Why Does Junior Put All His Eggs In One Basket? A Potential Rational Explanation for Holding Concentrated Portfolios
On-line Appendix

Appendix A. Proof of Proposition 1

Assume that $(c,(x,z))$ is a feasible strategy for initial conditions (W_t, Y_t). Then, for all $\alpha > 0$, we first show that $(\alpha c,(\alpha x, \alpha z))$ is a feasible strategy for initial conditions $(\alpha W_t, \alpha Y_t)$. Consider the dynamics for the wealth process W_α, with initial conditions $(\alpha W_t, \alpha Y_t)$, following the consumption and investment plan $(\alpha c,(\alpha x, \alpha z))$. We have

$$dW_{\alpha s} = \alpha W_s ds - \alpha c_s ds + \alpha Y_s ds + \alpha z_s^\intercal(\mu - r\bar{1})ds + \alpha z_s^\intercal \sigma dw_s = \alpha dW_s, \tag{1}$$

therefore, $W_{\alpha s} = \alpha W_s$. Similarly, we have $Y_{\alpha s} = \alpha Y_s$. The investment strategy satisfies the margin requirement since

$$\lambda^\intercal(\alpha z) = \alpha \lambda^\intercal z \le \alpha W. \tag{2}$$

It follows that

$$F(\alpha W, \alpha Y) \le \alpha^{1-\gamma} F(W, Y), \tag{3}$$

since the utility function in homogeneous of degree $1 - \gamma$. In addition

$$F(W, Y) = F(\alpha^{-1}\alpha W, \alpha^{-1}\alpha Y) \le \alpha^{\gamma-1} F(\alpha W, \alpha Y), \tag{4}$$

so given Eq. (3) in fact we have

$$F(\alpha W, \alpha Y) = \alpha^{1-\gamma} F(W, Y). \tag{5}$$

Appendix B. Proof of Proposition 2

To show that F is nondecreasing in (W_t, Y_t) is simple, since given an initial endowment (W_t, Y_t), it is clear that starting with wealth $W_t' > W_t$ or income $Y_t' > Y_t$ at time t, the optimal strategy for the initial condition (W_t, Y_t) is still admissible and potentially nonoptimal for the problem with initial conditions (W_t', Y_t'). This implies that F is nondecreasing in W and Y. To show concavity, consider two initial conditions (W_t, Y_t) and (W_t', Y_t') and $\alpha \in (0, 1)$. Denote $(c, (x, z))$ and $(c', (x', z'))$ the optimal strategies respectively for the two initial conditions. Then, the strategy

$$S : (\alpha c + (1-\alpha)c', \alpha x + (1-\alpha)x', \alpha z + (1-\alpha)z'), \tag{6}$$

is admissible for the initial condition

$$I : (\alpha W_t + (1-\alpha)W_t', \alpha Y_t + (1-\alpha)Y_t'). \tag{7}$$

Denoting W^α the wealth process associated with strategy S and initial condition I, for all times s, we have

$$W_s^\alpha = \alpha W_s + (1-\alpha)W_s', \tag{8}$$

and similarly for the income process

$$Y_s^\alpha = \alpha Y_s + (1-\alpha)Y_s'. \tag{9}$$

The margin constraint is satisfied since

$$\lambda^\intercal(\alpha z + (1-\alpha)z') = \alpha\lambda^\intercal z + (1-\alpha)\lambda^\intercal z' \leq \alpha W + (1-\alpha)W \leq W, \tag{10}$$

as both z and z' are feasible. Finally, by strict concavity of the utility function u, we have

$$E_t\left[\int_t^{\infty} u\left(\alpha c_s + (1-\alpha)c'_s\right) e^{-\theta s} ds\right] > E_t\left[\int_t^{\infty} \left(\alpha u(c_s) + (1-\alpha)u(c'_s)\right) e^{-\theta s} ds\right], \quad (11)$$

which implies that

$$F(\alpha W_t + (1-\alpha)W'_t, \alpha Y_t + (1-\alpha)Y'_t) > \alpha F(W_t, Y_t) + (1-\alpha)F(W'_t, Y'_t). \quad (12)$$

Appendix C. Proof of Proposition 3

We note that the assumption that $(\sigma\sigma^\intercal)^{-1}\eta \in \mathbb{R}_+^N$ ensures that all assets are held long in the portfolio when the margin requirement is not binding.

The assumption that all the entries off the diagonal of the inverse covariance matrix $(\sigma\sigma^\intercal)^{-1}$ are non-positive implies that $(\sigma\sigma^\intercal)^{-1} = \alpha I_N - P$, where $\alpha > 0$ and P is a matrix with non-negative elements. Since $(\sigma\sigma^\intercal)^{-1}$ is positive definite, all its eigenvalues are positive, which implies that the spectral radius of matrix P/α must be strictly less than one. In spectral theory, this class of matrices is called Z-matrices (or negated Metzler matrices). Note that we have

$$\sigma\sigma^\intercal = \frac{1}{\alpha}(I_N - \frac{P}{\alpha})^{-1} = \frac{1}{\alpha}\sum_{n=0}^{\infty}\left(\frac{P}{\alpha}\right)^n < \infty \quad (13)$$

as the spectral radius of P/α, is strictly less than one. We conclude that all the entries of the covariance matrix $\sigma\sigma^\intercal$ are non-negative, i.e. all the assets are pairwise positively correlated. The assumption is satisfied for instance when (i) the returns of all the N assets are independent, or (ii) when the returns of all the assets have pairwise the same non-negative coefficient of correlation $\rho \geq 0$.

To see this last point, consider the case where all pairwise correlations are positive and equal to $\rho > 0$. Let $M = (\sigma\sigma^\intercal)^{-1} = [m_{ij}]$. It is easy to check that

$$\begin{aligned} m_{ii} &= \frac{1 + (N-2)\rho}{(1-\rho)(1+(N-1)\rho)} \frac{1}{\sigma_i^2} > 0 \\ m_{ij} &= -\frac{\rho}{(1-\rho)(1+(N-1)\rho)} \frac{1}{\sigma_i \sigma_j} < 0, \; i \neq j. \end{aligned} \tag{14}$$

We proceed with the proof of Proposition 3 in three steps.

Step 1: Normalization of the Program.

First, we rewrite the optimization problem. When labor income is uncorrelated with the market, we have

$$\max_{\omega \in \mathbb{R}_+^N} \quad \omega^\intercal \eta - \frac{y}{2} \omega^\intercal (\sigma\sigma^\intercal) \omega, \qquad \text{s.t. } \omega^\intercal \lambda^+ \leq 1, \tag{15}$$

where $\eta = \mu - r\overline{1}$. For $k = 1, \ldots, N$, set $\widehat{\omega}_i = \omega_i / \lambda_i^+$, $\widehat{\eta}_i = \eta_i / \lambda_i^+$ and $\widehat{\sigma}_i = \sigma_i / \lambda_i^+$. The optimization program is equivalent to

$$\max_{\widehat{\omega} \in \mathbb{R}_+^N} \quad \widehat{\omega}^\intercal \widehat{\eta} - \frac{y}{2} (\widehat{\omega})^\intercal (\widehat{\sigma}\widehat{\sigma}^\intercal) \widehat{\omega}, \qquad \text{s.t. } \widehat{\omega}^\intercal \overline{1} \leq 1. \tag{16}$$

Observe that $(\sigma\sigma^\intercal)^{-1}\eta \in \mathbb{R}_+^N$ if and only if $(\widehat{\sigma}\widehat{\sigma}^\intercal)^{-1}\widehat{\eta} \in \mathbb{R}_+^N$. Thus, without loss of generality, we can assume that the margin coefficients are the same for all the assets and can be normalized to one.

Step 2: Reduced Effective Domain.

In Lemma H.1 in Appendix H, we discuss the dual formulation of the optimization problem, define the effective domain $\mathcal{N}_{a,b}$ and show that

$$\mathcal{N}_{a,b} = \{(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+^N, (1 - \lambda_i^+) a \leq b_i \leq (1 + \lambda_i^-) a\}. \tag{17}$$

When short sales are prohibited and margin coefficients are normalized to one, the effective domain is $\mathcal{N}_{a,b}$ reduces to $\mathbb{R}_+ \times \mathbb{R}_+^N$. The corresponding dual optimization program is

$$\min_{(a,b) \in \mathbb{R}_+ \times \mathbb{R}_+^N} ay + \frac{1}{2} (\eta + b - a\overline{1})^\intercal (\sigma\sigma^\intercal)^{-1} (\eta + b - a\overline{1}). \tag{18}$$

We note that asset i is not included in the portfolio if and only if $b_i^* > 0$, $i \in \{1, \ldots, N\}$.

Step 3: Supermodularity Property.

