concave relation between expected return and market beta, and a concave relationship between market alpha and market beta. An empirical analysis of U.S. stock market data confirms the existence of a significant concave cross-sectional relation between average return and estimated market beta. We estimate that the market risk premium is at least four to six percent per annum, substantially above traditional estimates. A practical implication for active portfolio managers is that the alpha of "betting against beta" strategies seems dominated by the medium-minus-high-beta spread rather than the low-minus-medium-beta spread. The success of such strategies thus largely depends on underweighting or short selling high-beta stocks.

Keywords: capital market equilibrium, asset pricing, investment restrictions, portfolio theory, market beta, stock selection.
JEL Classification: G12, C21.

---

[*] We thank Stephen Brown, Haim Levy, Lasse Pedersen, Pim van Vliet and seminar participants at Ross Business School for their many useful comments and suggestions.

[†] Department of Economics, School of Economics and Political Science, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland, Tel. +41 +71 2242430, Fax. +41 +71 224 28 94, email: enrico.degiorgi@unisg.ch.

[‡] Nazarbayev University, Graduate School of Business, Astana 010000, Kazakhstan; e-mail: thierry-post@hotmail.com.

[§] Ozyegin & Cekmekoy University – Graduate School of Business, Çekmeköy, 34794 Istanbul, Turkey, Tel. +90 216 564 9587, email: atakan.yalcin@ozyegin.edu.tr.

A Concave Security Market Line

Abstract

We provide theoretical and empirical arguments in favor of a diminishing marginal premium for market risk. In capital market equilibrium with binding portfolio restrictions, investors with different risk aversion levels generally hold different sets of risky securities. Whereas the traditional linear relation breaks down, equilibrium can be approximated by a concave relation between expected return and market beta, and a concave relationship between market alpha and market beta. An empirical analysis of U.S. stock market data confirms the existence of a significant concave cross-sectional relation between average return and estimated market beta. We estimate that the market risk premium is at least four to six percent per annum, substantially above traditional estimates. A practical implication for active portfolio managers is that the alpha of "betting against beta" strategies seems dominated by the medium-minus-high-beta spread rather than the low-minus-medium-beta spread. The success of such strategies thus largely depends on the underweighting or short selling of high-beta stocks.

Keywords: capital market equilibrium, asset pricing, investment restrictions, portfolio theory, market beta, stock selection.

JEL Classification: G12, C21.

Introduction

The Capital Asset Pricing Model predicts a linear relation between the expected return and the market beta of securities - the Security Market Line. This linear relation also arises as a special case in the Arbitrage Pricing Theory if one assumes that security returns obey the “market-model”. A wealth of empirical research suggests that linearity should be rejected. For example, Fama and French (1992) show that small-cap stocks carry a return premium that is unrelated to market beta and that the empirical return-risk relation appears flat after controlling for market capitalization.

It is well documented that the CAPM generally breaks down in case of binding investment restrictions. Black (1972) analyzes market equilibrium with restrictions on borrowing. In this case, the SML remains linear, although the intercept increases and the slope decreases compared with the CAPM predictions. Linearity generally breaks down if investors also face restrictions on the risky securities, such as short selling constraints and position limits. Ross (1977) and Sharpe (1991) stressed this important point before, but they stopped short of deriving an alternative shape for the SML. The present study extends Black (1972), Ross (1977) and Sharpe (1991) by providing theoretical and empirical arguments in favor of a concave shape for the SML, or a diminishing marginal premium for market risk in the general case of binding portfolio constraints. A general formulation is obtained by treating the most extreme feasible portfolios of individual securities (or the vertices of the portfolio possibilities set) as base assets.

Linearity is a necessary and sufficient condition for mean-variance efficiency of the market portfolio. However, deviations from linearity are not always good measures of deviations from efficiency, and vice versa. It is well-known that a relatively small (large) reduction in expected return or increase in standard deviation of the market index can sometimes lead to relatively large (small) deviations from the classical SML. This study focuses on the cross-sectional mean-beta relation rather than evaluating market portfolio efficiency. Portfolio efficiency tests require an explicit specification of the relevant set of investment restrictions, whereas we eye results that apply more generally for every set of restrictions. In addition, we are interested in characterizing equilibrium with an inefficient market portfolio, or specifying the "alternative hypothesis" rather than the "null hypothesis" of efficiency tests.

Our theoretical analysis in Section I develops an extension of the CAPM with investment restrictions in the spirit of Sharpe (1991). In our model, different investors generally include different sets of risky securities in their portfolios (and thus deviate from the market portfolio). Despite the differences in their composition, the individual portfolios generally will be positively correlated and share a common exposure to systematic risk. The general equilibrium conditions appear difficult to test empirically without detailed information about the composition of the individual portfolios and the distribution of risk tolerance and wealth, a common problem for heterogeneous-investor models. However, due to the investors' common exposure to systematic risk, equilibrium can be approximated by an increasing and concave relation between expected return and the traditional market beta - a concave security market line (CSML). This relation is a generalization of the general linear SML derived by Black (1972). The concave SML becomes linear if no investor faces binding restrictions for the risky securities, and the CAPM arises if riskless borrowing is also allowed.

Our empirical analysis in Section II applies the two-pass regression methodology of Fama and MacBeth (1973) to U.S. stock market data. Following the original study, we use the squared market beta as a natural measure of SML concavity. Whereas we find no significant linear cross-sectional relation between average returns and past beta estimates, we do find a significant and robust concave relation, consistent with our hypothesis. The inclusion of beta-squared in the regression yields an estimated market-risk premium of at least four to six percent per annum for the average stock, substantially higher than conventional estimates. In addition, beta-squared has a significantly negative coefficient, implying that the risk premium increases at a diminishing rate. Encouragingly, the concave pattern is robust to the inclusion of other stock characteristics and the selection of the cross-section and sample period. Concavity arises both in the analysis of individual stocks and for aggregated stock portfolios with stable size and beta properties, reducing possible concerns about estimation error and time-variation of stock-level betas.

Section III evaluates several active investment strategies that “bet against beta” by overweighting low-beta stocks and/or underweighting high-beta stocks. Confirming the results of Frazzini and Pedersen (2014), we find a large and robust alpha spread between low-beta and high-beta stocks. Interestingly, this low-minus-high-beta spread is dominated by the medium-minus-high-beta spread and the low-minus-medium-beta spread is materially smaller. The success of “betting against beta” thus seems to largely depend on underweighting or short selling high-beta stocks. These findings seem consistent with, and supportive of, our theoretical model, because a concave SML implies a concave relation between market alpha and market beta, so that the spread between medium-beta stocks and high-beta stocks is indeed predicted to exceed the spread between low-beta stocks and medium-beta stocks.

I Theory

A Preliminaries

We will first introduce and motivate our assumptions. We build on the earlier work of Sharpe (1991). It is not our objective to develop the most general asset pricing model, but rather to explore the effect of binding investment restrictions while maintaining a set of simplifying assumptions that support a linear SML in the absence of binding restrictions.

Assumption 1 (Securities)**.** *The investment universe consists of* $N+1$ *base assets, associated with returns* $\mathbf{r} \in \mathbb{R}^{N+1}$. *Throughout the text, we will use the index set* $\mathcal{I} = \{1, \ldots, N+1\}$ *to denote the different securities. The returns have mean* $\mathbb{E}[\mathbf{r}] = \boldsymbol{\mu}$ *and variance-covariance matrix* $\mathbb{E}\left[(\mathbf{r}-\boldsymbol{\mu})(\mathbf{r}-\boldsymbol{\mu})^T\right] = \boldsymbol{\Omega}$. *Security* $N+1$ *is risk-free and yields a sure return of* r_F.

Assumption 2 (Investors)**.** *There are* K *investors. Investor* k*'s wealth expressed as a proportion of the total wealth of all investors is* $w_k > 0$. *Investors may diversify between the securities, and we will use* $\boldsymbol{\lambda}_k \in \mathbb{R}^{N+1}$ *for the vector of optimal portfolio weights of investor* k. *We will use the index set* $\mathcal{K} = \{1, \ldots, K\}$ *to denote all investors. Investors possess mean-variance preferences, that is, the expected utility of wealth associated with a portfolio* $\boldsymbol{\lambda} \in \mathbb{R}^{N+1}$ *is*

$$U_{\zeta_k}(\boldsymbol{\lambda}) = \boldsymbol{\mu}^T \boldsymbol{\lambda} - \frac{1}{2\,\zeta_k} \boldsymbol{\lambda}^T \boldsymbol{\Omega}\, \boldsymbol{\lambda},$$

where $\zeta_k > 0$ *is investor* k*'s risk tolerance. The limiting cases with* $\zeta_k \to \infty$ *and* $\zeta_k \to 0$ *are risk neutrality and extreme risk aversion, respectively. We will use* $\mu_{\boldsymbol{\lambda}}$ *and* $\sigma^2_{\boldsymbol{\lambda}}$ *to denote the expected return and variance, respectively, of portfolio* $\boldsymbol{\lambda} \in \mathbb{R}^{N+1}$.

Mean-variance preferences can equivalently be represented by a quadratic utility function. Levy and Markowitz (1979) convincingly show that this specification generally gives an accurate second-order Taylor series approximation for many well-behaved utility functions on the typical return interval for stock portfolios. In some cases, the more general criteria of stochastic dominance are more appropriate. In this more general case, our arguments seem even more relevant, because differences in the general shape of the investors' utility functions will represent an additional source of non-linearity and market portfolio inefficiency.

Assumption 3 (Portfolios). *The portfolio possibilities are represented by the simplex*

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^{N+1} : \boldsymbol{\lambda} \geq \mathbf{0}_{N+1}, \, \mathbf{1}_{N+1}^T \boldsymbol{\lambda} = 1\},$$

where $\mathbf{0}_{N+1}$ *and* $\mathbf{1}_{N+1}$ *are the zero and unity vectors, respectively, with dimension* $N+1$*. We use* $\mathcal{A}_{\boldsymbol{\lambda}} = \{i : \lambda_i > 0, \, i \in \mathcal{I}\}$ *for the "active set" of portfolio* $\boldsymbol{\lambda}$*, or all securities that are included in the portfolio with a strictly positive weight.*

The simplex Λ is the convex hull of the base assets and covers the standard case without short selling and borrowing if the base assets are individual traded securities. However, the analysis also covers the general case where the portfolio possibilities are a general polytope. The Minkowski-Weyl Theorem says that any polytope can be represented as the convex hull of its vertices. Therefore, the base asset should be seen in general as the most extreme feasible portfolios, which may include long-short combinations of the individual securities.

For the sake of simplicity, we assume that all investors face the same set of investment restrictions. In the more general case with heterogenous restrictions, our arguments appear even more relevant, as differences in the efficient sets between investors would represent an additional source of non-linearity and market portfolio inefficiency. Although we use a single set of restrictions, investors with different levels of risk tolerance will generally have different sets of binding constraints.

Assumption 4 (Market portfolio)**.** *The markets for the risky securities clear, so that all risky securities are held by the K investors in the economy. We will use $\boldsymbol{\tau} \in \mathbb{R}^{N+1}$ for the vector of portfolio weights of the market portfolio of risky assets, where $\tau_i = \sum_{k \in \mathcal{K}} w_k \, \lambda_{i,k}/(1 - \lambda_{N+1,k})$ for $i = 1, \ldots, N$ and $\tau_{N+1} = 0$. The market portfolio is feasible, or $\boldsymbol{\tau} \in \Lambda$, and all risky securities have a strictly positive market capitalization, or $\tau_i > 0$ for all $i = 1, \ldots, N$.*

B General Security Market Line (GSML)

The above assumptions allow us to define portfolio optimality and mean-variance efficiency and to derive and analyze equilibrium conditions. A given portfolio $\boldsymbol{\lambda} \in \Lambda$ is optimal for an investor with risk tolerance $\zeta > 0$ if and only if it maximizes expected utility:

$$\boldsymbol{\lambda} = \arg\max_{\boldsymbol{\kappa} \in \Lambda} U_\zeta(\boldsymbol{\kappa}). \tag{1}$$

Following Sharpe (1991, Equation 12), we will make use of the Karush-Kuhn-Tucker (KKT) conditions for this optimization problem:

$$\mu_i = \frac{1}{\zeta} \operatorname{Cov}\left(r_i, \mathbf{r}^T \boldsymbol{\lambda}\right) + \theta_{\boldsymbol{\lambda}} + \alpha_{i,\boldsymbol{\lambda}}, \quad i \in \mathcal{I} \tag{2}$$

$$\alpha_{i,\boldsymbol{\lambda}} \leq 0, \quad \forall i \in \mathcal{I} \tag{3}$$

$$\alpha_{i,\boldsymbol{\lambda}} \, \lambda_i = 0, \quad \forall i \in \mathcal{I} \tag{4}$$

$$\mathbf{1}_N^T \boldsymbol{\lambda} = 1, \quad \lambda_i \geq 0, \quad \forall i \in \mathcal{I}. \tag{5}$$

In these conditions, $\theta_{\boldsymbol{\lambda}} \in \mathbb{R}$ and $\boldsymbol{\alpha}_{\boldsymbol{\lambda}} = (\alpha_{1,\boldsymbol{\lambda}}, \ldots, \alpha_{N+1,\boldsymbol{\lambda}})^T \in \mathbb{R}^{N+1}$ are Lagrange multipliers that measure the shadow prices of the budget constraint and the negative shadow prices of the no-short-selling/no-borrowing constraints, respectively.

The complementary slackness condition (4) implies that the active securities ($i \in \mathcal{A}_{\boldsymbol{\lambda}}$) must have a zero shadow price for the associated short-sales restriction ($\alpha_{i,\boldsymbol{\lambda}} = 0$), whereas strictly positive shadow prices are allowed for inactive securities ($\alpha_{i,\boldsymbol{\lambda}} \leq 0$ for $i \notin \mathcal{A}_{\boldsymbol{\lambda}}$). Each of the active securities must therefore have the same marginal utility. The common value of marginal utility for the active securities is the shadow price $\theta_{\boldsymbol{\lambda}}$, or the investor's marginal utility of wealth. The inactive securities must have a marginal utility less than or equal to of the active securities, and $\boldsymbol{\alpha}_{\boldsymbol{\lambda}}$ measures their shortfall.

The availability of the riskless security implies $r_F = \theta_{\boldsymbol{\lambda}} + \alpha_{N+1,\boldsymbol{\lambda}}$, and therefore, the shadow price of the budget constraint cannot fall below the risk-free rate: $\theta_{\boldsymbol{\lambda}} \geq r_F$. If the borrowing restriction is not binding for portfolio $\boldsymbol{\lambda}$, the complementary slackness condition (4) implies that $\alpha_{N+1,\boldsymbol{\lambda}} = 0$ and thus $\theta_{\boldsymbol{\lambda}} = r_F$.

For a risky portfolio $\boldsymbol{\lambda}$, that is, $\sigma^2_{\boldsymbol{\lambda}} > 0$, the optimality condition (2) can be expressed in terms of the exposure coefficient, or "beta", of security i with respect to portfolio $\boldsymbol{\lambda}$:

$$\mu_i = \left(\frac{1}{\zeta}\,\sigma^2_{\boldsymbol{\lambda}}\right)\beta_{i,\boldsymbol{\lambda}} + \theta_{\boldsymbol{\lambda}} + \alpha_{i,\boldsymbol{\lambda}}, \tag{6}$$

where

$$\beta_{i,\boldsymbol{\lambda}} = \frac{\operatorname{Cov}\left(r_i, \mathbf{r}^T\boldsymbol{\lambda}\right)}{\sigma^2_{\boldsymbol{\lambda}}}. \tag{7}$$

The optimality conditions (2) and (4) imply the following relation between an optimal portfolio's expected return $\mu_{\boldsymbol{\lambda}}$, its standard deviation $\sigma_{\boldsymbol{\lambda}}$, the shadow price of the budget constraint $\theta_{\boldsymbol{\lambda}}$ and the risk tolerance parameter ζ:

$$\frac{1}{\zeta}\,\sigma^2_{\boldsymbol{\lambda}} = \mu_{\boldsymbol{\lambda}} - \theta_{\boldsymbol{\lambda}}. \tag{8}$$

If the borrowing restriction is not binding for portfolio $\boldsymbol{\lambda}$, the right-hand side of Equation (8) is the risk premium $\mu_{\boldsymbol{\lambda}} - r_F$ of portfolio $\boldsymbol{\lambda}$.