As shown in Appendix I, the optimal control variable a^* is a non-increasing function of the lifetime relative risk aversion, y. For $y \geq y_B^*$, we know that $b^* \equiv 0$. Next, assume that $y < y_B^*$ so that $a^*(y) > 0$ and observe that

$$\frac{1}{2}(\eta+b-a^*(y)\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\eta+b-a^*(y)\overline{1}) = \frac{(a^*(y))^2}{2}(\widehat{\eta}^*(y)+b'-\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\widehat{\eta}^*(y)+b'-\overline{1}), \tag{19}$$

where $b' = b/a^*(y) \in \mathbb{R}_+^N$ and $\widehat{\eta}^*(y) = \eta/a^*(y)$. The dual optimization problem can be seen as an N player game, where player i chooses quantity $b_i' \in \mathbb{R}_+$ in order to maximize profit π_i where

$$\pi_i(b_i', b_i'^{-1}; y) = -\frac{1}{2}(\widehat{\eta}^*(y) + b' - \overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\widehat{\eta}^*(y) + b' - \overline{1}). \tag{20}$$

For all $(i, j) \in \{1, \ldots, N\}^2$, $i \neq j$, we have

$$\begin{aligned} \frac{\partial^2 \pi_i}{\partial b_i' \partial b_j'} &= -e_i^\intercal(\sigma\sigma^\intercal)^{-1}e_j \geq 0 \\ \frac{\partial^2 \pi_i}{\partial b_i \partial y} &= \frac{\partial a^*(y)}{\partial y}\frac{1}{(a^*(y))^2}e_i^\intercal(\sigma\sigma^\intercal)^{-1}\eta \leq 0. \end{aligned} \tag{21}$$

From Eq. (21), we have that this game satisfies the supermodularity conditions, so its unique Nash equilibrium $(b_1^*, b_2^*, \ldots, b_N^*)$ is non-increasing in the lifetime relative risk aversion, y, see Topkis (1998) and lecture notes by Levin (2006). This implies that, should asset i not be the asset with the largest expected excess return, if $y_{D,i}^* = \sup \{y \geq 0,\ b_i^*(y) = 0\}$, then for all $y > y_{D,i}^*$, $b_i^*(y) = 0$, i.e., asset i is optimally held in the portfolio as long as $y > y_{D,i}^*$, but is not included in the portfolio for all $y \leq y_{D,i}^*$. As in the case of assets with independent returns, there are $N + 1$ regions.

Appendix D. Proof of Proposition 4

For $y > y_B^*$, optimal allocation in asset k is given by

$$\omega_k^* = \frac{\mu_k - r}{y\sigma_k^2}. \tag{22}$$

For y slightly below y_B^*, we have

$$\omega_k^* = \frac{1}{y\sigma_k^2}\left(\mu_k - r - \psi_N(\lambda^*, y)\right), \tag{23}$$

and the Lagrange multiplier $\psi_N(\lambda^*, y)$ is equal to

$$\psi_N(\lambda^*, y) = \frac{\alpha^\intercal \xi - y}{\alpha^\intercal \overline{1}}, \tag{24}$$

where

$$\xi_k = \frac{\mu_k - r}{\lambda_k^*}, \tag{25}$$

and

$$\alpha_k = \left(\frac{\lambda_k^*}{\sigma_k}\right)^2, k \in \{1, \ldots, N\}. \tag{26}$$

Since for all $k \in \{1, \ldots, N\}$, we must have

$$\frac{\omega_k^*}{\lambda_k^*} \geq 0, \tag{27}$$

it implies that λ_k^* and $\mu_k - r$ must have the same sign, so that $\xi_k > 0$. It is possible to rewrite the optimal asset allocation as

$$\omega_k^* = \frac{\alpha_k}{y\lambda_k^*\alpha^\intercal\overline{1}}(y - \alpha^\intercal(\xi - \xi_k\overline{1})). \tag{28}$$

At $y = y_B^*$, we must have

$$\psi_N(\lambda^*, y_B^*) = 0, \tag{29}$$

which leads to

$$y_B^* = \alpha^\intercal \xi. \tag{30}$$

Next, without loss of generality, assume that $0 < \xi_N < \xi_{N-1} < \cdots < \xi_1$. Since $\xi \geq 0$ and $\alpha \geq 0$, it is easy to see that as y decreases, asset allocation ω_N^* is the first allocation to hit zero at

$$y_{N,D}^* = \alpha^\intercal(\xi - \xi_N \overline{1}). \tag{31}$$

More generally, for $K \in \{1, \ldots, N\}$ define the dropping cutoff

$$y_{K,D}^* = (I_K\alpha)^\intercal \left[I_K(\xi - \xi_K\overline{1})\right], \tag{32}$$

and by convention, set $y_{N+1,D}^* = y_B^*$; observe that $0 = y_{1,D}^* < y_{2,D}^* < \cdots < y_{N+1,D}^*$. When K assets are held in the portfolio, optimal allocation in asset k is given by

$$\omega_k^* = \frac{\mu_k - r - \psi_K(\lambda, y)\lambda_k}{y\sigma_k^2}. \tag{33}$$

It is easy to see that for $\frac{\omega_k^*}{\lambda_k}$ to be positive, we must have $\frac{\mu_k - r}{\lambda_k}$ positive as $\psi_K(\lambda, y) > 0$. This implies that the vector of margin coefficients must be the same for all $y \leq y_B^*$, i.e. $\lambda = \lambda^*$. Without loss of generality we can assume that the excess return of every asset is positive, and, by Proposition C, there are exactly $N+1$ regions: for $y_{K,D}^* < y < y_{K+1,D}^*$, only the first K assets are held in the portfolio, $K \in \{1, \ldots, N\}$ with

$$\omega_k^* = \frac{\alpha_k}{y\lambda_k^*} \frac{\left[y - y_{K,D}^*\right]^+}{(I_K\alpha)^\intercal\overline{1}}, \;\; k = 1, \ldots, N. \tag{34}$$

Appendix E. Proof of Proposition 5

The margin constraint is equivalent to 2^N linear constraints of the form $\lambda^\intercal z \leq W$, where $\lambda \in \Lambda$. Each linear constraint is defined by its vector λ. Note that *at most* N constraints can be binding at the same time. If exactly 2 constraints are binding, constraints p and q respectively defined by vectors $\lambda^{(p)}$ and $\lambda^{(q)}$, are binding, it must be the case that vectors $\lambda^{(p)}$ and $\lambda^{(q)}$ have $N-1$ components in common; if the kth component $\lambda_k^p \neq \lambda_k^q$, then $z_k^* = 0$, i.e. asset k is dropped out of the portfolio. More generally, if exactly $K+1$ constraints are binding, K assets have been dropped out of the portfolio and, the vectors $\left\{\lambda^{(i)}\right\}_{i=1}^{K+1}$ of the binding constraints must have $N-K$

components in common. The Hamilton-Jacobi-Bellman (HJB) equation for the primal value function F is

$$\theta F = \max_{\frac{z}{W}\in Q} \frac{\gamma (F_1)^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + (rW+Y)F_1 + mYF_2 + \frac{\Sigma^\intercal\Sigma + \Theta^\intercal\Theta}{2}Y^2F_{22} \\ + z^\intercal\left((\mu - r\bar{1})F_1 + \sigma\Sigma Y F_{12}\right) + \frac{z^\intercal\sigma\sigma^\intercal z}{2}F_{11}. \tag{35}$$

Since $F(W,Y) = Y^{1-\gamma}f(\frac{W}{Y})$, the maximization program is equivalent to

$$\max_{\omega\in Q} \quad \omega^\intercal(\eta + y\sigma\Sigma) - \tfrac{y}{2}\omega^\intercal\sigma\sigma^\intercal\omega, \tag{36}$$

with $\omega = z/W$ and lifetime relative risk aversion

$$y = -\frac{WF_{11}}{F_1} = -\frac{vf''(v)}{f'(v)}, \tag{37}$$

the program defined in Eq. (36) is well defined, since, for $y > 0$, the objective function is strictly concave and the margin constraint is convex, so there is a unique solution that, from the maximum theorem, is continuous in y.

Case $\eta = 0$.

In this case, the program defined in Eq. (36) is independent of the parameter y, so the fraction of wealth invested in each asset is constant. The unconstrained allocation is

$$\frac{z}{W} = (\sigma\sigma^\intercal)^{-1}\sigma\Sigma. \tag{38}$$

If

$$\max_{\lambda\in\Lambda} \left(\lambda^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma\right) \leq 1, \tag{39}$$

the margin constraint is never binding, so

$$\frac{z^*}{W} = (\sigma\sigma^\intercal)^{-1}\sigma\Sigma. \tag{40}$$

If, on the other hand,

$$\max_{\lambda\in\Lambda} \left(\lambda^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma\right) > 1, \tag{41}$$

the constraint is always binding. Depending of the parameters values, K assets are optimally held in the portfolio, with $K = 1, \dots, N$. More specifically, assuming that assets $N, N-1, \dots, K+1$ are dropped from the portfolio, K assets remain if and only if for exactly K assets

$$\max_{\lambda \in \Lambda} (\lambda_k e_k^\intercal I_K \omega^*) > 0, \quad k = 1, \dots, K, \tag{42}$$

with

$$I_K \omega^* = \frac{(I_K \sigma \sigma^\intercal I_K^\intercal)^{-1} I_K \sigma \Sigma + (1 - \lambda^\intercal I_K^\intercal (I_K \sigma \sigma^\intercal I_K^\intercal)^{-1} I_K \sigma \Sigma) I_K \lambda}{\lambda^\intercal I_K^\intercal (I_K \sigma \sigma^\intercal I_K^\intercal)^{-1} I_K \lambda} \tag{43}$$

and $z_k^* = 0$ for $k = K+1, \dots, N$. The proof is the same as in the case $\eta \neq 0$ (see below) and is therefore omitted.

Case $\eta \neq 0$.

Since we intend to achieve a maximum, the smaller the number of constraints that are binding, the higher the maximum value. First we look at the values of y such that the margin constraint is not binding.

Nonbinding region. The first order condition leads to

$$\omega^* = \frac{(\sigma \sigma^\intercal)^{-1}}{y} (\eta + y \sigma \Sigma). \tag{44}$$

To satisfy the margin constraint, we must have

$$\max_{\lambda \in \Lambda} (\omega^*)^\intercal \lambda \leq 1. \tag{45}$$

First, we characterize the binding cutoff y_B^*. As long as the constraint is not binding, the optimal asset allocation is given by Eq. (44). Define

$$y_B^* = \max_{\lambda \in \Lambda} \frac{\lambda^\top (\sigma \sigma^\top)^{-1} \eta}{1 - \lambda^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma}. \tag{46}$$

Since Λ is discrete and finite, the maximum is attained for some $\lambda = \lambda_B^*$; by construction, we have

$$(\lambda_B^*)^\top \frac{(\sigma \sigma^\top)^{-1}}{y_B^*} (\eta + y_B^* \sigma \Sigma) = 1, \tag{47}$$

so the constraint is binding at $y = y_B^*$. Using Eq. (44), and the condition on the matrix J_K in Eq. (21) of the paper, it is easy to see that for $y > y_B^*$, $\max_{\lambda \in \Lambda} \lambda^\top \left(\frac{z^*}{W}\right) < 1$, so the constraint is not binding. Finally, for the constraint to be binding at $y = y_B^*$, it is easy to verify that vector λ_B^* (at $y = y_B^*$), must be such that the sign of $\lambda_{B,i}^*$ and ω_i^* given by Eq. (44) is the same for all $i = 1, \ldots, N$.

Case $\Theta = 0$. Using Eq. (44), we obtain the following reduced HJB equation

$$\left(\theta + (\gamma - 1)(m - \gamma\frac{\Sigma^\intercal\Sigma}{2})\right) f(v) = \frac{\gamma}{1-\gamma}\left(f'(v)\right)^{\frac{\gamma-1}{\gamma}} + f'(v) + B^{-1} v f'(v) - \frac{1}{2}\eta^\intercal(\sigma\sigma^\intercal)^{-1}\eta \frac{\left(f'(v)\right)^2}{f''(v)}. \tag{48}$$

Consider the Legendre transform: $x = f'(v)$, $v = -J'(x)$ and $f(v) = J(x) - xJ'(x)$. It follows that function J must solve the following linear ODE

$$\begin{aligned}\left(\theta + (\gamma - 1)(m - \gamma\frac{\Sigma^\intercal\Sigma}{2})\right) J(x) = &\frac{\gamma}{1-\gamma} x^{\frac{\gamma-1}{\gamma}} + x \\ &+ (\theta - B^{-1} + (\gamma-1)(m - \gamma\frac{\Sigma^\intercal\Sigma}{2}) x J'(x) \\ &+ \frac{1}{2}\eta^\intercal(\sigma\sigma^\intercal)^{-1}\eta x^2 J''(x).\end{aligned} \tag{49}$$

The general solution is

$$J(x) = \frac{\gamma A x^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + Bx + \frac{\gamma K}{\beta - 1 + \gamma} x^{\frac{\beta-1+\gamma}{\gamma}} + \frac{\gamma L}{\delta - 1 + \gamma} x^{\frac{\delta-1+\gamma}{\gamma}}, \tag{50}$$

where K and L are constants and β and δ are respectively the positive and negative root of the quadratic

$$\frac{1}{2\gamma^2}\left(\eta^\intercal(\sigma\sigma^\intercal)^{-1}\eta\right) x^2 + \left(A^{-1} - B^{-1} - \frac{1}{2\gamma^2}\eta^\intercal(\sigma\sigma^\intercal)^{-1}\eta\right) x = A^{-1}. \tag{51}$$

We note that if x is a root of the quadratic

$$\left(\theta + (\gamma - 1)(m - \gamma\frac{\Sigma^\intercal\Sigma}{2})\right) = (\theta - B^{-1} + (\gamma-1)(m - \gamma\frac{\Sigma^\intercal\Sigma}{2}) x + \frac{1}{2}\eta^\top(\sigma\sigma^\intercal)^{-1}\eta x^2, \tag{52}$$

then $z = \gamma(x-1)+1$ is a root of the quadratic

$$\frac{1}{2}\left(\eta^\top(\sigma\sigma^\intercal)^{-1}\eta\right)x^2 + \left(A^{-1} - B^{-1} - \frac{1}{2}\eta^\top(\sigma\sigma^\intercal)^{-1}\eta\right)x = A^{-1}. \tag{53}$$

Differentiating Eq. (50) with respect to x and using the fact that $x = f'(v)$ and $v = -J'(x)$ leads to

$$v + B = Af'(v)^{-\frac{1}{\gamma}} + Kf'(v)^{\frac{\beta-1}{\gamma}} + Lf'(v)^{\frac{\delta-1}{\gamma}}. \tag{54}$$

Then, when v is large, the margin constraint is irrelevant: asymptotically, the solution $f'(v)$ must be the same as in the unconstrained case, so $f'(v)^{-\frac{1}{\gamma}} \underset{\infty}{\sim} A^{-1}v$. Since $\delta - 1 < 0$, we must have $L = 0$. Finally, K must be positive, otherwise for all v in the nonbinding region we have $f'(v) < f_0'(v)$, where f_0 is the unconstrained, reduced, value function. Integrating this relationship from v to $M > v$, we find that

$$f_0(v) < f(v) + f_0(M) - f(M). \tag{55}$$

Since in the limit when wealth goes to infinity, constrained and unconstrained value functions coincide, for any given v the previous relationship implies that $f_0(v) < f(v)$, which is impossible.

Binding region. We now assume that $y \leq y_B^*$. The Lagrangian for the maximization problem is

$$L = \omega^\intercal(\eta + y\sigma\Sigma) - \frac{1}{2}y\omega^\intercal\sigma\sigma^\intercal\omega - \psi(\omega^\intercal\lambda - 1), \tag{56}$$

where $\psi \geq 0$ is the Lagrange multiplier associated with the constraint. Let $\psi_K(\lambda, y)$ denote the value of the Lagrange multiplier ψ when only the first K assets are held in the portfolio for some level of risk aversion y and vector of margin coefficients $\lambda \in \Lambda$. The first order condition leads to

$$\omega^* = \frac{(\sigma\sigma^\intercal)^{-1}}{y}(\eta + y\sigma\Sigma - \psi_N(\lambda_B^*, y)\lambda_B^*). \tag{57}$$

Since the margin constraint is binding, $(\lambda_B^*)^\intercal\omega^* = 1$, we obtain that

$$\psi_N(\lambda_B^*, y) = \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\eta - (1 - (\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma)y}{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\lambda_B^*}. \tag{58}$$

This derivation is valid as long as for all $i = 1, \ldots, N$, $\omega_i^*/\lambda_{B,i}^* \geq 0$. At $y = y_B^*$, $\psi_N = 0$, exactly one constraint is binding and all asset allocations are different from zero until y becomes too small. More precisely, from Eqs. (57) and (58), it is easy to verify that $z_i^* = 0$ exactly when $y = y_{i,N}$ with

$$y_{i,N} = \frac{\left(\lambda_B^* - \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\lambda_B^*}{e_i^\intercal(\sigma\sigma^\intercal)^{-1}\lambda_B^*}e_i\right)^\intercal(\sigma\sigma^\intercal)^{-1}\eta}{1 - \left(\lambda_B^* - \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\lambda_B^*}{e_i^\intercal(\sigma\sigma^\intercal)^{-1}\lambda_B^*}e_i\right)^\top(\sigma\sigma^\intercal)^{-1}\sigma\Sigma}. \tag{59}$$

We can assume that $y_{N,N} = \max_{i=1,..,N}\{y_{i,N}\}$ and $y_{N,N} > 0$. When $y = y_{N,N}$, $z_N^* = 0$ and a second linear constraint becomes binding. Hence, we can conclude that for $y_{N,N} < y < y_B^*$, the margin constraint is binding and all assets are optimally held in the portfolio. For y slightly below $y_{N,N}$, at least two linear constraints are binding and allocation in asset N must be zero for y on some interval $[y_{N,N} - \varepsilon, y_{N,N}]$ for some $\varepsilon > 0$. To see this, we proceed by contradiction and assume that the position of asset N changes sign at $y = y_{N,N}$. We denote $\overline{\lambda}_B^*$ the vector of margin coefficients that has the same components as vector λ_B^*, except the last one. Since Λ is a discrete set, at $y = y_{N,N}$ we must have $\psi_N(\lambda_B^*, y_{N,N}) \neq \psi_N(\overline{\lambda}_B^*, y_{N,N})$, which is impossible by the continuity of the solution in the lifetime relative risk aversion y. As mentioned earlier, the vectors λ of these two linear constraints have their $N-1$ first components in common and only their last components differ. It follows that as risk aversion y decreases, the optimization problem is identical to program defined in Eq. (36) but possibly of smaller dimension (not holding some assets may be optimal) and for a different vector of margin coefficients $\lambda \in \Lambda$. Next, Lemma E.1 characterizes the optimal asset allocation when it is optimal only hold K assets in the portfolio. To simply the exposition, we assume, without loss of generality, that the first K assets are held in the portfolio while keeping in mind that several different K asset configurations can take place as y decreases. Finally, it should be clear from the previous N asset analysis that in general (except for a parameter degeneracy), it is optimal to hold K assets as long as y belongs to a nonempty interior interval.