For cases in which investors' risk tolerance and/or optimal portfolios are not specified, it is useful to consider the mean-variance efficient set, or all portfolios that are optimal for at least some risk tolerance level:

$$\mathcal{E} = \{\boldsymbol{\lambda} \in \Lambda \, : \, \boldsymbol{\lambda} = \arg\max_{\boldsymbol{\kappa} \in \Lambda} U_{\zeta}(\boldsymbol{\kappa}), \; \zeta > 0\}. \tag{9}$$

Under the above assumptions, the efficient set generally is not convex. Combining multiple efficient portfolios with different sets of binding restrictions generally produces an inefficient combined portfolio. The efficient set contains convex neighborhoods of efficient portfolios with the same sets of binding restrictions, but combining elements from these efficient subsets generally yields inefficient portfolios.

The aggregate market portfolio is a weighted average of the portfolios of all individual investors. Without convexity of the efficient set, optimizing behavior by individual investors does not guarantee that the market portfolio is optimal for a representative investor. To the contrary, given that different investors generally hold different portfolios with different sets of binding restrictions, the market portfolio generally is predicted to be inefficient.

Following Sharpe (1991, Equations 13-14), we can obtain the the following general relation between securities' expected returns and their market betas by taking a wealth-weighted average of the investors' KKT conditions:

Theorem 1 (Generalized Security Market Line)**.** *For all* $i \in \mathcal{I}$

$$\mu_i = \left(\frac{1}{\bar{\zeta}} \, \sigma_{\boldsymbol{\tau}}^2 \right) \beta_{i,\boldsymbol{\tau}} + \bar{\theta} + \bar{\alpha}_i \tag{10}$$

where $\bar{\zeta} = \sum_{k \in \mathcal{K}} w_k \, \zeta_k / (1 - \lambda_{N+1,k})$ *is the societal risk tolerance,* $\bar{\theta} = (1/\bar{\zeta}) \sum_{k \in \mathcal{K}} w_k \, \zeta_k \, \theta_{\boldsymbol{\lambda}_k} / (1 - \lambda_{N+1,k})$ *are weighted averages of the shadow prices of the borrowing restrictions and* $\bar{\alpha}_i =$

$(1/\bar{\zeta})\sum_{k\in\mathcal{K}} w_k\,\zeta_k\,\alpha_{i,\boldsymbol{\lambda}_k}/(1-\lambda_{N+1,k})$ *are weighted averages of the negative shadow prices the short-sales restrictions.*

The market portfolio is efficient if the constraints for the risky securities are not binding for all investors. In this case, all relevant shadow prices are zero ($\boldsymbol{\alpha}_{i,\boldsymbol{\lambda}_k} = 0$) and we obtain the classical linear relation between securities' expected returns and their market betas (SML), that is,

$$\mu_i = \left(\frac{1}{\bar{\zeta}}\,\sigma^2_{\boldsymbol{\tau}}\right)\beta_{i,\boldsymbol{\tau}} + \bar{\theta}. \tag{11}$$

The intercept $\bar{\theta}$ and the slope $\bar{\zeta}$ depend on the restrictions that are imposed on the risk-free security. If the borrowing restriction is not binding for any investor, then the intercept equals the risk-free rate ($\overline{\theta} = r_F$) and the slope equals the market-risk premium ($(1/\bar{\zeta})\,\sigma^2_{\boldsymbol{\tau}} = \mu_{\boldsymbol{\tau}} - r_F$; see Equation (8)). If borrowing is binding for some investors, then the intercept should be greater than or equal to the risk-free rate ($\overline{\theta} \geq r_F$) and the slope should be smaller than or equal to the market-risk premium ($(1/\bar{\zeta})\,\sigma^2_{\boldsymbol{\tau}} = \mu_{\boldsymbol{\tau}} - \overline{\theta} \leq \mu_{\boldsymbol{\tau}} - r_F$; see Equation (8)), as in Black's (1972) model.

The general non-linear GSML appears difficult to test without detailed information about individual portfolios ($\boldsymbol{\lambda}_k$) and the distribution of wealth (w_k) and risk tolerance (ζ_k), a common problem faced in heterogeneous investor models. Nevertheless, as we will show next, the GSML can be approximated by an increasing and concave function as a result of the investors' joint exposure to systematic risk.

C Concave Security Market Line (CSML)

Investors generally hold only a subset of the available securities and deviate from market weights. Nevertheless, the individual portfolios can be strongly correlated with the market portfolio, due to a common exposure to systematic risk. As a case in point, we created two monthly-rebalanced, non-overlapping stock portfolios of approximately equal market value from the stocks included our empirical analysis (see Section II): a value-weighted portfolio of stocks with below-median market betas and a value-weighted portfolio of stocks with above-median betas. Although the two portfolios each represent only half of the total market, the correlation of monthly returns with the stock market index is 96 percent for the low-beta portfolio and 98 percent for the high-beta portfolio.

If the portfolios are indeed strongly correlated, the correlation of an individual security with a given optimal portfolio ($\rho_{i,\boldsymbol{\lambda}_k}$) will tend to be closely related to its correlation with the market portfolio ($\rho_{i,\boldsymbol{\tau}}$). To capture this pattern, we will use the following correlation ratio:

$$\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \frac{\rho_{i,\boldsymbol{\lambda}_k}}{\rho_{i,\boldsymbol{\tau}}}. \tag{12}$$

It seems natural to assume that this correlation ratio shows relatively small differences between the active assets in a given optimal portfolio and that the ratio tends to be larger for active assets than for inactive assets. For example, a conservative investor is more likely to invest in large-cap low-beta stocks and an adventurous investor is more likely to invest in small-cap high-beta stocks. We formalize this assumption in the following way:

Assumption 5 (Correlation ratios)**.** *For all $k \in \mathcal{K}$, there exists $\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \geq 0$ such that*

$$\xi_{i,\boldsymbol{\lambda_k},\boldsymbol{\tau}} = \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \quad \textit{for all} \quad i \in \mathcal{A}_{\boldsymbol{\lambda}_k}, \quad \textit{and} \tag{13}$$

$$\xi_{i,\boldsymbol{\lambda_k},\boldsymbol{\tau}} \leq \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \quad \textit{for all} \quad i \notin \mathcal{A}_{\boldsymbol{\lambda}_k}. \tag{14}$$

This correlation structure arises for several relevant special cases. It is straightforward to see that the correlation applies in the case of the CAPM with a single optimal risky portfolio ($\boldsymbol{\lambda}_k = c_k\,\boldsymbol{\tau}$). A more general sufficient condition based on a general risk factor model will be discussed in Subsection D.

An analysis of the correlation coefficients of individual stocks relative to the low-beta portfolio, high-beta portfolio and market portfolio in our running example supports Assumption 6. Notably, the average low-beta stock has a higher correlation ratio relative to the low-beta portfolio than the average high-beta stock (0.92 vs. 0.87) but a lower correlation ratio relative to the high-beta portfolio (1.04 vs. 1.06).

Under Assumption 5, we can derive an interesting special case of the GSML:

Theorem 2 (Concave Security Market Line)**.** *Under Assumption 5 the following increasing and concave relation between expected return μ_i and market beta $\beta_{i,\boldsymbol{\tau}}$ holds for all $i = 1, \ldots, N$:*

$$\mu_i = \hat{\mu}_i = \min_{k \in \mathcal{K}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i,\boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right]. \tag{15}$$

The CSML is a piecewise-linear function of market beta. Every linear line segment reflects an individual investor's portfolio optimization problem; the intercept ($\theta_{\boldsymbol{\lambda}_k}$) is the investor's shadow price of the budget constraint and the slope ($(1/\zeta_k)\, \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}$) reflects her risk tolerance level and the covariance between her portfolio and the market portfolio.

Every individual line connects the investor's active securities and supports her inactive securities from above. Minimizing across these individual, increasing and linear functions yields an overall increasing and concave, piecewise linear shape.

For example, assume that $K = 3$, and

$$\left(\theta_{\boldsymbol{\lambda}_1}, \frac{1}{\zeta_1} \sigma_{\boldsymbol{\lambda}_1} \sigma_{\boldsymbol{\tau}} \xi_{\boldsymbol{\lambda}_1, \boldsymbol{\tau}} \right) = (0.055, 0.07);$$

$$\left(\theta_{\boldsymbol{\lambda}_2}, \frac{1}{\zeta_2} \sigma_{\boldsymbol{\lambda}_2} \sigma_{\boldsymbol{\tau}} \xi_{\boldsymbol{\lambda}_2, \boldsymbol{\tau}} \right) = (0.02, 0.11);$$

$$\left(\theta_{\boldsymbol{\lambda}_3}, \frac{1}{\zeta_3} \sigma_{\boldsymbol{\lambda}_3} \sigma_{\boldsymbol{\tau}} \xi_{\boldsymbol{\lambda}_3, \boldsymbol{\tau}} \right) = (0.04, 0.10).$$

In this case, the CSML relation amounts to the following three-piece function:

$$\mu_i = \begin{cases} 0.02 + 0.11\beta_{i,\boldsymbol{\tau}} & \beta_{i,\boldsymbol{\tau}} \leq \frac{2}{3} \\ 0.04 + 0.10\beta_{i,\boldsymbol{\tau}} & \frac{2}{3} < \beta_{i,\boldsymbol{\tau}} \leq \frac{3}{2} \\ 0.055 + 0.07\beta_{i,\boldsymbol{\tau}} & \frac{3}{2} < \beta_{i,\boldsymbol{\tau}} \end{cases} .$$

In general, the number of line segments increases with the number of different investors with different active sets and the function can approximate a smooth curve in case of a continuum of investors with different risk-tolerance levels.

The concavity of the SML follows from the mathematical result that the pointwise infimum of multiple concave functions preserves the concavity property. In economic terms, the relationship is concave, because the price of a unit of risk is relatively high (low) for low-beta (high-beta) securities that enter in the portfolio of conservative (adventurous) investors but not in the portfolio of adventurous (conservative) investors. Subsection E gives a full-blown numerical example.

The empirical analysis in Section II captures the concavity, or non-linearity, of the SML with the squared market beta, first introduced in empirical asset pricing by Fama and MacBeth (1973). This approach can be interpreted as providing a second-order Taylor series approximation to a smoothed version of the piece-wise linear CSML.

The non-linear shape of the CSML is reminiscent of asset pricing models that assign a role to higher-order and lower-partial co-moments with the market portfolio (for example, Bawa and Lindenberg 1977). However, our model exhibits a non-linear price of market beta rather than a price for non-linear market-risk exposure, and there are important differences between these two approaches.

First, the theoretical motivation is very different. The role of non-linear market-risk exposure follows from deviations from a normal return distribution and mean-variance risk preferences, whereas our non-linearity stems from differences between investors' risk tolerance and binding restrictions, an explanation that is perfectly consistent with mean-variance analysis. Second, the return distribution in typical applications seems sufficiently close to a normal distribution to apply the Levy and Markowitz (1979) argument for mean-variance analysis. Indeed, Dittmar (2002, Section IIID) and Post (2003, Section IV) show in a convincing way that higher-order and lower-partial co-moments with the market portfolio cannot rationalize market portfolio inefficiency. By contrast, the effect of binding restrictions occurs also under a normal distribution, as our model shows.

Our model allows the market portfolio to be mean-variance inefficient. In case of market portfolio inefficiency, market beta still plays a role, because the individual optimal portfolios have a strong joint exposure to market risk. However, we do not intend to measure the economic magnitude of deviations from market portfolio efficiency.

Ever since Roll (1977) and more recently Lustig, Van Nieuwerburgh, and Verdelhan (2013), the asset pricing literature has recognized that stock market indices are only a proxy for the latent market portfolio. In Roll and Ross (1994) and Kandel and Stambaugh (1995), divergence from the market portfolio (benchmark error) causes (spurious) violations of linearity. Unfortunately, stronger (weaker) non-linearities generally do not imply that the market portfolio is further from (closer to) the efficient frontier.

In our model, (true) violations of linearity arise from binding investment restrictions, as in Ross (1977) and Sharpe (1991). As in the case of benchmark error, larger (smaller) non-linearities generally do not mean that the market portfolio is further from (closer to) the frontier. In addition, a large distance from the frontier generally is not "economically worse" than a small distance, because the distance is not measured using the investor's (latent) optimal portfolio, but using the "nearest efficient portfolio".

One interesting implication of the CSML theorem is that the pricing errors of the CAPM (market alphas) will tend to display a concave pattern:

Corollary 1 (Concave Alpha-Beta Relationship)**.** *Let*

$$\alpha(\beta_{i,\boldsymbol{\tau}}) = \bar{\alpha}_i = \mu_i - \left[r_F + \beta_{i,\boldsymbol{\tau}}\left(\mu_{\boldsymbol{\tau}} - r_F\right)\right]$$

be market alpha as a function of market beta $\beta_{i,\boldsymbol{\tau}}$*. Under Assumption 5,* α *is a concave function of* $\beta_{i,\boldsymbol{\tau}}$ *and*

$$\alpha(\beta_{i,\boldsymbol{\tau}}) - \alpha(\beta_{i,\boldsymbol{\tau}} + \beta_{j,\boldsymbol{\tau}}) > \alpha(\beta_{i,\boldsymbol{\tau}} - \beta_{j,\boldsymbol{\tau}}) - \alpha(\beta_{i,\boldsymbol{\tau}}). \tag{16}$$

In our model, the market alphas are not pricing errors or deviations from equilibrium. Rather, they represent the negative shadow prices of binding investment restrictions to investors who deviate from the inefficient market portfolio. Clearly, an investor who would hold the market portfolio could improve her portfolio by over-weighting (under-weighting) the assets with positive (negative) market alphas in our model. However, in equilibrium, investors do not hold the market portfolio and the alphas relative to their personal optimal portfolios simply reflect the negative shadow price of binding restrictions. Clearly, if the set of restrictions were relaxed, for example, by means of financial engineering, investors may benefit by increasing (decreasing) their allocation to assets with positive (negative) alphas (relative to their personal portfolios). Section III will demonstrate that the market alpha of active investment portfolios indeed appears a concave function of the market beta.

The CSML (15) is intended as an approximation to the complex GSML (10) that is implied by the diverse optimality conditions of investors with different levels of risk tolerance and different active sets. Our model is not intended to describe the individual optimal portfolios. In fact, if the CSML gives a perfect fit, then the portfolios of different investors are predicted to show minimal overlap. Specifically, every pair of two different linear line segments in mean-beta space can share at most one mean-beta combination. Hence, the portfolios of two different investors (with different intercepts and slopes) will overlap only for more than one security if the overlapping securities have the same market beta. Clearly, this prediction seems equally unrealistic as the CAPM prediction that all investors hold all securities and use market capitalization weights. The general GSML does allow for general overlapping portfolios and the restrictions on portfolio overlap arise from considering the limiting case of neutral optimal non-market loadings.