Lemma E.1. *Assume that for* y *in* $\left[y^-_{N,K}, y^+_{N,K}\right]$ *with* $0 < y^-_{N,K} < y^+_{N,K} \leq y^*_B$*, the first* K *assets are optimally held in nonzero positions. Then, for all* $y \in \left[y^-_{N,K}, y^+_{N,K}\right]$*, there is a vector* $\lambda \in \Lambda$ *such that the optimal asset allocation is given by*

$$I_K\omega^* = \frac{I_K^\top (I_K\sigma\sigma^\intercal I_K^\top)^{-1} I_K \left(\eta - \psi_K(\lambda, y)\lambda\right)}{y}, \tag{60}$$

and satisfies that
(i) for all $i \in \{1, \ldots, K\}$*,* $\omega^*_i/\lambda_i \geq 0$
(ii) the Lagrange multiplier ψ_K *associated with the optimization problem is positive and given by*

$$\psi_K(\lambda, y) = \frac{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\top)^{-1} I_K\eta - \left(1 - (I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\top)^{-1} I_K\sigma\Sigma\right) y}{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\top)^{-1} I_K\lambda}. \tag{61}$$

(iii) Risky asset allocations are given by

$$I_K\omega^* = \frac{(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\eta}{y} + \left(M_K - \frac{L_K}{y}\right)(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\lambda, \tag{62}$$

with

$$\begin{aligned} L_K &= \frac{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\eta}{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\lambda} \\ M_K &= \frac{1}{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\lambda}. \end{aligned} \tag{63}$$

For y *in* $\left[y^-_{N,K}, y^+_{N,K}\right]$ *, no other asset configuration of dimension larger than* K *satisfies all the aforementioned properties.*

Proof of Lemma E.1: By assumption for all $\lambda \in \Lambda$, $(1 - \lambda^\intercal I_K^\top (I_K\sigma\sigma^\intercal I_K^\top)^{-1} I_K\sigma\Sigma) > 0$, so as y decreases, ψ_K remains positive. Consider the optimization problem $P_{N,J}$ when investors face a margin constraint, N assets are available but the last $N - J$ assets must be held in zero positions. Clearly, program $P_{N,J}$ is more stringent than program $P_{N,J+1}$, and for $(y, \lambda) \in \mathbb{R}_+ \times \Lambda$ given, the optimal solution of problem $P_{N,J}$ is an admissible (not necessarily optimal) allocation for problem $P_{N,J+1}$. Then, assume that for $(y, \lambda) \in \mathbb{R}_+ \times \Lambda$, there is a solution to program $P_{N,J+1}$ that is given by Eq. (60) where the Lagrange multiplier is given by Eq. (61) for $K = J + 1$. Given what precedes, it cannot be the case that the optimal solution to program $P_{N,J+1}$ and, a fortiori the optimal solution to program $P_{N,N}$, is given by Eq. (60) for $K = J$ unless asset $K + 1$ is held in a zero position. Whenever the asset position is given by Eq. (60), the investor is better off holding more rather than less assets.

Margin constraint binding for all $y \leq y_B^*$. The result follows from the fact that, if $K \leq N$ assets are optimally held in the portfolio as lifetime relative risk aversion y decreases, then Lagrange multiplier ψ_K given by Eq. (61) remains positive and therefore the constraint must be binding. This implies that once the constraint starts binding at $y = y_B^*$, it remains binding for all $y \leq y_B^*$.

A given optimal asset position may only be found as long as y belongs to a single interval. Assume that for y in $\left[y_{N,K}^-, y_{N,K}^+\right]$ with $0 < y_{N,K}^- < y_{N,K}^+ \leq y_B^*$, it is optimal to hold only K assets (without loss of generality the first K assets) in nonzero positions with vector of margin coefficient $I_K^\intercal \lambda$ and assume that $y_{N,K}^+$ is the largest value of y such that it is optimal to hold the (specific) asset combination. Lagrange multiplier $\psi_K(\lambda, y)$ given by Eq. (61) is a linear function that decreases with the second argument y, which implies that the components of the vector $yI_K\omega^*$ is also a linear function of y, where $I_K\omega^*$ is given by Eq. (60). Next, note that the components of vector $I_K\omega^*$ has a constant sign (the same sign as the component of vector $I_K\lambda$) on an interval. It remains to show that it is not possible to reintegrate some assets while keeping the first K assets and then dropping back the reintegrated assets to again hold only the first K assets. Since the constraint is binding, if asset $K + 1$ were to be reintegrated into the portfolio at $y = y_{N,K}^-$, the $K + 1$ asset's position as a function of y for y slightly below $y_{N,K}^-$ can be written as

$$\omega_{K+1}^* = \frac{A_{K+1} - B_{K+1}y}{y}, \tag{64}$$

with $B_{K+1} > 0$ (< 0) if the corresponding margin coefficient λ_{K+1} is equal to $\lambda^+(-\lambda^-)$, the first K components of vector λ being the same as when y is in $\left[y_{N,K}^-, y_{N,K}^+\right]$. If $B_{K+1} > 0$ (< 0), then for all values of $y < y_{N,K}^-$, $z_{K+1}^* > 0(< 0)$, which implies that asset $K + 1$ should not be dropped out of the portfolio *without* first dropping (at least) one of the first K assets held in the portfolio. It follows that a specific asset configuration can only hold for y within a single interval.

Asset reintegration condition. Assume that at $y = y_{N,J}^+ \leq y_B^*$, it is optimal to hold only (the first) J assets for some vector of margin coefficient $\lambda \in \Lambda$. By Lemma E.1, it is optimal to reintegrate asset $J+1$ into the portfolio at some lower level $y^-_{N,J} < y^+_{N,J}$ if and only $e^\intercal_{J+1}\omega^* \neq 0$ at all $y = y^-_{N,J} - \varepsilon$, with $\varepsilon > 0$ small, where allocation ω^* is given by Eq. (60) for $K = J+1$. This leads to the condition $\psi_{J+1}(\lambda, y^-_{N,J}) = \psi_J(\lambda, y^-_{N,J})$.

No asset reintegrated once asset with largest leveraged expected return held alone. Observe that for $y > 0$ small enough, assuming $\eta \neq 0$, the obvious optimal solution to the program defined by Eq. (36) is $\omega^* = (0, \ldots, \omega^*_i, \ldots, 0)$, with $\omega^*_i = 1/\lambda_i$, where asset i is such that $\eta_i/\lambda_i = \max_{k=1,..,N} \eta_k/\lambda_k$ for some $\lambda \in \Lambda$. Hence, for y small enough, only one asset is held in the portfolio. Next, observe that the program defined by Eq. (36) is equivalent to the following program

$$\max_{\omega \in Q} \omega^\top (\frac{\eta}{y} + \sigma\Sigma) - \frac{1}{2}\omega^\top \sigma\sigma^\top \omega, \tag{65}$$

where $\omega = \frac{z}{W}$. Without loss of generality, assume that at $y = y^*_1$ the solution of the optimization problem defined in Eq. (65) is $\omega^*_1 = \frac{1}{\lambda_1}$ with $\lambda_1 \in \{-\lambda^-, \lambda^+\}$, and $\omega^*_k = 0$ for $k = 2, \ldots, N$. We want to show that this is also the optimal solution for all $y < y^*_1$. The key is to observe that $y < y^*_1$

$$\max_{\omega \in Q} \omega^\top (\frac{\eta}{y} + \sigma\Sigma) - \frac{1}{2}\omega^\top \sigma\sigma^\top \omega \leq \max_{\omega \in Q} \left[\omega^\top (\frac{\eta}{y^*_1} + \sigma\Sigma) - \frac{1}{2}\omega^\top \sigma\sigma^\top \omega\right] + \max_{\omega \in Q} \omega^\top (\frac{\eta}{y} - \frac{\eta}{y^*_1}). \tag{66}$$

By assumption, the optimal solution to the optimization problem

$$\max_{\omega \in Q} \left[\omega^\top (\frac{\eta}{y^*_1} + \sigma\Sigma) - \frac{1}{2}\omega^\top \sigma\sigma^\top \omega\right], \tag{67}$$

is $(\frac{1}{\lambda_1}, 0, \ldots, 0)$ and it turns out that for $y < y^*_1$ the optimal solution to the optimization problem

$$\max_{\omega \in Q} \omega^\top (\frac{\eta}{y} - \frac{\eta}{y^*_1}), \tag{68}$$

is also $(\frac{1}{\lambda_1}, 0, \ldots, 0)$. The result follows.

Reduced HJB equation (first K assets held). Using the expressions for ω^* derived in Lemma E.1, we obtain the following reduced HJB equation

$$
\begin{aligned}
\left(\theta + (\gamma-1)(m - \gamma\frac{\Sigma^\intercal\Sigma + \Theta^\intercal\Theta}{2})\right) f(v) = &\frac{\gamma}{1-\gamma}\left(f'(v)\right)^{\frac{\gamma-1}{\gamma}} + f'(v) \\
&+ (B_K^{-1} + \gamma(\Sigma^\intercal\Sigma + \Theta^\intercal\Theta) - \gamma(I_K\Sigma)^\intercal I_K\Sigma + L_K)vf'(v) \\
&+ \frac{1}{2}\big(\Sigma^\intercal\Sigma + \Theta^\intercal\Theta + M_K^2(I_K\lambda)^\intercal(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}I_K\lambda \\
&\qquad -2M_K(I_K\lambda)^\intercal(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}I_K\sigma\Sigma\big)\, v^2 f''(v) \\
&- \frac{1}{2}\left((I_K\eta)^\intercal(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}I_K\eta - L_K^2(I_K\lambda)^\intercal(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}I_K\lambda\right)\frac{\left(f'(v)\right)^2}{f''(v)}
\end{aligned}
\tag{69}
$$

Note that the coefficient of the term $\left(f'(v)\right)^2/f''(v)$ is negative if $K>1$ by the Cauchy-Schwarz inequality, and equal to zero for $K=1$; the coefficient of the term $v^2f''(v)$ is equal to

$$
\Sigma^\intercal\Sigma + \Theta^\intercal\Theta - (I_K\Sigma)^\intercal I_K\Sigma + (M_K(I_K\sigma I_K^\intercal)^{-1}I_K\lambda + I_K\sigma\Sigma)^\intercal(M_K(I_K\sigma I_K^\intercal)^{-1}I_K\lambda + I_K\sigma\Sigma),
\tag{70}
$$

which is positive.

Deterministic income and general preferences.