Deviations from the correlation structure in Assumption 5 yield the following deviations of the CSML (15) from the general GSML (10):

Theorem 3 (CSML Errors in Expected Returns)**.** *Assume that* $\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \geq 0$ *for all* $k \in \mathcal{K}$ *and* $i = 1, \ldots, N$*. Then*

$$|\hat{\mu}_i - \mu_i| \leq |\mu_{\boldsymbol{\lambda}^\star} - r_f| \, \frac{\sigma_{\boldsymbol{\tau}}}{\sigma_{\boldsymbol{\lambda}^\star}} \, \beta_{i,\boldsymbol{\tau}} \max_{k\,:\,i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} - \xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \right| . \tag{17}$$

where $\boldsymbol{\lambda}^\star = \arg\max_{\boldsymbol{\kappa} \in \mathcal{E}} \left(\frac{\mu_{\boldsymbol{\kappa}} - \theta_{\boldsymbol{\kappa}}}{\sigma_{\boldsymbol{\kappa}}} \right)$ *is the mean-variance tangency portfolio.*

For a given security i, the error is bounded by its correlation ratios in portfolios that include the security, or, $\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}}$ for k such that $i \in \mathcal{A}_{\boldsymbol{\lambda}_k}$. In turn, these correlation ratios will be bounded if the optimal portfolios are well-diversified and have limited non-market risk exposure, as will be shown in the next subsection.

D Refinements Using a Risk Model

We will now derive additional results by introducing a general risk factor model:

Assumption 6 (Return generating process)**.** *The returns of the risky securities obey the following general risk factor model*

$$r_i = a_i + \sum_{l=1}^{L} b_{i,l} \, f_l + \epsilon_i \quad \forall i = 1, \ldots N, \tag{18}$$

where f_l*,* $l = 1, \ldots, L$ *are systematic risk factors with* $\mathbb{E}\left[(f_l - \mathbb{E}\left[f_l\right])^2\right] = \sigma_{f_l}^2$*,* $b_{i,l}$ *are the factor loadings of security* i*,* $a_i = \mu_i - \sum_{l=1}^{L} b_{i,l}\,\mathbb{E}\left[f_l\right]$*,* ϵ_i *is an idiosyncratic random factor for security* i*, with* $\mathbb{E}\left[\epsilon_i\right] = 0$*,* $\mathbb{E}\left[f_l\,\epsilon_i\right] = 0$ *for all* l *and* $\mathbb{E}\left[\epsilon_i\,\epsilon_{i'}\right] = 0$ *for all* $i \neq i'$*. Without loss of generality, we assume that the factors are orthogonal, or* $\mathbb{E}\left[f_l\,f_p\right] = 0$*, and the market return is the first factor, or* $f_1 = \mathbf{r}^T\boldsymbol{\tau}$*. We use* $\mathbf{b}_l = (b_{1,l}, \ldots, b_{N,l})^T$ *for the vector of factor loadings,* $\rho_{i,l}$ *for the correlation between security* i *and factor* l*,* $R_i^2 = \sum_{l=1}^{L} \rho_{i,l}^2$ *for the percentage explained variance, and* $z_{i,l} = \rho_{i,l}/R_i$ *for a standardized correlation coefficient that yields* $\sum_{l=1}^{L} z_{i,l}^2 = 1$*.*

A common specification is to use the single-factor “market model”($L = 1$). This specification is particularly relevant in the context of analyzing and testing efficiency and linearity. For example, the classical Gibbons, Ross, and Shanken (1989) test for market portfolio efficiency without portfolio restrictions assumes the market model. The market model is also of interest because it implies the classical linear SML as a special case in APT.

Empirically, it appears that the market return explains the bulk of the joint variation of stock returns, and additional risk factors, such as Fama and French’ (1993) hedge portfolio returns, explain much smaller (but sometimes significant) amounts. This finding is not surprising given that the market portfolio by construction is extremely diversified and therefore will be highly correlated with the first principal component of security returns, irrespective of the nature of the underlying risk factor model.

Despite the arguments for the market model, our analysis allows securities to have a joint exposure to non-market factors. A factor model is used here to describe the mutual correlation of asset returns. In turn, the correlation structure plays an important role for analyzing the goodness of the CSML approximation in this study. The use of a factor model should not be confused with using the APT to explain the cross-section of expected returns. Since our analysis allows for binding investment constraints, we cannot fix the intercepts a_i, $i = 1, ..., N$, using the APT, which is based on free portfolio formation.

Theorem 4 (Sufficient condition). *The correlation structure (13)-(14) applies when optimal portfolios* $\boldsymbol{\lambda}_k$*,* $k \in \mathcal{K}$*, have neutral non-market factor loadings, that is,* $\mathbf{b}_l^T \boldsymbol{\lambda}_k = 0$ *for all* $l = 2, \ldots, L$*, and are well diversified, that is,* $\lambda_{i,k}\, \sigma^2(\epsilon_i) \to 0$ *for all* $i \in \mathcal{A}_{\boldsymbol{\lambda}_k}$*. Moreover, we have*

$$\xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} = \frac{1}{\rho_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}}}. \tag{19}$$

The condition of neutral non-market loadings is more general and more plausible than the market model because it applies to optimal portfolios rather than individual assets. Diversified portfolios tend to have relatively small non-market loadings compared with market betas. Risk reduction through diversification encourages the investor to combine stocks with high and low non-market loadings, moving the portfolio's non-market loadings towards zero. In addition, extreme loadings to known non-market risk factors (such as Fama and French' SMB and HML factors) are mostly concentrated in the micro-cap market segment and therefore a broadly diversified portfolio generally has small non-market loadings.

We do not challenge the explanatory power of non-market factors for individual stocks or for benchmark portfolios formed to have maximal exposure to non-market risk factors. We merely say that optimal portfolios are likely to have limited exposure to non-market risk. As a case in point, the market model explains 91.5 (95.8) percent of the return variation of the above-mentioned low-beta (high-beta) portfolio and the Fama-French three-factor model explains only 0.8 (0.3) percent more, and the SMB and HML loadings of these two portfolios are not significant.[1]

Theorem 5 (Errors in Correlation Ratios)**.** *Under Assumption 6 we find*

$$|\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} - \xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}}| \leq \sqrt{\frac{1}{z_{i,\boldsymbol{\tau}}^2} - 1}\sqrt{\sum_{l=2}^{L} \rho_{\boldsymbol{\lambda}_k,l}^2 + \lambda_{i,k}\,\frac{\sigma^2(\epsilon_i)}{\sigma_i\,\sigma_{\boldsymbol{\lambda}_k}}}. \tag{20}$$

where

$$\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} - \sum_{l=2}^{L}\left(\frac{\rho_{\boldsymbol{\lambda}_k,l}}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}\right)\rho_{\boldsymbol{\lambda}_k,l} - \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}\sum_{i=1}^{I}\lambda_{i,k}^2\,\frac{\sigma^2(\epsilon_i)}{\sigma_{\boldsymbol{\lambda}_k}^2}.$$

Clearly, these errors go to zero if the non-market factor loadings go to zero and the portfolio is well-diversified. For example, assume that equilibrium is described by two optimal portfolios equal to the above-mentioned low-beta portfolio and high-beta portfolio and assume further that the Fama-French three-factor model captures all relevant risk factors. For the low-beta portfolio, the market model yields an R-squared of 91.5 percent and the three-factor model explains 92.3 percent. This implies that the first right-hand-side term of (20) equals $\sqrt{(0.923/0.915)-1}\,\sqrt{0.923-0.915} = 0.008$. The portfolio is well-diversified and we can ignore the second right-hand side term. Similarly, the high-beta portfolio involves R-squared values of 95.8 and 96.1 percent and the absolute error in (20) is bounded by $\sqrt{(0.961/0.958)-1}\,\sqrt{0.961-0.958} = 0.003$. The error bounds will be even smaller if the two optimal portfolios overlap.

E Numerical example

We will now illustrate the CSML approximation and the associated alpha-beta relationship with a numerical example using a small number of assets and investors. We use the single-factor model $r_i = a_i + b_i\,f + \epsilon_i$ for ten base assets ($N = 10$), where f is a systematic risk factor with standard deviation σ_f, b_i are the factor loadings and ϵ_i are idiosyncratic random factors with residual standard deviations σ_{ϵ_i}. The assets are assumed to have identical market values, and hence the market portfolio is simply the equal-weighted average of the ten assets. To use realistic parameter values, we calibrated the model parameters to the historical data set of size-neutral "beta decile stock portfolios" that is used in our empirical analysis in Section II. The risk factor is the first principal component of the returns to these ten portfolios.

---

[1] The SMB and HML factors are not orthogonal to the market return, contrary to our Assumption 6. The significance of these factors is therefore established by comparing the explanatory power of the market model and the three-factor model rather than analyzing their explanatory power in isolation.

Table I describes our procedure to calibrate the factor model and reports factor loadings, residual standard deviations, R-squared coefficients, and other relevant information.

We consider a simple case of our heterogeneous-investor model with one conservative investor and one adventurous investor. Both investors are assumed to make combinations of the ten base assets and a riskless Treasury bill without riskless borrowing and short selling. Both investors are assumed to have the same wealth level ($w_1 = w_2 = 0.5$), and the market portfolio therefore is equal to the equal-weighted average of the portfolios of the two investors. We chose the slopes ($(1/\zeta)\,\sigma_{\boldsymbol{\lambda}}$) and intercepts ($\theta_{\boldsymbol{\lambda}}$) of the investors' optimality conditions (see Equation 6) to satisfy the following market equilibrium: (i) the conservative investor optimally combines the five low-beta assets (P_1), (ii) the adventurous investor optimally combines the five high-beta assets (P_2), (iii) the two portfolios aggregate to the market portfolio. Table I shows the resulting expected returns of the individual assets and portfolios P_1 and P_2.

[Table I about here.]

Panel A of Figure 1 shows a mean-variance diagram with the expected returns and standard deviations of the base assets and the risk-free asset, together with the entire portfolio possibilities set and the efficient frontier for the case without riskless borrowing and short sales. We also report the portfolios of the conservative and adventurous investors (P_1 and P_2), together with the market portfolio.

The correlation between all possible portfolios is very high. For example, the correlation between the low-beta portfolio P_1 and the high-beta portfolio P_2 is higher than 98%, despite the two portfolios having no overlapping positions. Therefore, diversification between the base assets does not reduce variance in a material way, and, absent riskless borrowing and short selling, the investor adjusts the risk level of her portfolio by either combining low-beta assets and the risk-free asset and excluding high-beta assets (in the case of low risk tolerance) or combining high-beta assets and excluding low-beta assets and the risk-free asset (high risk tolerance). Paradoxically, diversifying across the optimal portfolios P_1 and P_2 produces inefficient portfolios. Most notably, the market portfolio, a weighted average of the two optimal portfolios, is mean-variance dominated by medium-beta assets. Clearly, the efficient set is not convex in this example.

Panel B and C illustrate the portfolio optimality conditions (2) to (5) for the two investors. For optimal portfolios, the included assets exhibit a linear relation between the expected return and the beta relative to the optimal portfolio; the excluded assets will lie below the line. For example, the conservative investor's portfolio (P_1) is characterized by a linear relation for the low-beta assets, the steep slope reflecting the investor's low risk tolerance. The expected return of the high-beta assets, which are excluded from P_1, is lower than predicted by this relation. This deviation does not reflect that P_1 is inefficient, but rather the positive shadow price of the restriction on short-selling high-beta assets. Panel D shows that the betas relative to P_2 are nearly proportional to the betas relative to P_1, reflecting the very high correlation between the two portfolios. This near-proportionality explains why Panel C shows a very similar, kinked shape as Panel B; the two panels are nearly identical, apart from a scalar multiplication of the betas (and a different active set).

Panel E shows the expected returns and market betas for the ten base assets. By construction, these values obey the GSML relation (see Theorem 1), because the individual portfolios obey the optimality conditions and the market clears in this example. Since portfolios P_1 and P_2 are highly correlated with the market portfolio, the pattern is similar to that in Panel B and C. The dashed, straight line represents Black's (1972) general linear SML, or, in this case, the average of the two straight lines in Panel B and C. Clearly, the market portfolio does not obey the mean-variance optimality conditions. It includes all assets and therefore efficiency requires a linear relation. However, the expected return to medium-beta assets is substantially higher than what is obtained through linear interpolation of the expected returns of low-beta assets and high-beta assets. The solid, kinked line represents our CSML approximation (see Theorem 3), which combines the two sets of optimality conditions under the assumption that P_1 and P_2 have zero non-market factor loadings (see Theorem 2).

Panel F shows the concave relation between market betas and market alphas. The solid, kinked line represents again the CSML approximation (15) of this relation. Whereas the alphas of the low-beta assets show a linear increasing pattern, the high-beta alphas show a linear decreasing pattern, resulting in an overall concave shape. The two-piece linear shape in Panel E and F reflects the use of only two investors in this example; a more general piece-wise linear shape arises if we introduce additional investors with different risk tolerance levels.

[Figure 1 about here.]

II Empirical Analysis

Empirical tests of the CAPM broadly fall into two categories: (i) structural tests of mean-variance efficiency of the market portfolio and (ii) regression analysis of the linear cross-sectional relationship between expected return and market beta (SML). Most empirical tests in the first category focus on the intercepts in time-series regressions of asset returns on the market index in the spirit of Black, Jensen, and Scholes (1972). Tests in the second category build on the classic two-pass regression procedure of Fama and MacBeth (1973). Shanken (1992), Jagannathan and Wang (1998) and Shanken and Zhou (2007) provide important discussions of the statistical properties and performance of this approach.

For several reasons, this section emphasizes the two-pass regression method. This approach has more flexibility to analyze and test the functional form of the SML. The efficiency tests use the null hypothesis that the market portfolio is efficient, or, equivalently, the SML is linear. By contrast, our null hypothesis is that different individual investors hold different efficient portfolios that are dominated by market risk and that combine to yield a concave SML. Portfolio efficiency tests can test this null only if we specify the relevant set of portfolio restrictions, the composition of the individual portfolios and the distribution of wealth. In addition, the two-pass regression method has the flexibility to control for the effect of empirically relevant co-variates (that do not enter in structural models), such as a stock's non-systematic risk and market capitalization.

This section applies the regression method to a set of 100 stock portfolios that are based on the market capitalization and market beta of individual stocks, as well as to the entire cross-section of individual stocks (Section IID).

A Data

We obtain daily and monthly stock returns from the Center for Research in Security Prices (CRSP) and annual balance sheet data from Compustat. Monthly market excess returns (MKTRF) are from Kenneth French's data library.

Our main analysis is based on the period from January 1927 to December 2016. This long sample period seems particularly useful, because it includes the bear markets of the 1930s, 1970s, and 2000s which seems important to avoid the 'Peso problem' for studying the return-risk relation. Following the convention, our analysis focuses on ordinary common U.S. stocks listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and Nasdaq markets, excluding ADRs, REITs, closed-end-funds, units of beneficial interest, and foreign stocks. A stock is excluded from the analysis if price information is no longer available. In that case, the delisting return or partial monthly return provided by CRSP is used as the last return observation. We require stocks to have at least 24 observations available in the past 60 months.

Market betas are estimated using Ordinary least Squares (OLS) regression in a rolling window of 60 months. In order to avoid the well-known sorting bias, or clustering of negative (positive) estimation errors in beta estimates for low-beta (high-beta) portfolios, we use a Bayesian shrinkage method. Following Vasicek (1973) and Elton *et al.* (2014), we shrink the 60-month beta estimate ($\hat{\beta}_{i,ts}$) toward the cross-sectional mean ($\hat{\beta}_{avg}$) as follows:

$$\hat{\beta}_i = w_i \, \hat{\beta}_{i,ts} + (1 - w_i) \, \hat{\beta}_{avg} \tag{21}$$

The Vasicek (1973) shrinkage factor is given by $w_i = 1 - \hat{\sigma}^2_{i,ts}/(\hat{\sigma}^2_{i,ts} + \hat{\sigma}^2_{avg})$ where $\hat{\sigma}^2_{i,ts}$

is the variance of the estimated beta for stock i, and $\hat{\sigma}^2_{avg}$ is the cross-sectional variance of stock betas. This *ex-ante* beta estimate places more weight on the historical time-series estimate when it has a lower variance or when there is a large dispersion in the cross-section of betas. Correcting for the sorting bias is particularly relevant for the cross-sectional regressions. Using the (biased) historical time-series betas would artificially flatten the estimated cross-sectional risk-return relationship.