The Hamilton-Jacobi-Bellman equation for the primal value function F is

$$
\theta F = \max_{\frac{z}{W}\in Q}\quad \widetilde{u}(F_1) + (rW+Y)F_1 + mYF_2 + z^\intercal(\mu - r\overline{1})F_1 + \tfrac{1}{2}z^\intercal\sigma\sigma^\intercal zW^2F_{11},
\tag{71}
$$

where $\widetilde{u}$ is the convex conjugate of u. This maximization problem is the same as the one solved for the CRRA preferences case so all the results found in the CRRA preference case apply. Furthermore, note that since $\Sigma=0$, margin coefficient $\lambda_{B,i}^*$ must have the same sign as $e_i^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1})$.

Appendix F. Proof of Proposition 6

When $N = 2$ and $\Sigma = 0$, the program defined by Eq. (36) becomes

$$\max_{\omega \in Q} \omega^\top(\mu - r\overline{1}) - \frac{y}{2}\omega^\top \sigma\sigma^\top \omega - \psi_2(\omega^\top\lambda - 1), \tag{72}$$

where $\lambda \in \Lambda$ and $\psi_2 \geq 0$ is the Lagrange multiplier. The first order condition is

$$\frac{z^*}{W} = \frac{(\sigma\sigma^\top)^{-1}}{y}(\mu - r\overline{1} - \psi_2\lambda), \tag{73}$$

and

$$\psi_2 = \frac{\lambda^\top(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}) - y}{\lambda^\top(\sigma\sigma^\intercal)^{-1}\lambda}. \tag{74}$$

The constraint starts binding at

$$y = y_B^* = \max_{\lambda\in\Lambda}\lambda^\top(\sigma\sigma^\top)^{-1}(\mu - r\overline{1}), \tag{75}$$

so that

$$\lambda_B^* = \arg\max_{\lambda\in\Lambda}\lambda^\top(\sigma\sigma^\top)^{-1}(\mu - r\overline{1}). \tag{76}$$

The covariance matrix is

$$\sigma\sigma^\top = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \tag{77}$$

so that

$$(\sigma\sigma^\top)^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}. \tag{78}$$

For $y \geq y_B^*$

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{y(1-\rho^2)\sigma_1^2\sigma_2^2}\left[\sigma_2^2(\mu_1 - r) - \rho\sigma_1\sigma_2(\mu_2 - r)\right] \\ \frac{z_2^*}{W} &= \frac{1}{y(1-\rho^2)\sigma_1^2\sigma_2^2}\left[\sigma_1^2(\mu_2 - r) - \rho\sigma_1\sigma_2(\mu_1 - r)\right], \end{aligned} \tag{79}$$

For $y \leq y_B^*$

$$\frac{z_1^*}{W} = \frac{1}{\lambda_1\left[\left(\frac{\sigma_1}{\lambda_1}\right)^2 + \left(\frac{\sigma_2}{\lambda_2}\right)^2 - 2\rho\frac{\sigma_1}{\lambda_1}\frac{\sigma_2}{\lambda_2}\right]}\left[\left(\frac{\mu_1 - r}{\lambda_1} - \frac{\mu_2 - r}{\lambda_2}\right)\frac{1}{y} + \left(\frac{\sigma_2}{\lambda_2}\right)^2 - \rho\frac{\sigma_1}{\lambda_1}\frac{\sigma_2}{\lambda_2}\right]$$

$$\frac{z_2^*}{W} = \frac{1}{\lambda_2\left[\left(\frac{\sigma_1}{\lambda_1}\right)^2 + \left(\frac{\sigma_2}{\lambda_2}\right)^2 - 2\rho\frac{\sigma_1}{\lambda_1}\frac{\sigma_2}{\lambda_2}\right]}\left[\left(\frac{\mu_2 - r}{\lambda_2} - \frac{\mu_1 - r}{\lambda_1}\right)\frac{1}{y} + \left(\frac{\sigma_1}{\lambda_1}\right)^2 - \rho\frac{\sigma_1}{\lambda_1}\frac{\sigma_2}{\lambda_2}\right]. \tag{80}$$

Let us assume that asset 1 is ultimately selected so that $\frac{\mu_1-r}{\lambda_1^*} > \max_{\lambda_2\in\{-\lambda^-,\lambda^+\}} \frac{\mu_2-r}{\lambda_2}$, for some $\lambda_1^* \in \{-\lambda^-, \lambda^+\}$ and set $\overline{\lambda}_1^* = \lambda^+$ if $\lambda_1^* = -\lambda^-$, $\overline{\lambda}_1^* = -\lambda^-$ if $\lambda_1^* = \lambda^+$.

General properties. For $y \leq y_B^*$, asset allocations (z_1, z_2) are given by Eq. (80) *provided that* it is possible to find a pair $(\lambda_1, \lambda_2) \in \Lambda$ such asset $\frac{z_i}{\lambda_i} \geq 0$, $i = 1, 2$, otherwise, only one asset is held in the portfolio. Second, recall that if asset 1 is held alone at $y = \overline{y}$, then it is optimal to hold only asset 1 for all $y \leq \overline{y}$ (Proposition 5). Third, it is never optimal to hold only asset 1 in a position (say long) for y in some interval and only asset 1 in the opposite position (say short) for y in some other interval. Fourth, define the dropping (D) and reintegrating (R) asset cutoffs

$$\begin{aligned} y_{2,D}^L &= \left[\frac{\frac{\mu_1-r}{\lambda_1^*} - \frac{\mu_2-r}{\lambda^+}}{\left(\frac{\sigma_1}{\lambda_1^*}\right)^2 - \rho\frac{\sigma_1}{\lambda_1^*}\frac{\sigma_2}{\lambda^+}}\right]^+ \text{ and } y_{2,D}^S = \left[\frac{\frac{\mu_1-r}{\lambda_1^*} + \frac{\mu_2-r}{\lambda^-}}{\left(\frac{\sigma_1}{\lambda_1^*}\right)^2 + \rho\frac{\sigma_1}{\lambda_1^*}\frac{\sigma_2}{\lambda^-}}\right]^+ \\ y_{1,R}^* &= \left[\frac{\frac{\mu_1-r}{\lambda_1^*} - \frac{\mu_2-r}{\lambda_2}}{-\left(\frac{\sigma_2}{\lambda_2}\right)^2 + \rho\frac{\sigma_1}{\lambda_1^*}\frac{\sigma_2}{\lambda_2}}\right]^+ \text{ and } y_{1,D}^* = \left[\frac{\frac{\mu_1-r}{\overline{\lambda}_1^*} - \frac{\mu_2-r}{\lambda_2}}{-\left(\frac{\sigma_2}{\lambda_2}\right)^2 + \rho\frac{\sigma_1}{\overline{\lambda}_1^*}\frac{\sigma_2}{\lambda_2}}\right]^+, \end{aligned} \tag{81}$$

and $y_{2,D}^* = \min^*\{y_{2,D}^L, y_{2,D}^S\}$. The value of λ_2 cannot change and is determined by the sign of asset 2 position at $y = y_B^*$ using Eq. (79). Given what precedes, by inspection, it is easy to check that the maximum number of regions that can be encountered is equal to five, namely $0 < y_{2,D}^* < y_{1,R}^* < y_{1,D}^* < y_B^*$. The special case $y_{1,R}^* = y_{1,D}^*$ occurs if and only if $S_{P_1} = \rho S_{P_2}$, where S_{P_1}, S_{P_2} are the Sharpe ratios of assets 1, 2 respectively. By inspection of Eq. (79) it must be the case that $y_{1,R}^* = y_{1,D}^* = y_B^* = \frac{\mu_2-r}{\lambda_2\sigma_2^2}$. On $\left[0, y_{2,D}^*\right]$ asset 1 is held alone and on $\left[y_{2,D}^*, y_B^*\right]$ both assets are held in the portfolio with the same sign. Thus, it is not possible to have four regions. Alternatively, we may have only three regions $0 < y^*_{2,D} < y^*_B$. Finally, the special case $y^*_{2,D} = y^*_B$ occurs if and only if $S_{P_2} = \rho S_{P_1}$ and there are only two regions.

To illustrate the three possible dynamics of the portfolio concentration once the margin constraint is binding, $y \le y^*_B$, we consider three parameter configurations below.

Five regions. We assume that $0 < \mu_2 - r < \mu_1 - r$ and $\frac{\mu_1 - r}{\sigma_1} < \rho \frac{\mu_2 - r}{\sigma_2}$, which implies that $0 < \sigma_2 < \rho\sigma_1$. Note that $\lambda^*_1 = \lambda^+$ and $\left(\lambda^*_{B,1}, \lambda^*_{B,2}\right) = (-\lambda^-, \lambda^+)$ so

$$y^*_B = (-\lambda^-, \lambda^+)^\top (\sigma\sigma^\top)^{-1}(\mu - r\overline{1}) > 0. \tag{82}$$