Empirical research on stock returns has identified several other relevant stock characteristics in addition to market beta; see, for example, Basu (1977,1983), Banz (1981), De Bondt and Thaler (1985), Jegadeesh (1990), Jegadeesh and Titman (1993), Acharya and Pedersen (2005) and Ang, Hodrick, Xing, and Zhang (2006). Our analysis will include popular corrections for book-to-market equity ratio (BtM), market capitalization of equity (ME), short-term reversal (R_1), momentum (R_{12-2}) and long-term reversal (R_{60-13}), illiquidity ($Illiq$), and idiosyncratic volatility ($\hat{\sigma}'(\epsilon)$).

Following Fama and French (1992, 1993), we define size as the market capitalization in billions of dollars. We define the book-to-market equity ratio for the twelve month period starting at the end of June of year t as the ratio of the book value of equity at the end of the fiscal year ending in calendar year $t-1$, to the market capitalization of the stock at the end of December of year $t-1$. For observations where the book value of equity is negative, we set book-to-market ratio to be missing. Following Brennan, Chordia, and Subrahmanyam (1998), the past return variables R_1, R_{12-2} and R_{60-13} are appropriately lagged relative to the test month and are all set equal to the logarithms of cumulative past gross returns. $Illiq$ the prior 12-month moving average of monthly Amihud's 2002 illiquidity ratio, where dollar trading volume is defined in billions of dollars. The idiosyncratic volatility is defined as the residual standard deviation, scaled by the standard deviation of the excess market return over the same time period.

Two additional control variables are based on the growing literature of the low-beta anomaly. A stock's sensitivity to funding constraints is motivated by the evidence provided in Frazzini and Pedersen (2014). We measure a stock's exposure to funding liquidity as the slope coefficient from a regression of the stock's excess returns on the T-bill rate over the prior 60-month period. Following Bali and Cakici (2008), we measure a stock's lottery demand as the average of the five highest daily returns during the prior month (max).

Table II presents summary statistics for the sample of stocks that satisfy our data requirements.

[Table II about here.]

The market betas of individual stocks can be difficult to estimate with high accuracy, due to relatively high stock-specific risk, and possible structural and cyclical changes over time. These problems can be mitigated by forming portfolios of stocks and periodically rebalancing these portfolios, a tradition that goes back to Blume (1970). It is particularly useful to form "beta portfolios" based on the past market beta estimates of individual stocks, as in, for example, Fama and MacBeth (1973). This approach yields relatively accurate and stable market beta estimates and a wide spread of the betas, increasing the statistical power of the analysis. Furthermore, a stable market beta avoids the possible bias that arises if market beta is correlated over time with the market-risk premium, and the associated need to use a conditional model specification; see, for example, Jagannathan and Wang (1996) for a discussion of this problem.

Small-cap stocks appear to carry a return premium that is uncorrelated with their market betas. To disentangle the effects of size and betas, we apply a double-sorting routine based on market capitalization and market beta, following Fama and French (1992).

At the end of each month, all stocks which fulfill the data requirements are assigned to portfolios based on decile breakpoints for the market capitalization of NYSE stocks, and independently based on decile breakpoints for the past 60-month beta estimates for the entire cross-section. The result is 100 test portfolios with independent variation in market capitalization and market beta. The portfolio formation starts in December 1931.

For each test portfolio, 60-month formation period and test-month portfolio returns are computed as equal-weighted averages of the individual stocks' returns. Our rationale for using equal-weighted averages is they appear to reduce the noise level in a more effective way than value-weighted averages. A value-weighted portfolio can be easily dominated by the return and idiosyncratic risk of a few large stocks. However, very similar results are obtained using value-weighted portfolio returns.

We estimate portfolio level beta, idiosyncratic volatility and past return variables the same way as individual stock betas. Portfolio-level BtM is an aggregate ratio: sum of stock-level book values divided by sum of stock-level market values. The remaining variables (ME, *Illiq*) are aggregated to the portfolio level by taking the median of the stock-level values.

B Methodology

We employ the two-pass regression methodology of Fama and MacBeth (1973). The analysis is built around the following (second-pass) cross-sectional regression equation:

$$r_{p,t} - r_{F,t} = \gamma_{0t} + \gamma_{1t}\,\hat{\beta}_{p,t-1} + \gamma_{2t}\,(\hat{\beta}_{p,t-1} - \bar{\beta}_{t-1})^2 + \epsilon_{p,t}. \tag{22}$$

In this equation, $\hat{\beta}_{p,t-1}$ is the estimated market beta for portfolio p and $\bar{\beta}_{t-1}$ is the cross-sectional average of the betas. The regressors are lagged by one month relative to the test month.

The use of squared beta was proposed in Fama and Macbeth (1973). The use of the square of the centered beta is a common reparameterization to mitigate the effect of multicollinearity in a polynomial regression.

The regression equation can be seen as a quadratic approximation to the CSML Equation (15). The theoretical model predicts a general concave, piece-wise linear relation with a different linear line segment for every investor. It is difficult to implement this functional form in practice, because it requires the specification of the relevant number of linear pieces and the associated return intervals. However, the piece-wise linear relation between expected return and market beta approximates a real analytical function $\mathbb{E}\left[r_p - r_F\right] = \mu(\beta_p)$ in case of a continuum of investors with different risk tolerance levels.

A Taylor series expansion around $\beta = \bar{\beta}$ gives

$$\mu(\beta_p) = \gamma_0 + \gamma_1\beta_p + \gamma_2(\beta_p - \bar{\beta})^2 + R_2(\beta_p), \tag{23}$$

where $\gamma_0 = \mu(\bar{\beta}) - \mu'(\bar{\beta})\,\bar{\beta}$, $\gamma_1 = \mu'(\bar{\beta})$, $\gamma_2 = \frac{1}{2}\mu''(\bar{\beta})$, and $R_2(\beta_p)$ is the remainder term. Given that the typical range for betas is limited (say $\beta_p \in [0, 2]$), the remainder term $R_2(\beta_p)$ is expected to be small and the second-degree Taylor polynomial is expected to be accurate, using Taylor's theorem.

Another possible specification is a four-piece linear function with kinks at the first, second and third quartile of the cross-sectional beta distribution. Alternative approaches are the inclusion of higher-order powers of beta to provide a higher-order Taylor series approximation and the use of a non-parametric kernel regression. Interestingly, these approaches yield results that are very similar to those based on beta-squared.

The parameters in the quadratic regression model (22) can be further restricted. Specifically, if it is assumed that the risk-free asset and the market portfolio obey the quadratic mean-beta relation, or $R_2(0) = 0$ and $R_2(1) = 0$, then the following two parameter restrictions are obtained:

$$\gamma_0 = -\gamma_2 \bar{\beta}^2; \tag{24}$$

$$\gamma_1 = (\mu_{\boldsymbol{\tau}} - r_F) + \gamma_2 \left(2\overline{\beta} - 1\right). \tag{25}$$

Imposing these parameter restrictions, the quadratic relation (22) reduces to

$$r_{p,t} - r_{F,t} = (\hat{\mu}_{\boldsymbol{\tau},t-1} - r_{F,t})\, \hat{\beta}_{p,t-1} + \gamma_{2t} \left(\hat{\beta}^2_{p,t-1} - \hat{\beta}_{p,t-1}\right) + \epsilon_{p,t}. \tag{26}$$

Here, $\hat{\mu}_{\boldsymbol{\tau},t-1}$ is an estimate for the expected market return.

The cross-sectional regression is estimated every month using OLS regression analysis to generate a time series of monthly coefficient estimates. The time-series averages and standard deviations of the coefficient estimates are used for testing hypotheses regarding the unknown parameters γ_{0t}, γ_{1t} and γ_{2t}.

The monthly coefficient estimates may show patterns of positive or negative autocorrelation that can cause the original Fama-McBeth standard errors to be biased. This study corrects for autocorrelation by using a Newey and West (1987) correction with 12 monthly lags.

Our results are robust to explicit corrections for cross-serial correlation and heteroskedasticity. Notably, we find similar results for beta and beta-squared with the cluster-robust method that is analyzed by Petersen (2009) or a full-fledged mixed model of fixed-random effects panel method. We therefore report only the standard Fama-McBeth results (with a Newey-West correction) in this study.

The methodological robustness of the results is perhaps unsurprising given that our data set fulfills the conditions that Petersen (2009) identified as favorable for the Fama-MacBeth method, namely a sample in which the autoregressive pattern dies off fast and the time series is long.

Apart from the individual regression coefficients, we also analyze the “average beta premium” (ABP), or the predicted return premium for bearing a market beta of one rather than zero:

$$ABP_t = \left(\gamma_{0t} + \gamma_{1t}\,(1) + \gamma_{2,t}\,(1 - \bar{\beta}_t)^2\right) - \left(\gamma_{0t} + \gamma_{1t}\,(0) + \gamma_{2,t}\,(0 - \bar{\beta}_t)^2\right) = \gamma_{1t} + \gamma_{2t}\,(1 - 2\,\bar{\beta}_t).$$

Given that the market portfolio has a market beta of one, the value of the ABP is expected to be close to the historical excess return to the market portfolio. ABP is also relevant to address possible concerns about multi-collinearity between beta and beta-squared; it is a single parameter that captures the joint effect of beta and beta-squared.

A logarithmic transformation is applied for several regressors (ME, BtM, $Illiq$) in order to allow for a diminishing effect and/or to mitigate the effect of outliers. These considerations are particularly relevant when analyzing individual stocks (rather than stock portfolios), given the relatively wide sample range and high noise level in this case. A logarithmic transformation for market beta is however unusual, as the standard theory explicitly predicts a linear SML. Although our theory predicts a concave SML, we do not consider a logarithmic specification suitable for our analysis. First, it is desirable to use a specification that includes the linear SML as a special case, in order to test the hypothesis of linearity. Second, the logarithmic specification allows for the risk premium to decrease with the beta, but it also assumes that the effect is strongest for low-beta stocks, while the data suggest otherwise. Indeed, if log beta is used instead of beta in our regressions, the empirical fit improves, but log-beta-squared remains significant.

For the sake of brevity, we do not report estimation results for the squared terms of regressors other than market beta. Neither our theoretical analysis nor the empirical literature assigns a role to, for instance, squared residual risk or squared market capitalization; in fact, these regressors are deliberately defined as concave transformations of an underlying variable (the logarithm of market capitalization and the square root of residual variance, for instance). Indeed, none of the additional squared terms has a significant effect or changes the results for beta-squared in a material way in our analysis.

C Portfolio-Level Results

Table III summarizes our results for the entire cross-section of 100 size-beta portfolios and the full time-series of monthly portfolio returns from January 1932 to December 2016. The beta-only regression does not yield a significant market risk premium. Nevertheless, the average Adjusted R-squared is about 20 percent, suggesting that market beta does play a signifcant role in the individual months. The difference between the two statistical goodness measures partly reflects time-variation in the coefficient value: the coefficient takes substantially different, but significant, values in different months. In addition, the R-squared is inflated by the cross-sectional correlation and heteroskedasticity of the SML error terms, whereas the Fama-MacBeth t-statistic avoides the use of cross-sectional standard errors.

The conclusions change considerably if we add beta-squared to the regression. The center panel reveals a significant and concave relationship between return and beta. The "average beta premium" (ABP) is estimated to be 45.8 basis points per month, or 5.5 percent per annum, and significant (t-statistic: 2.69). Consistent with our hypothesis, the coefficient estimate for beta-squared is significantly negative (t-statistic: -3.68), and the risk premium increases at a diminishing rate. Interestingly, the cross-sectional R-squared improves only marginally to 22 percent compared with the linear specification, reflecting that it does not penalize excessive time-series variation and low average values of the estimated beta risk premium.

Including ME and BtM has no material effect on the concave relation and the estimated ABP. Adding the other stock characteristics also does not materially change the results and conclusions about beta and beta squared. Consistent with what is documented elsewhere in the empirical literature, we find negative premiums for ME, R_1, R_{60-13} and max, and positive premiums for BtM, R_{12-2} and $Illiq$.

[Table III about here.]

The bottom panel adds residual risk as an additional regressor. In the first regression, the coefficient of residual risk is estimated to be 68.3 basis points per month, or about 8.2 percent per annum. Residual risk appears to have significant explanatory power, but it does not materially affect the results and conclusions about beta and beta-squared; the return-beta relation remains significantly concave and the estimated ABP even increases to 54 basis points per month. In addition, the inclusion of other control variables significantly diminishes the role of residual risk. ABP, on the other hand, varies from 4.4 to 6.2 percent per annum, similar in magnitude to the historical equity premium of about 6 percent in this period. These results are supportive of our model, which builds on systematic risk and does not assign a role to residual risk.

The weak role of residual risk may reflect that we analyze size-beta portfolios rather than individual stocks, and that the estimated betas of these portfolios are more accurate and stable than the estimated betas of individual stocks. These results confirm the conclusion of Bali and Cakici (2008) that there exists no robust, significant relation between average stock returns and residual risk.

In addition to analyzing the full cross section and time series, we also ran the cross-sectional regressions after excluding the pre-1963 period and/or stocks with share prices below \$5. The concave mean-beta relation appears very robust and the estimated level of ABP and t-statistic even increase slightly in every subsmaple.

Finally, we ran the restricted regression (26), which imposes the coefficient restrictions (24) and (25). The results again confirm the prediction that the inclusion of beta-squared improves the cross-sectional fit. Starting from the "full house" specification (the last row of Panel C), imposing the coefficient restrictions (24) and (25), and excluding beta-squared, reduces the R-squared from 41.4% to 35.1%. Including beta-squared (but retaining the coefficient restrictions) increases the R-squared to 37.9% and yields a t-statistic of -7.24 (!) for the quadratic term.

D Stock-Level Results

A possible concern about the portfolio-level results is that the sorting of stocks and the formation of portfolios affects the statistical size and power of the analysis. Sorting stocks based on ME and beta can produce artificially high sample variation for the portfolio values of those characteristics and artificially low sample variation for others. To address this concern, this section analyzes the cross-section of individual stocks rather than portfolios.

Table IV summarizes our results for the full time-series of monthly excess returns from January 1932 to December 2016. Similar to the results in Table III for size-beta portfolios, the mean-beta relation appears significantly concave and the estimated ABP is economically and statistically significant, hovering around four to six percent per annum.

[Table IV about here.]

Again, the results are robust to changes in the cross section and time series. Excluding the pre-1963 period and/or stocks with share prices below \$5 has limited effect on the estimated ABP and its t-statistic.

III Active Portfolio Construction

A concave mean-beta relation has important implications for active investment strategies that select stocks based on past market-beta estimates. Frazzini and Pedersen (2014) convincingly show that overweighting low-beta stocks and/or underweighting high-beta stocks generates large positive abnormal returns or alphas.

Our model predicts that the alpha is a concave function of market beta (Corollary 1). Therefore, we would expect that the abnormal return of active bets against beta is driven primarily by the high-beta stocks. In our model, these stocks are held by the most adventurous investors (highest risk tolerance) and have the lowest marginal risk premium. To examine this prediction, we evaluate the performance of ten "size-neutral beta-decile portfolios", and various long-short combinations of these portfolios, using time-series regression analysis. Our beta-decile portfolios are constructed from size-dependent beta sort portfolios and by averaging across the ten size segments. This aggregation method helps to control for the confounding effect of market capitalization and also to reduce the noise level.