In this case, $\overline{\lambda}^*_1 = -\lambda^-$ and we have

$$y^*_{1,D} = \frac{\frac{\mu_1 - r}{\lambda^-} + \frac{\mu_2 - r}{\lambda^+}}{\left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho\frac{\sigma_1}{\lambda^-}\frac{\sigma_2}{\lambda^+}},\ y^*_{1,R} = \frac{\frac{\mu_1 - r}{\lambda^+} - \frac{\mu_2 - r}{\lambda^+}}{-\left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho\frac{\sigma_1}{\lambda^+}\frac{\sigma_2}{\lambda^+}},\ y^*_{2,D} = \frac{\frac{\mu_1 - r}{\lambda^+} - \frac{\mu_2 - r}{\lambda^+}}{\left(\frac{\sigma_1}{\lambda^+}\right)^2 - \rho\frac{\sigma_1}{\lambda^+}\frac{\sigma_2}{\lambda^+}}. \tag{83}$$

and one can check that indeed, $0 < y^*_{2,D} < y^*_{1,R} < y^*_{1,D} < y^*_B$.
On $\left[y^*_{1,D}, y^*_B\right]$ asset 1 is held (short) and asset 2 is held (long)

$$\begin{aligned}
\frac{z^*_1}{W} &= \frac{1}{\lambda^-\left[\left(\frac{\sigma_1}{\lambda^-}\right)^2 + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + 2\rho\frac{\sigma_1}{\lambda^-}\frac{\sigma_2}{\lambda^+}\right]}\left[\left(-\frac{\mu_1 - r}{\lambda^-} - \frac{\mu_2 - r}{\lambda^+}\right)\frac{1}{y} + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho\frac{\sigma_1}{\lambda^-}\frac{\sigma_2}{\lambda^+}\right] < 0 \\
\frac{z^*_2}{W} &= \frac{1}{\lambda^-\left[\left(\frac{\sigma_1}{\lambda^-}\right)^2 + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + 2\rho\frac{\sigma_1}{\lambda^-}\frac{\sigma_2}{\lambda^+}\right]}\left[\left(\frac{\mu_2 - r}{\lambda^+} + \frac{\mu_1 - r}{\lambda^-}\right)\frac{1}{y} + \left(\frac{\sigma_1}{\lambda^-}\right)^2 + \rho\frac{\sigma_1}{\lambda^-}\frac{\sigma_2}{\lambda^+}\right] > 0,
\end{aligned} \tag{84}$$

on $\left[y^*_{1,R}, y^*_{1,D}\right]$ only asset 2 is held (long)

$$\begin{aligned}
\frac{z^*_1}{W} &= 0 \\
\frac{z^*_2}{W} &= \frac{1}{\lambda^+},
\end{aligned} \tag{85}$$

on $\left[y^*_{2,D}, y^*_{1,R}\right]$ asset 1 is held (long) and asset 2 is held (long)

$$\begin{aligned}
\frac{z^*_1}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\left[\frac{\mu_1 - \mu_2}{y} + \frac{\sigma_2}{\lambda^+}(\sigma_2 - \rho\sigma_1)\right] > 0 \\
\frac{z^*_2}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\left[\frac{\mu_2 - \mu_1}{y} + \frac{\sigma_1}{\lambda^+}(\sigma_1 - \rho\sigma_2)\, y\right] > 0,
\end{aligned} \tag{86}$$

and finally on $\left[0, y^*_{2,D}\right]$ only asset 1 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\lambda^+} \\ \frac{z_2^*}{W} &= 0. \end{aligned} \tag{87}$$

Three regions. We assume that $0 < \mu_2 - r < \mu_1 - r$, $\frac{\mu_2 - r}{\sigma_2} > \rho \frac{\mu_1 - r}{\sigma_1}$, which implies $0 < \rho\sigma_2 < \sigma_1$. We have $\lambda_1^* = \lambda^+$, $\lambda^*_{B,1} = \lambda^*_{B,2} = \lambda^+$ and

$$y_B^* = (\lambda^+, \lambda^+)^\top (\sigma\sigma^\top)^{-1}(\mu - r\overline{1}) > 0. \tag{88}$$

Margin coefficient $\overline{\lambda}_1^*$ is irrelevant. We have $0 < y^*_{2,D} < y^*_B$ with

$$y^*_{2,D} = \frac{\mu_1 - \mu_2}{\frac{\sigma_1}{\lambda^+}(\sigma_1 - \rho\sigma_2)}. \tag{89}$$

On $\left[y^*_{2,D}, y^*_B\right]$ asset 1 is held (long) and asset 2 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\left[\frac{\mu_1 - \mu_2}{y} + \frac{\sigma_2}{\lambda^+}(\sigma_2 - \rho\sigma_1)\right] > 0 \\ \frac{z_2^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\left[\frac{\mu_2 - \mu_1}{y} + \frac{\sigma_1}{\lambda^+}(\sigma_1 - \rho\sigma_2)\right] > 0, \end{aligned} \tag{90}$$

and on $\left[0, y^*_{2,D}\right]$ only asset 1 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\lambda^+} \\ \frac{z_2^*}{W} &= 0. \end{aligned} \tag{91}$$

Two regions. We assume that $0 < \mu_2 - r < \mu_1 - r$, $\frac{\mu_2 - r}{\sigma_2} = \rho \frac{\mu_1 - r}{\sigma_1}$. We have $\lambda_1^* = \lambda^+$. It follows that

$$y^*_{2,D} = y^*_B = \lambda^+ \frac{\mu_1 - r}{\sigma_1^2} > 0. \tag{92}$$

Only asset 1 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\max\{y, y^*_B\}} \frac{\mu_1 - r}{\sigma_1^2} > 0 \\ \frac{z_2^*}{W} &= 0. \end{aligned} \tag{93}$$

Appendix G. Proof of Proposition 7

We assume that for $(i, j) = \{1, \ldots, K\}^2$,

$$\frac{\eta_i}{\lambda_i^*} = \frac{\eta_j}{\lambda_j^*} > 0, \tag{94}$$

Assume that, for $y \in \left[y_{N,K}^-, y_{N,K}^+\right]$, it is optimal to hold the first K assets in such a way that the condition in Eq. (94) is satisfied for $k = 1, \ldots, K$. The Lagrange multiplier given by Eq. (61) can be written as

$$\psi_K(y, \lambda^*) = \frac{\eta_1}{\lambda_1^*} - \frac{1 - (\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\intercal I_K^\top)^{-1} I_K \sigma \Sigma}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\intercal I_K^\top)^{-1} I_K \lambda^*} y, \tag{95}$$

which leads to the following optimal portfolio allocation

$$I_K \omega^* = (I_K \sigma \sigma^\intercal I_K^\top)^{-1} \left(\sigma \Sigma + \frac{1 - (\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\intercal I_K^\top)^{-1} I_K \sigma \Sigma}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\intercal I_K^\top)^{-1} I_K \lambda^*} \lambda^* \right). \tag{96}$$

Asset allocations are independent of assets' excess return η as well as the lifetime risk aversion y.

Remark. If we assume that $\frac{e_i^\top \sigma \Sigma}{\lambda_i^*} = \frac{e_j^\top \sigma \Sigma}{\lambda_j^*}$ for (i, j) in $\{1, \ldots, K\}^2$ we obtain that

$$I_K \omega^* = \frac{(I_K \sigma \sigma^\intercal I_K^\top)^{-1} \lambda^*}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\intercal I_K^\top)^{-1} I_K \lambda^*}, \tag{97}$$

which is only depends on the covariance matrix and the margin coefficients of the first K assets.

We now show that if all assets have the same leveraged expected excess return, i.e. the condition in Eq. (94) holds for $K = N$, then if at $y \in \left[y_{N,K}^-, y_{N,K}^+\right]$ exactly K assets are held in the portfolio, then the same K assets will be held in the same position for all $y \leq y_{N,K}^-$.

Step 1: We know that for $y \leq y_B^*$, the margin constraint is binding. For $y = 0$ the investor is indifferent between assets, therefore threshold y_K is well defined for some $K \in \{1, \ldots, N\}$.

Step 2: If no security among the K assets held is first dropped out of the portfolio, reintegrating asset $i \in \{K+1, \ldots, N\}$ into the portfolio is not optimal otherwise for all $\varepsilon > 0$ small enough, at $y = y^-_{N,K} + \varepsilon$ (respectively $y = y^-_{N,K} - \varepsilon$), the fraction of wealth invested in each asset is given by Eq. (96) for asset K (respectively asset $K+1$). Observe that these expressions are independent of risk aversion y, so as ε goes to 0 there will be a jump in asset allocations at $y = y^-_{N,K}$, which is impossible because of the continuity of the solution in parameter y.

Step 3: From Eq. (96), the expression for the asset allocation is, by assumption, optimal for values of the lifetime relative risk aversion y in $\left[y^-_{N,K}, y^+_{N,K}\right]$ and admissible for all y below $y^+_{N,K}$. From step 2, to reintegrate asset $K+1$, one asset among the K assets held must first be dropped, which cannot be optimal, since, from Lemma E.1, should it be possible to hold K assets whose positions are given by Eq. (96), holding only $K-1$ assets will be a dominated investment strategy.

Appendix H. Dual Approach: Fictitious Financial Market

Let a, b and κ be, respectively, an 1×1, an $N \times 1$ and an $M \times 1$ adapted stochastic processes to filtration $\mathbb{F}$ and consider the following fictitious financial market that consists of:

- a riskless bond $\widehat{B}$ with dynamics given by

$$d\widehat{B}_t = (r+a)\widehat{B}_t dt, \tag{98}$$

- N risky, nondividend paying securities whose prices evolve according to:

$$d\widehat{S}_t = I_{\widehat{S}_t}(\mu + b)dt + I_{\widehat{S}_t}\sigma dw_t, \tag{99}$$

- M additional, nondividend paying securities whose prices evolve according to:

$$d\widehat{P}_t = I_{\widehat{P}_t}\widehat{\mu}dt + I_{\widehat{P}_t}\widehat{\sigma}dw^Y_t, \tag{100}$$

where $\widehat{\mu}$ and $\widehat{\sigma}$ are respectively an $M \times 1$ and $M \times M$ adapted stochastic processes to filtration $\mathbb{F}$, such that $\kappa = -\widehat{\sigma}^{-1}(\widehat{\mu} - r\overline{1})$.

Dual formulation

A state price density $\pi^{a,b,\kappa}$ is an adapted stochastic process to filtration $\mathbb{F}$ defined by $\pi_0^{a,b,\kappa} = 1$ and

$$d\pi_t^{a,b,\kappa} = \pi_t^{a,b,\kappa} \left(-(r + a_t)dt - \left(\sigma^{-1}\left(b_t - a_t\overline{1} + \mu - r\overline{1}\right)\right)^\intercal dw_t + \kappa_t^\intercal dw_t^Y\right), \tag{101}$$

where a, b and κ are, respectively, an 1×1, an $N \times 1$ and an $M \times 1$ adapted stochastic process to filtration $\mathbb{F}$.

Effective domain

For $(a, b, \kappa) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^M$, let

$$e(a, b, \kappa) = \sup_{\frac{z}{z+x} \in Q} - ax - b^\intercal z. \tag{102}$$

The effective domain $\mathcal{N}$ is defined by

$$\mathcal{N} = \left\{(a, b, \kappa) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^M,\ e(a, b, \kappa) < \infty\right\}. \tag{103}$$

Lemma H.1. *Under the margin constraint, Eq. (4) of the paper, the effective domain is given by*

$$\mathcal{N} = \{(a, b, \kappa) \in \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathbb{R}^M,\ \kappa^+ a \leq b_i \leq \kappa^- a,\ i = 1, \ldots, N\}, \tag{104}$$

and $e(a, b, \kappa) \equiv 0$, *for all* $(a, b, \kappa) \in \mathcal{N}$.

Proof of Lemma H.1. The relationship $e(a, b, \kappa) \equiv 0$ comes from the fact that Q is a cone. Then, it is easy to see that we must have $a \geq 0$, $b_i \geq 0, i = 1, \ldots, N$. If $z_i \geq 0, i = 1, \ldots, N$ we have

$$- ax - b^\intercal z = -a\left(x + (1 - \lambda^+)\sum_{i=1}^{N} z_i\right) - \sum_{i=1}^{N}(b_i - (1 - \lambda^+)a)z_i. \tag{105}$$

Since $z_i \geq 0, i = 1, \ldots, N$ we must have $b_i - (1 - \lambda^+)a \geq 0$, $i = 1, \ldots, N$. Similarly, when $z_i \leq 0, i = 1, \ldots, N$, we have

$$-ax - b^\intercal z = -a\left(x + (1 + \lambda^-)\sum_{i=1}^{N} z_i\right) - \sum_{i=1}^{N}(b_i - (1 + \lambda^-)a)z_i. \qquad (106)$$

Since $z_i \leq 0, i = 1, \ldots, N$, we must have $b_i - (1 + \lambda^-)a \leq 0, i = 1, \ldots, N$. Since $\lambda^+ = \kappa^+ + 1$ and $\lambda^- = \kappa^- - 1$, the result follows.

Following the derivation in Cuoco (1997), for *some* suitable price density $\pi^* = \pi^{a^*,b^*,\kappa^*}$, the optimization problem, given in Eq. (9) of the paper, is equivalent to

$$\begin{aligned} F(W_0, Y_0) &= \max_c E_0\left[\int_0^\infty u(c_s)e^{-\theta s}ds\right] \\ \text{such that } E_0\left[\int_0^\infty \pi_s^* c_s ds\right] &= W_0 + E_0\left[\int_0^\infty \pi_s^* Y_s ds\right], \end{aligned} \qquad (107)$$

with $W_0 > 0$ and $Y_0 > 0$ given.

Appendix I. Dual Approach

To ensure that the optimization problem, given by Eqs. (9) of the paper, and (107) are equivalent, it is enough to determine the saddle point $(c^*, \phi^*, (a^*, b^*, \kappa^*))$ of the functional

$$\mathcal{L}(c, \psi, (a, b, \kappa)) = E_0\left[\int_0^\infty u(c_s)e^{-\theta s}ds\right] - \phi\left(E_0\left[\int_0^\infty \pi_s^{a,b,\kappa}(c_s - Y_s)ds\right] - W_0\right). \qquad (108)$$

The maximization over c yields $u'(c_s^*)e^{-\theta s} = \phi\pi_s^{a,b,\kappa}$ and the Lagrange multiplier ϕ^* is determined by the budget constraint

$$E_0\left[\int_0^\infty \pi_s^{a,b,\kappa}(I(\phi^*\pi_s^{a,b,\kappa}e^{\theta s}) - Y_s)ds\right] = W_0, \qquad (109)$$

where I is the inverse of the marginal utility function. We define the process $X^{a,b,\kappa}$:

$$X_t^{a,b,\kappa} = \phi^*\pi_t^{a,b,\kappa}e^{\theta t}. \qquad (110)$$

The dual value function J is given by

$$J(X_0, Y_0) = \min_{(a,b,\kappa)\in\mathcal{N}} E_0 \left[\int_0^\infty \left(\widetilde{u}(X_s^{a,b,\kappa}) + X_s^{a,b,\kappa} Y_s \right) e^{-\theta s} ds \right], \tag{111}$$

where $\widetilde{u}(X) = \max_{c\geq 0} u(c) - Xc$ is the convex conjugate of u. The solution of this minimization problem (a^*, b^*, κ^*) allows us to recover the state price density $\pi^* = \pi^{a^*,b^*,\kappa^*}$. For CRRA preferences, the convex conjugate is given by

$$\widetilde{u}(X) = \begin{cases} \frac{\gamma X^{\frac{\gamma-1}{\gamma}}}{1-\gamma} & , \quad \gamma \neq 1, \\ -\ln X - 1 & , \quad \gamma = 1. \end{cases} \tag{112}$$

Properties of the dual value function

Primal variables (F, W) and dual variables (J, X) are linked by the following Legendre transformation

$$W = -J_1(X, Y) \text{ and } X = F_1(W, Y). \tag{113}$$

As explained in He and Pagès (1993), J is nonincreasing and strictly convex in X. It is also easy to check that J is nondecreasing and concave in Y. For the case of a CRRA investor, the dual value function J can be written as $J(X, Y) = X^{\frac{\gamma-1}{\gamma}} h(X^{\frac{1}{\gamma}} Y)$, for some smooth function h. For convenience, let us write $\mathcal{N} = \mathcal{N}_{a,b} \times \mathbb{R}^M$. The dual value function J satisfies the following Hamilton-Jacobi-Bellman equation:

$$\begin{aligned}\theta J = & \frac{\gamma X^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + XY + (\theta - r) X J_1 + mY J_2 + \frac{\Sigma^\intercal \Sigma + \Theta^\intercal \Theta}{2} Y^2 J_{22} - \frac{\Sigma^\intercal \Sigma}{2} \frac{J_{12}^2}{J_{11}} \\ & + \min_{\kappa \in \mathbb{R}^M} \left\{ \frac{\kappa^\intercal \kappa}{2} X^2 J_{11} + \kappa^\intercal \Theta X Y J_{12} \right\} \\ & + \min_{(a,b)\in\mathcal{N}_{a,b}} \left\{ -aXJ_1 + \frac{X^2}{2} \left(b + \mu - (r+a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right)^\intercal (\sigma\sigma^\intercal)^{-1} \left(b + \mu - (r+a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right) J_{11} \right\}\end{aligned} \tag{114}$$

We obtain that $\kappa^* = -\frac{\Theta XYJ_{12}}{X^2 J_{11}}$, which leads to

$$\theta J = \frac{\gamma X^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + XY + (\theta - r)XJ_1 + mYJ_2 + \frac{\Sigma^\intercal\Sigma + \Theta^\intercal\Theta}{2} Y^2 \left(J_{22} - \frac{J_{12}^2}{J_{11}}\right)$$
$$+ \min_{(a,b)\in\mathcal{N}_{a,b}} \left\{ -aXJ_1 + \frac{X^2}{2}\left(b + \mu - (r+a)\bar{1} - \frac{\sigma\Sigma Y J_{12}}{XJ_{11}}\right)^\intercal (\sigma\sigma^\intercal)^{-1} \left(b + \mu - (r+a)\bar{1} - \frac{\sigma\Sigma Y J_{12}}{XJ_{11}}\right) J_{11} \right\} \tag{115}$$

Using the fact that $\gamma XJ_{11} = -J_1 + YJ_{12}$ and $-XJ_{11}/J_1 = 1/y$, the minimization problem is equivalent to

$$\min_{(a,b)\in\mathcal{N}_{a,b}} a + \frac{1}{2y}\left(\eta + y\sigma\Sigma + b - a\bar{1}\right)^\intercal (\sigma\sigma^\intercal)^{-1}\left(\eta + y\sigma\Sigma + b - a\bar{1}\right). \tag{116}$$

The minimization problem given by Eq. (116) and the maximization problem, given by Eq. (16) of the paper, are dual programs of one another: the solution a^* of the dual problem is equal to the Lagrange multiplier ψ of the primal problem. Within the nonbinding region, we find that $b_i^* = a^* = 0$. When K assets are optimally held — without loss of generality we can always assume the first K assets — the solution of program given by Eq. (116) is

$$\begin{aligned} a^* = \psi_K &= \frac{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\eta - (1 - (I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K\sigma\Sigma)y}{(I_K\lambda)^\intercal (I_K\sigma\sigma I_K^\intercal)^{-1} I_K\lambda} \\ b_k^* &= (1-\lambda_k)a^*,\ k = 1,\ldots,K, \end{aligned} \tag{117}$$

for some $\lambda \in \Lambda$, and the fraction of wealth invested in risky assets z^*/W is given by

$$I_K\frac{z^*}{W} = \frac{(I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}}{y} I_K(\eta - y\sigma\Sigma + b^* - a^*\bar{1}). \tag{118}$$

The last $N-K$ constraints of set $\mathcal{N}_{a,b}$ are non binding and the last $N-K$ components of vector b^* are such that $z_k^* = 0$, for $k = K+1, K+2, \ldots, N$.

Remark. Observe that the right hand side of Eq. (107) represents the lifetime resources of the investor. Even though an individual is not allowed to pledge his future labor income in any investment strategy and can only use his financial wealth W_0, his lifetime resources may by far exceed W_0. The margin requirement imposes a limit on the investor's maximum exposure to risky assets. When the margin requirement binds, the investor becomes fairly risk tolerant, which leads him to sacrifice diversification and load up his portfolio with assets that deliver a high expected return.