Our analysis is based on the single-factor market model, the Fama and French three-factor and the Carhart four-factor model. The market model corrects excess returns only for market risk exposure (MKTRF); the three-factor model also corrects for exposures to small-cap stocks (SMB) and value-stocks (HML). Given that our benchmark portfolios are size-neutral, these corrections can be expected to have less effect here than in some other applications. The Carhart model accounts also for medium-term momentum (UMD) exposure, which seems relevant for our purposes, because the stock-level market betas are estimated from the previous 60 monthly returns and the estimates can therefore be confounded with the levels of past return.

We analyze the alphas of the ten beta deciles (BD1 to BD10) and four hedge portfolios formed from the deciles: (i) long decile 1 and short an equal dollar amount in decile 10 ("BD1-10"), (ii) long decile 1 and short decile 5 ("BD1-5"), (iii) long decile 6 and short decile 10 ("BD6-10"), and (iv) long deciles 5 and 6, and short deciles 1 and 10 ("BD5,6-1,10"). The first portfolio, BD1-10, is a pure bet against beta that is consistent with the zero-beta model. To discriminate between the zero-beta model and the CSML, we compare the "low-minus-medium-beta portfolio" BD1-5 with the "medium-minus-high-beta portfolio" BD6-10. Corollary 1 predicts that BD6-10 will have a larger alpha than BD1-5. The "CSML portfolio" BD56-110 takes a long position in medium-beta stocks and a short position in both high-beta stocks and low-beta stocks. Our theory predicts that this portfolio will have a positive alpha, while the zero-beta model predicts that it will have a zero alpha.

In contrast to the cross-sectional regressions in Section IIB, the performance evaluation in this section does not require a Bayesian shrinkage method. The alphas and factor loadings are estimated using the full-sample time-series regressions of the portfolio's post-formation returns, a method that does not suffer from a sorting bias. The estimated factor loadings are therefore not shrunk to the average.

Table V summarizes our estimation results for the base case of equal-weighted portfolios constructed from the entire cross-section of stocks that fulfill our data requirements (see Section II) and the full time-series from January 1932 to December 2016. The first column shows that the four-factor-model alphas display the concave pattern that is predicted by the CSML. For the lowest-beta stocks (BD1), the alpha is 2.1 percent per annum. The alpha first increases with market beta to about 2.2 percent per annum for medium-beta stocks, and then decreases sharply to negative values for higher beta levels. Most notably, the highest-beta portfolio (BD10) has an alpha of negative 3.2 percent per annum. The portfolio BD1-10 yields a combined alpha of 5.3 percent per annum (corresponding 3-factor and 1-factor alphas are about 9.2 and 8.4 percent per annum and highly statistically significant).

This pattern confirms the profitability of betting against beta that is documented in Frazzini and Pedersen (2014). The concave pattern of the alpha implies that the BD1-10 alpha stems mostly from the high-beta stocks. Indeed, the low-minus-medium-beta portfolio BD1-5 generates an alpha of negative 0.1 percent , while the medium-minus-high-beta portfolio BD6-10 generates 4.5 percent abnormal return. The "CSML portfolio" BD5,6-1,10, which longs medium-beta stocks and shorts high- and low-beta stocks, yields a large positive alpha of 4.6 percent per annum (t-stat: 2.84).

Our objective here is not to horse race the CSML portfolio (BD5,6-1,10) against the pure bet-against-beta (BD1-10), but to show that the profitability of both portfolios stems from the alpha spread in the high-beta market segment (BD6-10) rather than the low-beta segment (BD1-5). The success of betting-against-beta strategies thus seems to largely depend on the underweighting or short selling of high-beta stocks.

Due to the use of size-neutral portfolios, the differences in the non-market factor loadings between the ten portfolios are modest. Still, the non-market exposures have a non-trivial effect and the alpha spread widens if we use either the market model or the three-factor model, as shown in the last two columns of Table V. The concave pattern is apparent for these two model specifications as well.

[Table V about here.]

Figure 2 plots the estimated four-factor-model alphas against the estimated market betas and clearly illustrates the predicted concave pattern. Apart from the full-sample time-series regression estimates of the market betas, the figure also includes the time-series average of the 60-month Bayesian estimates (used in Section IIA). The two sets of estimates are remarkably similar, which further illustrates robustness of our results.

[Figure 2 about here.]

As the cross-sectional regressions, the time-series regressions are robust to the exclusion of the early sample period. Our results are also robust to using the market model or three-factor model and excluding stocks with share price below \$5. In every sample, we see that the medium-minus-high-beta spread (BD6-10) significantly exceeds the low-minus-medium-beta spread (BD1-5).

As a final step in our analysis, we attempt to include the non-linearity in the factor models using Non-linear Least Squares (NLS). We model the intercept of the regressions as a quadratic function of the market beta, and jointly solve the alpha estimation problem for the 10 beta-decile portfolios. In line with our earlier arguments and findings, the resulting alphas are economically and statistically insignificant.

Overall, these findings seem to be consistent with, and supportive of, the conclusions from our theoretical model (Section I) and cross-sectional regression analysis (Section II). These results also illustrate the possible practical relevance of our analysis for active portfolio managers who select stocks based on past market betas. Whether a portfolio manager can benefit from bets against beta in practice will of course depend on the investment restrictions and transactions costs that she faces.

IV Conclusions

Binding investment restrictions generally distort the classical linear relation between expected return and market beta (SML) in a systematic way. In this case, the expected return to a given risky security will reflect the relative risk tolerance of investors who include the security in their optimal portfolios and the covariance of the security with those portfolios. To the extend that the optimal portfolios are correlated and share a common exposure to market risk, expected return will tend to be a concave function of the market beta (rather than the traditional linear function). Indeed, a high positive correlation and substantial joint exposure to market risk is typical for the portfolios of many mutual funds and institutional investors, even if their portfolio composition shows large differences. The concave relation arises in the framework of Sharpe (1991) with mean-variance preferences and without short-selling and borrowing, but also for more general preferences and restrictions (see the discussion in Section IA).

Sharpe (1991) concludes that advances in financial technology and knowledge relax investment constraints and will move markets closer the assumptions of the CAPM. However, closeness to the assumptions is not sufficient for closeness to the equilibrium conditions. Restrictions tend to make the individual portfolios more similar and relaxations introduce more possibilities for differences between individual portfolios. In addition, relaxations generally will not affect all investors in the same way. Many institutional investors and mutual funds face legal or contractual constraints on borrowing and short selling that cannot be relaxed by means of clever financial engineering. In such situations, financial innovation may in effect increase differences between investors' active sets. For these reasons, it is not obvious that financial innovation moves the market closer to a linear SML.

The pattern of expected returns ultimately seems an empirical issue. Our empirical analysis reveals a significant and robust, concave relation between average return and estimated market beta for stocks, consistent with our hypothesis. The inclusion of beta-squared in the regression yields a beta-risk premium of at least four to six percent for the average stock (with a market beta of one), substantially higher than conventional estimates and statistically highly significant. In addition, beta-squared has a significantly negative coefficient, implying that the risk premium increases at a diminishing rate. Encouragingly, the role of concavity is robust to the inclusion of other stock characteristics and the selection of the cross-section and sample period. Concavity appears not only in the analysis of individual stocks but also for aggregated size-beta portfolios, which reduces concerns about estimation error and time-variation of stock-level betas. Similar results are found when using alternative functional forms and using a variety of regression methods.

A similar relation can however not be expected for sets of portfolios which have a limited beta spread or dynamic betas. An unreported analysis replicates our results using double-sorted portfolios formed on market capitalization and one of three other characteristics: BtM, Operating Profitability and Investments.

All three sets of portfolios show limited cross-sectional variation of the market betas and high cyclical variation of the betas (as in Lettau and Ludvigson (2001)). It is therefore not surprising that the estimated betas and their squares in these data sets have limited explanatory power for the cross-sectional variation of the average returns.

These empirical findings confirm our theoretical analysis and contrast with the empirical results of Fama and MacBeth (1973), who conclude that beta-squared plays no significant role and that the SML appears linear. A closer look at the original results of Fama and MacBeth (1973, p. 623, Table 3D) reveals that the coefficient of beta-squared is actually significantly negative in the only sub-period that shows a significant role for market beta (1946-1955). However, beta-squared is not significant in their full sample period, 1935-1968, which is comparable to our early sub-period, 1932-1963. A further analysis reveals that the different results and conclusions can be explained by differences in the portfolio construction procedure. The original study uses twenty beta-portfolios, without controlling for market capitalization, making it difficult to disentangle the competing effects of market beta and market capitalization. Most notably, (single-sorted) high-beta portfolios tend to include a disproportional number of micro-cap and small-cap stocks, and these portfolios benefit from the return premium that these stocks earn (independent of their risk levels), obscuring the underlying mean-beta relationship. Analyzing size-beta portfolios or individual stocks introduces more independent variation in market beta and market capitalization and allows for disentangling the two competing effects.

Consistent with our regression results, the alpha spread of active ‘bets against beta’ is dominated by the medium-minus-high-beta spread rather than the low-minus-medium-beta spread. The success of such strategies thus largely depends on the underweighting or short selling of high-beta stocks. Whether an active manager can benefit from bets against beta will of course depend on the investment restrictions and transactions costs that she faces in practice.

A Proofs

A Proof of Theorem 1

Multiplying the optimality condition (2) for investor k by ζ_k implies

$$\zeta_k \, \mu_i = \operatorname{Cov}\left(r_i, \mathbf{r}^T \boldsymbol{\lambda}_k\right) + \zeta_k \, \theta_{\boldsymbol{\lambda}_k} + \zeta_k \, \alpha_{i,\boldsymbol{\lambda}_k} \tag{27}$$

for all $i \in \mathcal{I}$ and all $k \in \mathcal{K}$. Aggregating (27) across investors $k \in \mathcal{K}$ using wealth shares $w_k/(1+\lambda_{N+1,k})$ invested in the risk securities yields

$$\sum_{k\in\mathcal{K}} (w_k/(1+\lambda_{N+1,k}))\, \zeta_k \, \mu_i = \sum_{k\in\mathcal{K}} (w_k/(1+\lambda_{N+1,k})) \left(\operatorname{Cov}\left(r_i, \mathbf{r}^T \boldsymbol{\lambda}_k\right) + \zeta_k \, \theta_{\boldsymbol{\lambda}_k} + \zeta_k \, \alpha_{i,\boldsymbol{\lambda}_k}\right)$$

which is equivalent to

$$\bar{\zeta} \, \mu_i = \operatorname{Cov}\left(r_i, \mathbf{r}^T \boldsymbol{\tau}\right) + \bar{\zeta} \, \bar{\theta} + \bar{\zeta} \, \bar{\alpha}_i. \tag{28}$$

Replacing $\operatorname{Cov}\left(r_i, \mathbf{r}^T \boldsymbol{\tau}\right)$ with $\beta_{i,\boldsymbol{\tau}} \, \sigma^2_{\boldsymbol{\tau}}$ and dividing by $\overline{\zeta}$ implies (10). □

B Proof of Theorem 2

The optimality conditions (2)-(5) imply the following set of inequalities and equalities

$$\begin{cases} \mu_i \le \left(\frac{1}{\zeta_k} \, \sigma^2_{\boldsymbol{\lambda}_k}\right) \beta_{i,\boldsymbol{\lambda}_k} + \theta_{\boldsymbol{\lambda}_k} & \forall i \in \mathcal{I} \text{ and } k \in \mathcal{K} \text{ such that } i \notin \mathcal{A}_{\boldsymbol{\lambda}_k}, \\ \mu_i = \left(\frac{1}{\zeta_k} \, \sigma^2_{\boldsymbol{\lambda}_k}\right) \beta_{i,\boldsymbol{\lambda}_k} + \theta_{\boldsymbol{\lambda}_k} & \forall i \in \mathcal{I} \text{ and } k \in \mathcal{K} \text{ such that } i \in \mathcal{A}_{\boldsymbol{\lambda}_k}. \end{cases} \tag{29}$$

The portfolio beta

$$\beta_{i,\boldsymbol{\lambda}} = \frac{\operatorname{Cov}(r_i, x^T \boldsymbol{\lambda})}{\sigma^2_{\boldsymbol{\lambda}}} = \rho_{i,\boldsymbol{\lambda}} \, \frac{\sigma_i}{\sigma_{\boldsymbol{\lambda}}}$$

relative to risky portfolio $\boldsymbol{\lambda} \in \Lambda$ can be rewritten in terms of the market beta:

$$\beta_{i,\boldsymbol{\lambda}_k} = \frac{\sigma_{\boldsymbol{\tau}}}{\sigma_{\boldsymbol{\lambda}_k}} \, \frac{\rho_{i,\boldsymbol{\lambda}_k}}{\rho_{i,\boldsymbol{\tau}}} \, \beta_{i,\boldsymbol{\tau}} = \frac{\sigma_{\boldsymbol{\tau}}}{\sigma_{\boldsymbol{\lambda}_k}} \, \xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \, \beta_{i,\boldsymbol{\tau}} \quad \forall i = 1, \ldots, N, \; k \in \mathcal{K}. \tag{30}$$

Under Assumption 5, $\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}$ for all $i \in \mathcal{A}_{\boldsymbol{\lambda}_k}$ and $\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \le \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}$ for all $i \notin \mathcal{A}_{\boldsymbol{\lambda}_k}$, and, therefore, (29) can be written as

$$\begin{cases} \mu_i \le \left(\frac{1}{\zeta_k} \, \sigma_{\boldsymbol{\lambda}_k} \, \sigma_{\boldsymbol{\tau}} \, \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}\right) \beta_{i,\boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} & \forall i = 1, \ldots, N \text{ and } k \in \mathcal{K} \text{ such that } i \notin \mathcal{A}_{\boldsymbol{\lambda}_k}, \\ \mu_i = \left(\frac{1}{\zeta_k} \, \sigma_{\boldsymbol{\lambda}_k} \, \sigma_{\boldsymbol{\tau}} \, \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}\right) \beta_{i,\boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} & \forall i = 1, \ldots, N \text{ and } k \in \mathcal{K} \text{ such that } i \in \mathcal{A}_{\boldsymbol{\lambda}_k}. \end{cases} \tag{31}$$

In equilibrium any risky security $i = 1, \ldots, N$ is included in at least some portfolio, that is, $i \in \mathcal{A}_{\boldsymbol{\lambda}_k}$, for some $k \in \mathcal{K}$. Therefore, aggregating (31) across investors $k \in \mathcal{K}$ yields

$$\mu_i = \min_{k \in \mathcal{K}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right].$$

The slope of the relation between μ_i and $\beta_{i,\boldsymbol{\tau}}$ is positive since $\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \geq 0$. The minimum over increasing linear functions is always an increasing and concave function. $\square$

C Proof of Theorem 3

The CSML (15) can be written as

$$\hat{\mu}_i = \min_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right].$$

Moreover, inspection of the proof of Theorem 2 shows that the GSML (10) is equivalent to

$$\mu_i = \min_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right].$$

It follows that

$$\hat{\mu}_i - \mu_i = \min_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right] - \min_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left[\left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}}\, \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} + \theta_{\boldsymbol{\lambda}_k} \right]$$

and therefore

$$\begin{aligned}
|\hat{\mu}_i - \mu_i| &\leq \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \left(\frac{1}{\zeta_k}\, \sigma_{\boldsymbol{\lambda}_k}\, \sigma_{\boldsymbol{\tau}} \right) \left(\xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} - \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} \right| \\
&= \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \left(\frac{\mu_{\boldsymbol{\lambda}_k} - \theta_{\boldsymbol{\lambda}_k}}{\sigma_{\boldsymbol{\lambda}_k}}\, \sigma_{\boldsymbol{\tau}} \right) \left(\xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} - \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} \right| \\
&\leq \sigma_{\boldsymbol{\tau}} \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \frac{\mu_{\boldsymbol{\lambda}_k} - r_f}{\sigma_{\boldsymbol{\lambda}_k}} \right| \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \left(\xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} - \xi_{\boldsymbol{i}, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right) \beta_{i, \boldsymbol{\tau}} \right| \\
&\leq \sigma_{\boldsymbol{\tau}} \left| \frac{\mu_{\boldsymbol{\lambda}^\star} - r_f}{\sigma_{\boldsymbol{\lambda}^\star}} \right| \beta_{i, \boldsymbol{\tau}} \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} - \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right| \\
&= |\mu_{\boldsymbol{\lambda}^\star} - r_f| \frac{\sigma_{\boldsymbol{\tau}}}{\sigma_{\boldsymbol{\lambda}^\star}} \beta_{i, \boldsymbol{\tau}} \max_{k \,:\, i \in \mathcal{A}_{\boldsymbol{\lambda}_k}} \left| \xi_{\boldsymbol{\lambda}_k, \boldsymbol{\tau}} - \xi_{i, \boldsymbol{\lambda}_k, \boldsymbol{\tau}} \right|.
\end{aligned}$$

The last two inequalities use $\mu_{\boldsymbol{\lambda}_k} - \theta_{\boldsymbol{\lambda}_k} \leq \mu_{\boldsymbol{\lambda}_k} - r_f$ and $\max_{k \in \mathcal{K}} \left| \frac{\mu_{\boldsymbol{\lambda}_k} - r_f}{\sigma_{\boldsymbol{\lambda}_k}} \right| \leq \left| \frac{\mu_{\boldsymbol{\lambda}^\star} - r_f}{\sigma_{\boldsymbol{\lambda}^\star}} \right|$, respectively. $\square$

D Proof of Theorem 4

Under Assumption 6, we find:

$$\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} + \sum_{l=2}^{L} \left(\frac{\rho_{i,l}}{\rho_{i,\boldsymbol{\tau}}}\right) \rho_{\boldsymbol{\lambda}_k,l} + \lambda_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma_i\, \sigma_{\boldsymbol{\lambda}_k}}. \tag{32}$$

Using

$$\frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} = \rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} + \frac{1-\rho^2_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}.$$

and

$$1 = \rho_{\boldsymbol{\lambda}_k,\boldsymbol{\lambda}_k} = \rho^2_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} + \sum_{l=2}^{L} \rho^2_{\boldsymbol{\lambda}_k,l} + \sum_{i=1}^{I} \lambda^2_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma^2_{\boldsymbol{\lambda}_k}}$$

we obtain

$$\frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} = \rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} + \sum_{l=2}^{L} \left(\frac{\rho_{\boldsymbol{\lambda}_k,l}}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}\right) \rho_{\boldsymbol{\lambda}_k,l} + \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} \sum_{i=1}^{I} \lambda^2_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma^2_{\boldsymbol{\lambda}_k}}.$$

It follows that

$$\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} - \sum_{l=2}^{L} \left(\frac{\rho_{\boldsymbol{\lambda}_k,l}}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}\right) \rho_{\boldsymbol{\lambda}_k,l} - \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} \sum_{i=1}^{I} \lambda^2_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma^2_{\boldsymbol{\lambda}_k}}.$$

Placing this equation in (32) gives

$$\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} \quad = \quad \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} + \sum_{l=2}^{L} \left(\frac{\rho_{i,l}}{\rho_{i,\boldsymbol{\tau}}}\right) \rho_{\boldsymbol{\lambda}_k,l} + \lambda_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma_i\, \sigma_{\boldsymbol{\lambda}_k}} \tag{33}$$

where

$$\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} - \sum_{l=2}^{L} \left(\frac{\rho_{\boldsymbol{\lambda}_k,l}}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}\right) \rho_{\boldsymbol{\lambda}_k,l} - \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}} \sum_{i=1}^{I} \lambda^2_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma^2_{\boldsymbol{\lambda}_k}}.$$

Finally, using $\rho_{\boldsymbol{\lambda}_k,l} = \mathbf{b}_l^T\, \boldsymbol{\lambda}_k\, \sigma_{f_l}/\sigma_{\boldsymbol{\lambda}_k} = 0$ for all $l = 2,\ldots,L$ and $\lambda_{i,k}\, \sigma^2(\epsilon_i) \to 0$ for all $i \in \mathcal{A}_{\boldsymbol{\lambda}_k}$ and $\lambda_{i,k} = 0$ for all $i \notin \mathcal{A}_{\boldsymbol{\lambda}_k}$, Equation (33) implies

$$\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}$$

for all $i = 1,\ldots,N$ and

$$\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} = \frac{1}{\rho_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}}}.$$

E Proof of Theorem 5

Inspection of the proof of Theorem 4 (see Equation (33)) shows that

$$|\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} - \xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}}| = \left| \sum_{l=2}^{L} \left(\frac{\rho_{i,l}}{\rho_{i,\boldsymbol{\tau}}} \right) \rho_{\boldsymbol{\lambda}_k,l} + \lambda_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma_i \, \sigma_{\boldsymbol{\lambda}_k}} \right| \leq \left| \sum_{l=2}^{L} \left(\frac{\rho_{i,l}}{\rho_{i,\boldsymbol{\tau}}} \right) \rho_{\boldsymbol{\lambda}_k,l} \right| + \lambda_{i,k} \frac{\sigma^2(\epsilon_i)}{\sigma_i \, \sigma_{\boldsymbol{\lambda}_k}}.$$

We have

$$\sum_{l=2}^{L} \left(\frac{\rho_{i,l}}{\rho_{i,\boldsymbol{\tau}}} \right) \rho_{\boldsymbol{\lambda}_k,l} = \sum_{l=2}^{L} \left(\frac{z_{i,l}}{z_{i,\boldsymbol{\tau}}} \right) \rho_{\boldsymbol{\lambda}_k,l} = \sum_{l=2}^{L} m_{i,l} \, \rho_{\boldsymbol{\lambda}_k,l} = f(m_{i,2}, \ldots, m_{i,L})$$

where $m_{i,l} = z_{i,l}/z_{i,\boldsymbol{\tau}}$ for $l = 2, \ldots, L$. Using $z_{i,\boldsymbol{\tau}}^2 + \sum_{l=2}^{L} z_{i,l}^2 = 1$ we obtain $\sum_{l=2}^{L} m_{i,l}^2 = 1/z_{i,\boldsymbol{\tau}}^2 - 1$.

We now maximize $f(m_{i,2}, \ldots, m_{i,L})$ over $m_{i,2}, \ldots, m_{i,L}$ under the constraint $\sum_{l=2}^{L} m_{i,l}^2 = 1/z_{i,\boldsymbol{\tau}}^2 - 1$. The Lagrange function is

$$L(m_{2,l}, \ldots, m_{i,L}, \eta) = \sum_{l=2}^{L} m_{i,l} \, \rho_{\boldsymbol{\lambda}_k,l} - \eta \left(\sum_{l=2}^{L} m_{i,l}^2 - 1/z_{i,\boldsymbol{\tau}}^2 + 1 \right)$$

and the first order conditions are

$$\rho_{\boldsymbol{\lambda}_l,f_l} - 2\,\eta\, m_{i,l} = 0 \tag{34}$$

$$\sum_{l=2}^{L} m_{i,l}^2 - 1/z_{i,\boldsymbol{\tau}}^2 + 1 = 0. \tag{35}$$

Equation (34) implies

$$m_{i,l}^{\star} = \frac{\rho_{\boldsymbol{\lambda}_l,l}}{2\,\eta}.$$

We insert this latter expression in (35) and obtain

$$\frac{1}{2\,\eta} = \pm \frac{\sqrt{\frac{1}{z_{i,\boldsymbol{\tau}}^2} - 1}}{\sqrt{\sum_{l=2}^{L} \rho_{\boldsymbol{\lambda}_k,l}^2}}$$

and thus

$$m_{i,l}^{\star} = \pm \frac{\sqrt{\frac{1}{z_{i,\boldsymbol{\tau}}^2} - 1}}{\sqrt{\sum_{l=2}^{L} \rho_{\boldsymbol{\lambda}_k,l}^2}} \, \rho_{\boldsymbol{\lambda}_k,l}.$$

It follows

$$\left|\sum_{l=2}^{L} m_{i,l}\,\rho_{\boldsymbol{\lambda}_k,l}\right| \leq \sum_{l=2}^{L} \left|m^{\star}_{i,l}\right|\,\rho_{\boldsymbol{\lambda}_k,l} = \sqrt{\frac{1}{z^2_{i,\boldsymbol{\tau}}}-1}\,\sqrt{\sum_{l=2}^{L}\rho^2_{\boldsymbol{\lambda}_k,l}}.$$

Therefore

$$|\xi_{\boldsymbol{\lambda}_k,\boldsymbol{\tau}} - \xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}}| \leq \sqrt{\frac{1}{z^2_{i,\boldsymbol{\tau}}}-1}\,\sqrt{\sum_{l=2}^{L}\rho^2_{\boldsymbol{\lambda}_k,l}} + \lambda_{i,k}\,\frac{\sigma^2(\epsilon_i)}{\sigma_i\,\sigma_{\boldsymbol{\lambda}_k}}.$$

□

References

Acharya, V. V., and L. H. Pedersen (2005): "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77(2), 375–410.

Amihud, Y. (2002): "Illiquidity and Stock Returns: Cross-section and Time-series Effects," *Journal of Financial Markets*, 5(1), 31 – 56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61(1), 259–299.

Bali, T. G., and N. Cakici (2008): "Idiosyncratic Volatility and the Cross Section of Expected Returns," *Journal of Financial and Quantitative Analysis*, 43(1), 29–58.

Banz, R. (1981): "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9(1), 3–18.

Basu, S. (1977): "Investment Performance of Common Stocks in Relationship to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, 32(3), 663–682.

Basu, S. (1983): "The Relationship Between Earnings' Yield, Market Value and Return for NYSE Common Stocks: Further Evidence," *Journal of Financial Economics*, 12(1), 129–156.

Bawa, V., and E. Lindenberg (1977): "Capital Market Equilibrium in a Mean-Lower Partial Moment Framework," *Journal of Financial Economics*, 5(2), 189–200.

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

Black, F., M. Jensen, and M. Scholes (1972): "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets*, ed. by M. Jensen, vol. 81, pp. 79–121. Praeger Publishers, New York.

Blume, M. E. (1970): "Portfolio Theory: A Step Toward Its Practical Application," *Journal of Business*, 43(2), 152–73.

Brennan, M. J., T. Chordia, and A. Subrahmanyam (1998): "Alternative Factor Specifications, Security Characteristics, and the Cross-section of Expected Stock Returns," *Journal of Financial Economics*, 49(3), 345–373.

De Bondt, A., and R. Thaler (1985): "Does the Stock Market Overreact," *Journal of Finance*, 40(3), 793–805.

Dittmar, R. F. (2002): "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns," *Journal of Finance*, 57(1), 369–403.

Elton, E. J., M. J. Gruber, S. J. Brown, and W. N. Goetzmann (2014): *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons.

Fama, E., and K. French (1992): "The Cross-section of Expected Stock Returns," *Journal of Finance*, 47(2), 427–465.

——— (1993): "Common Risk-factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1), 3–56.

Fama, E. F., and J. D. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81(3), 607–636.

Frazzini, A., and L. H. Pedersen (2014): "Betting Against Beta," *Journal of Financial Economics*, 111(1), 1–25.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989): “A Test of the Efficiency of a Given Portfolio,” *Econometrica*, 57(5), 1121–1152.

Jagannathan, R., and Z. Wang (1996): “The Conditional CAPM and the Cross-Section of Expected Returns,” *Journal of Finance*, 51(1), 3–53.

——— (1998): “An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression,” *Journal of Finance*, 53(4), 1285–1309.

Jegadeesh, N. (1990): “Evidence of Predictable Behavior of Security Returns,” *Journal of Finance*, 45(3), 881–898.

Jegadeesh, N., and S. Titman (1993): “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *The Journal of Finance*, 48(1), 65–91.

Kandel, S., and R. F. Stambaugh (1995): “Portfolio Inefficiency and the Cross-Section of Expected Returns,” *Journal of Finance*, 50(1), 157–184.

Lettau, M., and X. Ludvigson (2001): “Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying,” *Journal of Political Economy*, 109(6), 1238–1287.

Levy, H., and H. Markowitz (1979): “Approximating Expected Utility by a Function of Mean and Variance,” *American Economic Review*, 69(3), 308–317.

Lustig, H., S. Van Nieuwerburgh, and A. Verdelhan (2013): “The Wealth-Consumption Ratio,” *The Review of Asset Pricing Studies*, 3(1), 38–94.

Newey, W. K., and K. D. West (1987): “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55(3), 703–708.

Petersen, M. A. (2009): "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, 22(1), 435–480.

Post, T. (2003): "Empirical Tests for Stochastic Dominance Efficiency," *Journal of Finance*, 58(5), 1905–1932.

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

Roll, R., and S. A. Ross (1994): "On the Cross-sectional Relation between Expected Returns and Betas," *Journal of Finance*, 49(1), 101–121.

Ross, S. A. (1977): "The Capital Asset Pricing Model (CAPM), Short-Sale Restrictions and Related Issues," *Journal of Finance*, 32(1), 177–183.

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

Shanken, J., and G. Zhou (2007): "Estimating and testing beta pricing models: Alternative methods and their performance in simulations," *Journal of Financial Economics*, 84(1), 40–86.

Sharpe, W. F. (1991): "Capital Asset Prices with and without Negative Holdings," *Journal of Finance*, 46(2), 489–509.

Vasicek, O. A. (1973): "A Note on Using Cross-Sectional Information in Baysian Estimation of Security Betas," *The Journal of Finance*, 28(5), 1233–1239.

Table I: Numerical example. We use the single-factor model $r_i = a_i + b_i\, f + \epsilon_i$ for ten base assets ($N = 10$) with identical market values. We calibrated the model parameters to the historical data set of ten beta decile stock portfolios that is used in Section III of our empirical analysis. The risk factor f is the first principal component of the returns to these ten portfolios. The risk factor f has a standard deviation $\sigma_f = 23.1\%$ and displays a correlation of 90.2% with CRSP market returns. The loadings b_i, the residual standard deviations σ_{ϵ_i} and R-squares are obtained by regressing the portfolio returns against the risk factor f. The table also shows expected asset returns that are consistent with an equilibrium with one conservative investor who optimally holds all five low-beta assets (P_1) and one adventurous investor who optimally holds all five high-beta assets (P_2). The two investors are assumed to have identical wealth levels ($w_1 = w_2 = 0.5$), and therefore the market portfolio (Mkt) is required to equal the equal-weighted average of the two portfolios. The table also reports the negative of shadow prices $\alpha_{\boldsymbol{\lambda}}$ and the market alphas $\bar{\alpha}_i$.

	Factor model			Descriptives			Weights($\boldsymbol{\lambda}_k$)			Corr. ratios ($\xi_{i,\boldsymbol{\lambda}_k,\boldsymbol{\tau}}$)		Alphas		
Base asset	Loading (b_i)	Residual (σ_{ϵ_i})	R-squared	Exp. return	St. dev.	Mkt beta	P_1	P_2	Mkt	P_1	P_2	P_1	P_2	Mkt
Low	0.183	2.008	0.815	1.035	4.647	0.593	0.200	0.000	0.100	1.016	0.985	0.000	-0.320	0.129
2	0.225	1.767	0.896	1.173	5.538	0.741	0.200	0.000	0.100	1.007	0.991	0.000	-0.193	0.119
3	0.247	1.409	0.942	1.231	5.896	0.806	0.200	0.000	0.100	1.001	0.995	0.000	-0.140	0.113
4	0.274	1.302	0.959	1.305	6.414	0.883	0.200	0.000	0.100	1.000	0.996	0.000	-0.072	0.109
5	0.297	1.189	0.971	1.383	6.976	0.965	0.200	0.000	0.100	1.000	0.996	0.000	0.000	0.105
6	0.317	1.034	0.980	1.388	7.391	1.027	0.000	0.200	0.100	0.993	1.001	-0.048	0.000	0.048
7	0.343	1.014	0.984	1.395	8.020	1.115	0.000	0.200	0.100	0.993	1.000	-0.126	0.000	-0.034
8	0.366	0.948	0.988	1.400	8.497	1.184	0.000	0.200	0.100	0.993	1.000	-0.187	0.000	-0.098
9	0.392	1.444	0.975	1.407	9.141	1.268	0.000	0.200	0.100	0.992	1.001	-0.259	0.000	-0.176
High	0.432	2.595	0.936	1.418	10.358	1.418	0.000	0.200	0.100	0.987	1.005	-0.383	0.000	-0.314
T-bill	0.000	0.000	0.000	0.311	0.000	0.000	0.000	0.000	0.000			-0.145	-1.000	0.000
P_1	0.245	0.688	0.985	1.225	5.701	0.797			0.500					0.115
P_2	0.371	0.698	0.993	1.402	8.580	1.203			0.500					-0.115
Mkt	0.308	0.497	0.993	1.313	7.123	1.000								0.000

Table II: Descriptive statistics. This table reports summary statistics over the sample period from Januray 1932 to December 2016. The sample contains all ordinary U.S. shares (share codes 10 and 11) that are listed on the NYSE, AMEX and Nasdaq. Panel A presents the time-series averages of cross-sectional means and standard deviations of firm characteristics, whereas Panel B reports the time-series averages of cross-sectional correlation coefficients. The variables reported are monthly excess return in percent per month (ret), market beta ($\hat{\beta}$), centered beta squared ($(\hat{\beta}-\bar{\beta})^2$), standardized residual risk ($\hat{\sigma}'(\epsilon)$), log market capitalization (ME), log book-to-market ratio (BtM), one-month lagged one-month return (R_1), one-month lagged 11-month return (R_{12-2}), 12-month lagged 48-month return (R_{60-13}) (all three past return measures are defined as log cumulative past returns), log illiquidity ($Illiq$), lottery demand in percent per month (max) and Treasury Bill sensitivity ($\hat{\beta}_{RF}$).

Panel A: Time-series averages of cross-sectional means and standard deviations.

	ret	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$
mean	1.05	1.13	0.21	2.45	-2.85	-0.23	0.00	0.04	0.23	-2.44	3.51	7.67
std	11.12	0.45	0.29	1.25	1.80	0.85	0.11	0.35	0.69	2.42	2.55	17.37

Panel B: Time-series averages of cross-sectional correlation coefficients.

	ret	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$
ret	1	-0.02	0.00	-0.03	0.00	0.02	-0.06	0.04	0.00	0.00	-0.05	0.01
$\hat{\beta}$	-0.02	1	0.06	0.42	-0.06	0.02	-0.02	-0.04	-0.04	-0.01	0.29	-0.14
$(\hat{\beta}-\bar{\beta})^2$	0.00	0.06	1	-0.02	-0.01	0.00	0.00	-0.01	-0.01	0.00	-0.01	0.03
$\hat{\sigma}'(\epsilon)$	-0.03	0.42	-0.02	1	-0.53	0.05	-0.04	-0.08	-0.19	0.40	0.48	-0.15
ME	0.00	-0.06	-0.01	-0.53	1	-0.32	0.08	0.20	0.32	-0.82	-0.34	0.11
BtM	0.02	0.02	0.00	0.05	-0.32	1	0.02	0.02	-0.37	0.29	0.11	0.01
R_1	-0.06	-0.02	0.00	-0.04	0.08	0.02	1	0.02	0.01	-0.05	0.34	0.03
R_{12-2}	0.04	-0.04	-0.01	-0.08	0.20	0.02	0.02	1	0.02	-0.21	-0.15	0.07
R_{60-13}	0.00	-0.04	-0.01	-0.19	0.32	-0.37	0.01	0.02	1	-0.28	-0.22	0.01
$Illiq$	0.00	-0.01	0.00	0.40	-0.82	0.29	-0.05	-0.21	-0.28	1	0.31	-0.09
max	-0.05	0.29	-0.01	0.48	-0.34	0.11	0.34	-0.15	-0.22	0.31	1	-0.05
$\hat{\beta}_{RF}$	0.01	-0.14	0.03	-0.15	0.11	0.01	0.03	0.07	0.01	-0.09	-0.05	1

Table III: Cross-sectional regressions for stock portfolios. We compute cross-sectional Fama-MacBeth regressions of monthly excess returns on stock characteristics for 100 bivariate independent-sort size-beta stock portfolios. We require individual stocks to have at least 24 valid return observations available in the past 60 months to be included in a given portfolio. Regressions are run each month from January 1932 to December 2016. The characteristics include past 60-month market beta ($\hat{\beta}$), centered beta squared ($(\hat{\beta}-\bar{\beta})^2$), standardized residual risk ($\hat{\sigma}'(\epsilon)$), log market capitalization (ME), log book-to-market ratio (BtM), one-month lagged one-month return (R_1), one-month lagged 11-month return (R_{12-2}), 12-month lagged 48-month return (R_{60-13}) (all three past return measures are defined as log cumulative past returns), log illiquidity ($Illiq$), lottery demand (max), and Treasury Bill sensitivity ($\hat{\beta}_{RF}$). The reported coefficients are time-series averages of the monthly regression slopes ($\bar{\gamma}$). The Newey-West corrected t-statistics of these averages are shown in parentheses. We also report adjusted R-squares (Adj.R^2) for the regressions, and the average beta premium ($ABP = \gamma_1 + \gamma_2\,(1-2\,\bar{\beta})$).

Panel A: Beta-only.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.931	0.138											0.204	0.138
	(6.42)	(0.98)												(0.98)
$\bar{\gamma}$	0.799	0.106			-0.146								0.283	0.106
	(5.13)	(0.77)			(-3.50)									(0.77)
$\bar{\gamma}$	0.857	0.107			-0.059	0.444							0.303	0.107
	(5.80)	(0.85)			(-1.91)	(5.54)								(0.85)
$\bar{\gamma}$	0.718	0.195			-0.057	0.334	-4.994	1.265	-0.194				0.358	0.195
	(4.53)	(1.40)			(-2.05)	(4.44)	(-6.91)	(5.96)	(-2.13)					(1.40)
$\bar{\gamma}$	0.938	0.244			-0.007	0.280	-6.416	1.194	-0.180	0.062			0.376	0.244
	(5.17)	(1.69)			(-0.30)	(3.93)	(-8.95)	(5.71)	(-2.18)	(2.22)				(1.69)
$\bar{\gamma}$	0.997	0.203			-0.026	0.264	-6.422	1.177	-0.153	0.042	-0.031		0.388	0.203
	(6.21)	(1.63)			(-1.07)	(3.91)	(-8.66)	(5.93)	(-1.90)	(1.77)	(-0.77)			(1.63)
$\bar{\gamma}$	0.795	0.065			-0.024	0.241	-6.529	1.192	-0.178	0.034	-0.026	-0.036	0.395	0.065
	(5.02)	(0.50)			(-1.00)	(3.60)	(-8.78)	(5.41)	(-2.16)	(1.51)	(-0.65)	(-2.32)		(0.50)

Panel B: Beta and beta-squared.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.957	0.166	-0.247										0.221	0.458
	(6.50)	(1.21)	(-3.68)											(2.69)
$\bar{\gamma}$	0.848	0.140	-0.303		-0.156								0.300	0.510
	(5.35)	(1.04)	(-4.29)		(-3.69)									(2.86)
$\bar{\gamma}$	0.906	0.115	-0.246		-0.069	0.413							0.320	0.407
	(5.94)	(0.93)	(-3.49)		(-2.24)	(5.17)								(2.49)
$\bar{\gamma}$	0.744	0.196	-0.207		-0.065	0.306	-5.471	1.233	-0.167				0.371	0.438
	(4.54)	(1.39)	(-3.01)		(-2.34)	(4.16)	(-7.69)	(5.84)	(-1.79)					(2.42)
$\bar{\gamma}$	0.986	0.248	-0.228		-0.016	0.247	-6.953	1.139	-0.153	0.060			0.389	0.522
	(5.01)	(1.66)	(-3.13)		(-0.68)	(3.59)	(-9.88)	(5.47)	(-1.78)	(2.04)				(2.60)
$\bar{\gamma}$	1.044	0.211	-0.236		-0.034	0.229	-6.918	1.135	-0.127	0.041	-0.030		0.399	0.501
	(6.10)	(1.64)	(-3.37)		(-1.42)	(3.50)	(-9.45)	(5.67)	(-1.52)	(1.66)	(-0.78)			(2.75)
$\bar{\gamma}$	0.888	0.118	-0.203		-0.029	0.216	-6.904	1.162	-0.145	0.039	-0.024	-0.024	0.404	0.349
	(5.44)	(0.87)	(-2.77)		(-1.17)	(3.32)	(-9.33)	(5.32)	(-1.69)	(1.63)	(-0.64)	(-1.58)		(1.78)

Panel C: The full house.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.791	-0.084	-0.464	0.683									0.280	0.516
	(5.56)	(-0.73)	(-5.14)	(3.28)										(2.89)
$\bar{\gamma}$	0.809	0.055	-0.371	0.243	-0.119								0.321	0.515
	(4.98)	(0.44)	(-4.83)	(1.87)	(-3.22)									(2.84)
$\bar{\gamma}$	0.858	0.065	-0.296	0.146	-0.052	0.374							0.338	0.418
	(5.45)	(0.54)	(-4.16)	(1.26)	(-1.72)	(5.10)								(2.50)
$\bar{\gamma}$	0.749	0.142	-0.253	0.152	-0.050	0.268	-5.762	1.215	-0.210				0.384	0.448
	(4.54)	(1.03)	(-3.46)	(1.25)	(-1.72)	(3.90)	(-8.15)	(5.72)	(-2.11)					(2.45)
$\bar{\gamma}$	1.014	0.196	-0.263	0.101	-0.012	0.220	-7.168	1.127	-0.195	0.052			0.401	0.514
	(5.21)	(1.37)	(-3.50)	(0.85)	(-0.42)	(3.40)	(-10.30)	(5.36)	(-2.17)	(1.83)				(2.56)
$\bar{\gamma}$	1.092	0.193	-0.263	0.096	-0.026	0.207	-6.913	1.130	-0.172	0.044	-0.042		0.409	0.519
	(6.18)	(1.53)	(-3.57)	(0.89)	(-0.90)	(3.33)	(-9.36)	(5.54)	(-1.98)	(1.74)	(-1.16)			(2.81)
$\bar{\gamma}$	0.934	0.094	-0.235	0.113	-0.020	0.198	-6.894	1.166	-0.190	0.040	-0.037	-0.029	0.414	0.367
	(5.48)	(0.70)	(-3.06)	(1.05)	(-0.65)	(3.18)	(-9.26)	(5.22)	(-2.13)	(1.68)	(-1.00)	(-1.85)		(1.84)

Table IV: Cross-sectional regressions for individual stocks. We compute cross-sectional Fama-MacBeth regressions of monthly excess returns on stock characteristics for individual stocks. We require stocks to have at least 24 valid return observations available in the past 60 months. Regressions are run each month from January 1932 to December 2016. The characteristics include past 60-month market beta ($\hat{\beta}$), centered beta squared ($(\hat{\beta}-\bar{\beta})^2$), standardized residual risk ($\hat{\sigma}'(\epsilon)$), log market capitalization (ME), log book-to-market ratio (BtM), one-month lagged one-month return (R_1), one-month lagged 11-month return (R_{12-2}), 12-month lagged 48-month return (R_{60-13}) (all three past return measures are defined as log cumulative past returns), log illiquidity ($Illiq$), lottery demand (max), and Treasury Bill sensitivity ($\hat{\beta}_{RF}$). The reported coefficients are time-series averages of the monthly regression slopes (gamma). The Newey-West corrected t-statistics of these averages are shown in parentheses. We also report adjusted R-squares (Adj.R^2) for the regressions, and the average beta premium ($ABP = \gamma_1 + \gamma_2\,(1 - 2\,\bar{\beta})$).

Panel A: Beta-only.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.842	0.160											0.031	0.160
	(6.02)	(0.98)												(0.98)
$\bar{\gamma}$	0.418	0.065			-0.115								0.050	0.065
	(2.53)	(0.43)			(-2.64)									(0.43)
$\bar{\gamma}$	0.583	0.091			-0.078	0.270							0.057	0.091
	(3.77)	(0.66)			(-1.98)	(5.16)								(0.66)
$\bar{\gamma}$	0.591	0.057			-0.061	0.212	-8.061	0.878	-0.137				0.088	0.057
	(4.28)	(0.43)			(-2.03)	(4.16)	(-16.64)	(6.05)	(-1.89)					(0.43)
$\bar{\gamma}$	0.492	0.025			-0.102	0.219	-7.982	0.921	-0.119	-0.033			0.091	0.025
	(2.61)	(0.19)			(-3.20)	(4.30)	(-15.89)	(6.50)	(-1.75)	(-1.79)				(0.19)
$\bar{\gamma}$	0.644	0.146			-0.145	0.191	-6.947	0.874	-0.159	-0.037	-0.137		0.096	0.146
	(3.54)	(1.23)			(-5.10)	(3.91)	(-12.40)	(6.12)	(-2.48)	(-2.07)	(-7.40)			(1.23)
$\bar{\gamma}$	0.597	0.110			-0.143	0.190	-6.884	0.906	-0.169	-0.036	-0.137	-0.002	0.097	0.110
	(3.29)	(0.93)			(-5.13)	(3.89)	(-12.43)	(6.16)	(-2.60)	(-2.01)	(-7.54)	(-0.35)		(0.93)

Panel B: Beta and beta-squared.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.873	0.158	-0.176										0.033	0.390
	(6.05)	(0.97)	(-2.48)											(1.89)
$\bar{\gamma}$	0.454	0.061	-0.213		-0.115								0.051	0.345
	(2.77)	(0.41)	(-2.86)		(-2.63)									(1.76)
$\bar{\gamma}$	0.617	0.088	-0.206		-0.078	0.269							0.058	0.364
	(3.99)	(0.64)	(-2.66)		(-1.96)	(5.17)								(1.90)
$\bar{\gamma}$	0.616	0.063	-0.205		-0.062	0.214	-8.071	0.871	-0.131				0.089	0.346
	(4.38)	(0.48)	(-2.62)		(-2.03)	(4.19)	(-16.68)	(5.98)	(-1.83)					(1.93)
$\bar{\gamma}$	0.516	0.034	-0.208		-0.102	0.221	-7.987	0.916	-0.115	-0.033			0.092	0.319
	(2.67)	(0.26)	(-2.55)		(-3.21)	(4.34)	(-15.94)	(6.44)	(-1.70)	(-1.77)				(1.71)
$\bar{\gamma}$	0.671	0.157	-0.221		-0.146	0.193	-6.943	0.869	-0.154	-0.037	-0.138		0.097	0.460
	(3.60)	(1.31)	(-2.71)		(-5.11)	(3.95)	(-12.39)	(6.08)	(-2.43)	(-2.04)	(-7.45)			(2.61)
$\bar{\gamma}$	0.636	0.123	-0.199		-0.144	0.192	-6.885	0.904	-0.166	-0.036	-0.138	-0.001	0.099	0.395
	(3.40)	(1.02)	(-2.42)		(-5.14)	(3.93)	(-12.43)	(6.17)	(-2.57)	(-1.97)	(-7.56)	(-0.22)		(2.23)

Panel C: The full house.

	const	$\hat{\beta}$	$(\hat{\beta}-\bar{\beta})^2$	$\hat{\sigma}'(\epsilon)$	ME	BtM	R_1	R_{12-2}	R_{60-13}	$Illiq$	max	$\hat{\beta}_{RF}$	Adj.R^2	ABP
$\bar{\gamma}$	0.996	0.212	-0.257	0.018									0.051	0.540
	(6.77)	(1.80)	(-3.42)	(0.22)										(3.36)
$\bar{\gamma}$	0.593	0.338	-0.319	-0.237	-0.199								0.060	0.738
	(3.48)	(2.71)	(-4.15)	(-5.05)	(-5.15)									(4.29)
$\bar{\gamma}$	0.712	0.304	-0.300	-0.232	-0.163	0.197							0.066	0.682
	(4.43)	(2.55)	(-3.74)	(-5.06)	(-4.56)	(4.05)								(3.87)
$\bar{\gamma}$	0.734	0.223	-0.269	-0.182	-0.125	0.147	-8.189	0.903	-0.150				0.095	0.573
	(5.06)	(1.91)	(-3.30)	(-3.71)	(-4.76)	(3.09)	(-16.79)	(6.26)	(-2.12)					(3.41)
$\bar{\gamma}$	0.631	0.206	-0.266	-0.205	-0.173	0.155	-8.128	0.956	-0.132	-0.035			0.098	0.551
	(3.23)	(1.77)	(-3.16)	(-4.22)	(-6.38)	(3.21)	(-16.05)	(6.73)	(-1.99)	(-1.90)				(3.15)
$\bar{\gamma}$	0.728	0.248	-0.258	-0.133	-0.187	0.150	-7.161	0.904	-0.167	-0.037	-0.120		0.101	0.589
	(3.86)	(2.20)	(-3.10)	(-2.95)	(-7.00)	(3.18)	(-12.93)	(6.32)	(-2.62)	(-2.05)	(-8.04)			(3.43)
$\bar{\gamma}$	0.686	0.205	-0.231	-0.130	-0.183	0.150	-7.098	0.941	-0.179	-0.036	-0.120	-0.003	0.103	0.510
	(3.62)	(1.80)	(-2.75)	(-2.94)	(-6.93)	(3.19)	(-13.00)	(6.43)	(-2.75)	(-1.97)	(-8.11)	(-0.78)		(2.93)

Table V: Time-series regressions. We perform time-series regressions of portfolio excess returns on factor returns. For each of the ten equal-weighted, size-neutral beta-decile portfolios (BD1 to BD10), we run a single time-series regression over the full sample period from January 1932 to December 2016. Shown are (annualized) estimates of four-factor-model alphas (4f-$\hat{\alpha}$) and factor loadings ($\hat{\beta}_{mkt}$, $\hat{\beta}_{smb}$, $\hat{\beta}_{hml}$ and $\hat{\beta}_{umd}$), as well as three-factor-model and market-model alphas (3f-$\hat{\alpha}$ and 1f-$\hat{\alpha}$). Newey-West corrected t-statistics are presented in parentheses. The table also reports results for four hedge portfolios constructed from the ten decile portfolios: (i) long decile 1 and short decile 10 "BD1-10"), (ii) long decile 1 and short decile 5 ("BD1-5"), (iii) long decile 6 and short decile 10 ("BD6-10"), and (iv) long deciles 5 and 6, and short deciles 1 and 10 ("BD5,6-1,10").

Portolio	4f-$\hat{\alpha}$	$\hat{\beta}_{mkt}$	$\hat{\beta}_{smb}$	$\hat{\beta}_{hml}$	$\hat{\beta}_{umd}$	3f-$\hat{\alpha}$	1f-$\hat{\alpha}$
BD1	2.083	0.560	0.617	0.199	0.022	2.332	3.894
	(2.39)	(25.21)	(6.81)	(4.15)	(0.70)	(2.67)	(3.39)
BD2	2.701	0.692	0.657	0.326	-0.042	2.234	4.486
	(3.77)	(42.65)	(6.93)	(7.30)	(-1.30)	(3.29)	(4.22)
BD3	2.892	0.794	0.635	0.333	-0.066	2.155	4.456
	(4.22)	(54.37)	(12.54)	(10.47)	(-2.32)	(3.25)	(3.95)
BD4	2.044	0.884	0.691	0.353	-0.083	1.117	3.604
	(3.63)	(54.43)	(12.02)	(10.14)	(-3.06)	(1.93)	(3.05)
BD5	2.183	0.953	0.774	0.363	-0.096	1.112	3.767
	(3.89)	(62.93)	(14.58)	(11.42)	(-3.97)	(1.93)	(3.08)
BD6	1.312	1.013	0.814	0.340	-0.138	-0.236	2.460
	(2.31)	(73.60)	(17.44)	(10.86)	(-5.74)	(-0.44)	(1.99)
BD7	0.494	1.106	0.883	0.316	-0.152	-1.204	1.504
	(0.61)	(40.30)	(19.92)	(7.47)	(-4.21)	(-1.76)	(1.10)
BD8	-0.562	1.180	0.939	0.271	-0.201	-2.816	-0.125
	(-0.70)	(64.87)	(25.94)	(7.03)	(-5.78)	(-4.13)	(-0.09)
BD9	-1.916	1.257	1.009	0.231	-0.251	-4.729	-2.018
	(-1.69)	(55.15)	(16.42)	(3.60)	(-4.83)	(-4.90)	(-1.29)
BD10	-3.222	1.382	1.177	0.049	-0.329	-6.908	-4.579
	(-1.84)	(45.93)	(13.83)	(0.47)	(-3.48)	(-5.09)	(-2.35)
BD1-10	5.305	-0.821	-0.560	0.149	0.352	9.240	8.473
	(2.52)	(-19.71)	(-3.43)	(1.15)	(3.28)	(5.12)	(4.31)
BD1-5	-0.099	-0.393	-0.157	-0.164	0.118	1.220	0.128
	(-0.12)	(-16.09)	(-2.37)	(-4.00)	(4.29)	(1.48)	(0.15)
BD6-10	4.533	-0.369	-0.363	0.291	0.191	6.672	7.038
	(2.84)	(-12.21)	(-3.19)	(2.62)	(2.01)	(4.99)	(4.74)
BD5,6-1,10	4.633	0.024	-0.206	0.455	0.073	5.452	6.910
	(2.84)	(0.66)	(-1.83)	(3.89)	(0.74)	(3.85)	(4.30)

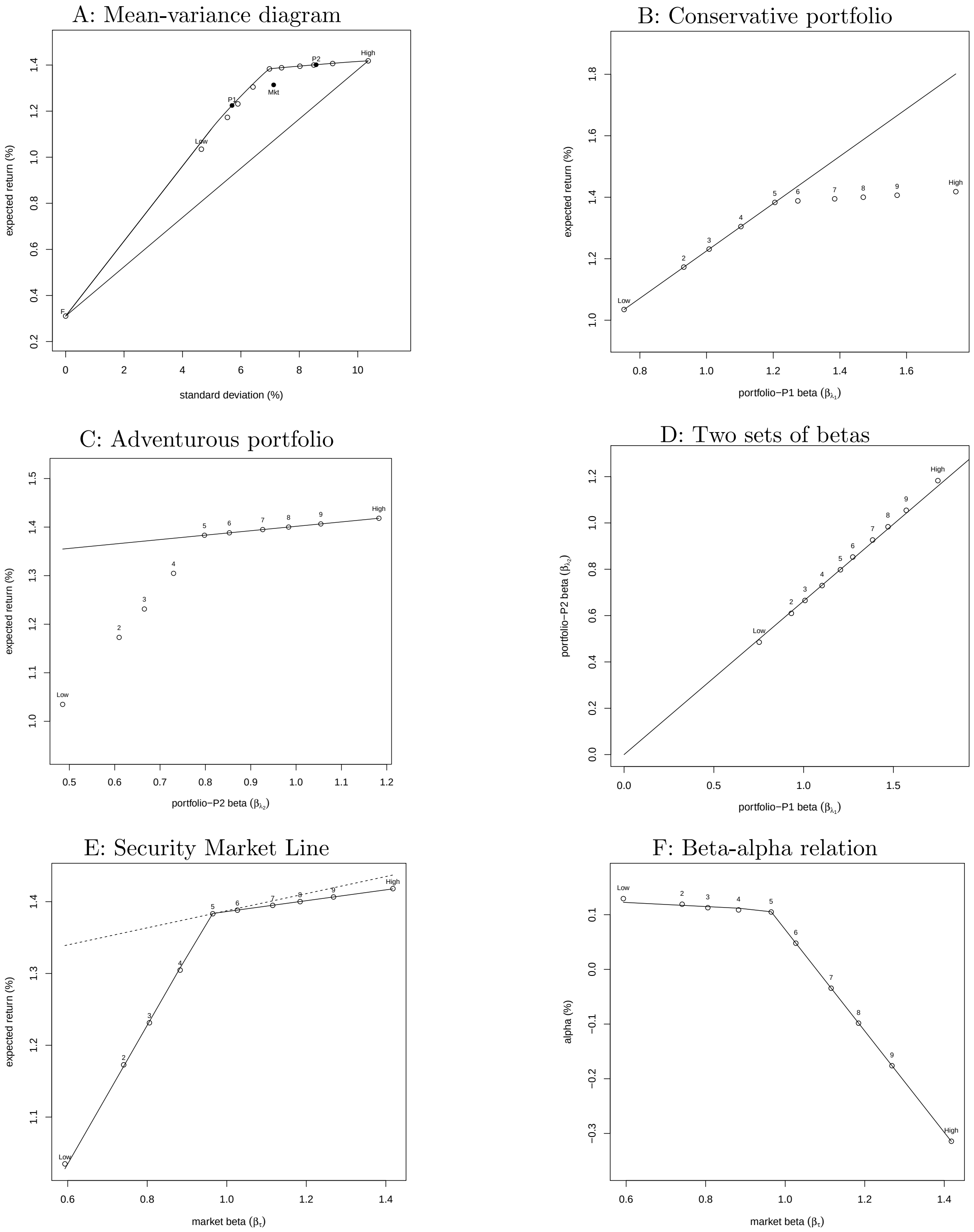

Figure 1: Numerical example. The figure illustrates the 11-asset, two-investor equilibrium described in Table II. Panel A shows the full portfolio possibilities set (or all convex combinations of the 11 base assets) and adds the optimal portfolios (P_1 and P_2) of the two investors, together with the market portfolio (Mkt). Panel B plots the expected returns against the betas relative to the conservative investor's portfolio (P_1). The solid line represents the portfolio optimality condition; the line connects the conservative investor's active assets and envelops her inactive assets. Panel C shows a similar plot for the adventurous investor and her portfolio P_2. Panel D shows that the betas relative to P_1 are nearly proportional to the betas relative to P_2, reflecting the very high correlation between the two portfolios. Panel E shows the relation between expected return and market beta. The dashed, straight line represents represents the classical SML, or average of the two lines shown in Panel B and C. The solid, kinked line represents our CSML approximation (15), which combines the two sets of optimality conditions under the assumption that P_1 and P_2 have zero non-market factor loadings. Panel F shows the relation between market beta and market alpha. The solid, kinked line represents again the CSML approximation of this relation.

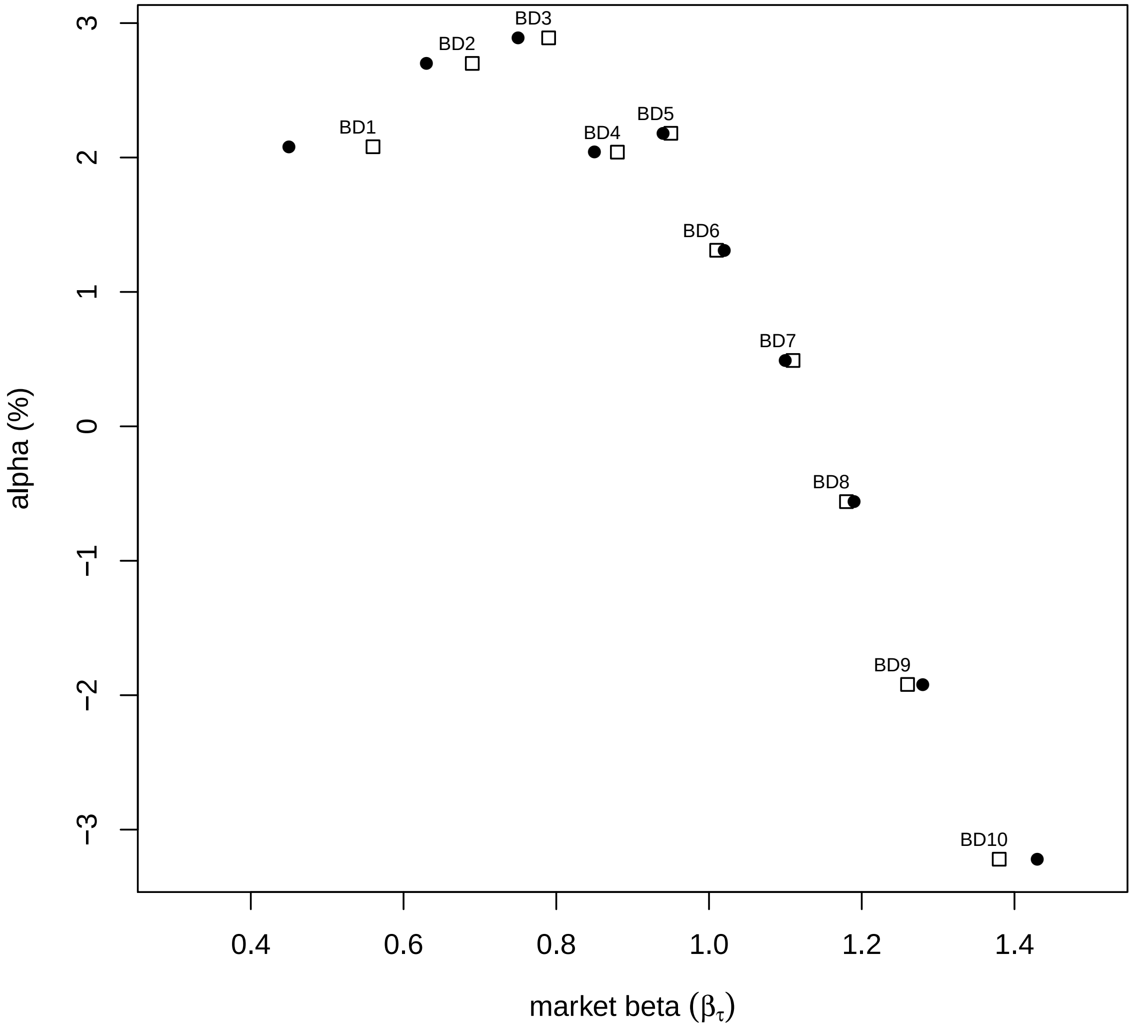

Figure 2: Alpha versus beta. The figure plots the estimated four-factor alphas versus estimated market betas (clear squares) for our ten size-neutral beta stock portfolios. Time-series regressions for each of the ten portfolios are run once from January 1932 to December 2016 using ex-post equal-weighted monthly returns. Also shown are the time series averages of the 60-month Bayesian shrinkage estimates of the market betas (filled circles).