Remark. For the particular case of deterministic income and independent returns, the investor's choice can be thought of in terms of an adjusted Sharpe ratio for asset k, $\widehat{S}_{P,k}$, defined by

$$\widehat{S}_{P,k} = \frac{\mu_k + b_k^* - (r + a^*)}{\sigma_k}. \tag{119}$$

Inside the nonbinding region, for every asset k, the adjusted Sharpe ratio $\widehat{S}_{P,k}$ and the true Sharpe ratio $S_{P,k} = (\mu_k - r)/\sigma_k$ coincide since, when the constraint is not binding, $b_k^* = a^* = 0$. Inside the binding region with N assets, we have $b_k^* = (1 - \lambda_{B,k}^*)a^*$, for $k = 1, \ldots, N$ so indeed

$$\left| \widehat{S}_{P,k} \right| < \left| S_{P,k} \right|, \tag{120}$$

since $\mu_k - r$ and $\lambda_{B,k}^*$ have the same sign. Asset k is dropped out of the portfolio as soon as its adjusted Sharpe ratio $\widehat{S}_{P,k}$ becomes zero. Inside the binding region with only K assets, as the margin constraint becomes more binding, the adjusted Sharpe ratio of the remaining K risky assets shrinks, since a^* rises when y decreases. This result is in line with empirical findings by Ivković, Sialm, and Weisbenner (2008) who report that concentrated portfolios have lower Sharpe ratios.

Appendix J. Proof of Proposition 8

We prove Proposition 8 first for several special cases when shorting is not allowed and then for the general case. We also provide a complete characterization for the special case when the returns of the risky assets are independent, shorting is not allowed, and the margin requirement is the same for the market index fund and the market-weighted portfolio of risky assets. In this special case, the market index fund is the first asset dropped from the portfolio, irrespective of the characteristics of the risky assets. Both before and after the market index fund drops from the investor's portfolio, assets whose beta is less than, or equal to, one and that have volatility larger than the volatility of the market index fund, may remain in the investor's portfolio.

The maximization program is

$$\begin{aligned} \max_{(\omega,\omega_M)\in\mathbb{R}^{N+1}} & (\omega,\omega_M)^\intercal(\eta + y\sigma\Sigma, \eta_M + y\pi^\intercal\sigma\Sigma) - \frac{y}{2}(\omega,\omega_M)^\intercal V(\omega,\omega_M) \\ & \text{s.t. } (\omega^+)^\intercal\lambda^+ + (\omega^-)^\intercal\lambda^- + \lambda_M^+\omega_M^+ + \lambda_M^-\omega_M^- \leq 1, \end{aligned} \tag{121}$$

where $(\lambda^+)^\intercal = (\lambda_1^+, \lambda_2^+, \ldots, \lambda_N^+)^\intercal$ and $(\lambda^-)^\intercal = (\lambda_1^-, \lambda_2^-, \ldots, \lambda_N^-)^\intercal$.

First, note that since the objective function is continuous and the set over which the maximum is sought, $\{(\omega,\omega_M) \in \mathbb{R}^{N+1}, (\omega^+)^\intercal\lambda^+ + (\omega^-)^\intercal\lambda^- + \lambda_M^+\omega_M^+ + \lambda_M^-\omega_M^- \leq 1\}$ is compact, the maximum is achieved and at least one solution exists. The constraint $(\omega^+)^\intercal\lambda^+ + (\omega^-)^\intercal\lambda^- + \lambda_M^+\omega_M^+ + \lambda_M^-\omega_M^- \leq 1$ can be rewritten as: $\lambda^\intercal\omega + \lambda_M\omega_M \leq 1$, where $\lambda^\intercal = (\lambda_1, \lambda_2, \ldots, \lambda_N)^\intercal$, and $\lambda_k \in \{-\lambda_k^-, \lambda_k^+\}$, and $\omega_k/\lambda_k \geq 0$ for all $k \in \{1, \ldots, N\}$.

The Lagrangian of the maximization problem is

$$\begin{aligned} L =& (\omega,\omega_M)^\intercal(\eta + y\sigma\Sigma, \eta_M + y\pi^\intercal\sigma\Sigma) - \frac{y}{2}(\omega,\omega_M)^\intercal V(\omega,\omega_M) \\ & - \psi\left[(\omega,\omega_M)^\intercal(\lambda,\lambda_M) - 1)\right] + (\frac{\omega}{\lambda}, \frac{\omega_M}{\lambda_M})^\intercal(\varphi,\varphi_M), \end{aligned} \tag{122}$$

which leads to the following optimal condition

$$(\eta + y\sigma\Sigma, \eta_M + y\pi^\intercal\sigma\Sigma) - yV(\omega^*,\omega_M^*) - \psi(\lambda,\lambda_M) + (\frac{\varphi}{\lambda}, \frac{\varphi_M}{\lambda_M}) = 0, \tag{123}$$

where $\psi \geq 0$, with $\psi\left[(\omega^*,\omega_M^*)^\intercal(\lambda,\lambda_M) - 1)\right] = 0$ and $(\varphi/\lambda, \varphi_M/\lambda_M) \in \mathbb{R}_+^N \times \mathbb{R}_+$, such that $\varphi_k\omega_k^*/\lambda_k = 0$ and $\varphi_M\omega_M^*/\lambda_M = 0$, for $k = 1, \ldots, N$. We note that, by convention, if $(x,y) \in \mathbb{R}^N \times \mathbb{R}^N$, then $z^\intercal = \left(\frac{x}{y}\right)^\intercal = \left(\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots, \frac{x_N}{y_N}\right)$.

The optimal condition on the index fund holding is redundant. Since the system admits a (non-unique) solution the Lagrange multipliers must satisfy

$$\psi(\lambda_M - \pi^\intercal\lambda) - \frac{\varphi_M}{\lambda_M} + \pi^\intercal\frac{\varphi}{\lambda} = 0. \tag{124}$$

Manipulating the $N \times N$ system, we obtain that

$$\omega^* + \omega_M^*\pi = \frac{1}{y}(\sigma\sigma^\intercal)^{-1}(\eta + \frac{\varphi}{\lambda} + y\sigma\Sigma - \psi\lambda). \tag{125}$$

Observe that $\omega_k^* + \omega_M^* \pi_k$ is the total exposure of the portfolio to asset k, either by directly holding asset k with weight ω_k^*, and/or through the market index fund with weight $\omega_M^* \pi_k$.

We now investigate two special cases when short sales are prohibited.

Special Case 1: No Short Sales, Labor Income Uncorrelated with the Risky Assets.

Our analysis is divided in two parts: when the margin requirement is not binding and when the margin requirement is binding.

Part 1: Analysis when the margin requirement is not binding

First, observe that since $\pi \in \mathbb{R}^N_{++}$, holding all the securities long for large values of the lifetime relative risk aversion, y, is a feasible strategy. For $y > y_B^*$, the margin requirement is not binding and we have

$$\omega^* + \omega_M^* \pi = \frac{1}{y}(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) = \frac{\varpi\pi}{y} \in \mathbb{R}^N_+. \tag{126}$$

Let $(\lambda^+)^\intercal = (\lambda_1^+, \lambda_2^+, \ldots, \lambda_N^+)^\intercal$ denote the vector of long margin coefficients for the securities and λ_M^+ the margin coefficient for the market index fund. In order to determine the value of the binding threshold for the margin requirement, y_B^*, we need to distinguish several cases.

Special Case 1.1. Equal margin requirements: $\lambda_M^+ = \pi^\intercal \lambda^+$.

From Eq. (126), at $y = y_B^*$ we obtain that

$$1 = \frac{(\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1})}{y_B^*} + (\lambda_M^+ - \pi^\intercal\lambda^+)\omega_M^*, \tag{127}$$

so that

$$y_B^* = (\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}). \tag{128}$$

This is the same binding threshold as in the case when the market index fund is not available.

Special Case 1.2. Margin requirement for market index fund greater than weighted margin requirement for individual assets: $\lambda_M^+ > \pi^\intercal \lambda^+$.

Again at $y = y_B^*$, we have

$$1 = \frac{(\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1})}{y_B^*} + (\lambda_M^+ - \pi^\intercal\lambda^+)\omega_M^*. \tag{129}$$

Since the investor is always better off when the margin requirement is not binding, the optimal strategy must be such that the margin requirement starts binding at the lowest possible value of the lifetime relative risk aversion, y. As we assume $\lambda_M^+ > \pi^\intercal\lambda^+$, choosing $\omega_M^* = 0$ is optimal and

$$y_B^* = (\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}). \tag{130}$$

In this case the binding threshold is the same as in the case when the market index fund is not available. In addition, the market index fund is not held when the margin requirement starts binding.

Special Case 1.3. Margin requirement for market index fund smaller than weighted margin requirement for individual assets: $\lambda_M^+ < \pi^\intercal\lambda^+$.

Again, it is optimal to choose ω_M^* is such a way that the margin requirement starts binding for the smallest possible level of the lifetime relative risk aversion, y. As we assume that $\lambda_M^+ < \pi^\intercal\lambda^+$, choosing ω_M^* as large as possible, while compatible with $\omega^* + \omega_M^*\pi \in \mathbb{R}_+^N$, is optimal. We can choose $\omega_M^* = 1/\lambda_M^+$ and therefore we must have $\omega^* = 0$. It follows that

$$y_B^* = \lambda_M^+\bar{1}^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) = \frac{\lambda_M^+}{\varpi} \tag{131}$$

Observe that

$$y_B^* < (\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) \tag{132}$$

as

$$\frac{\lambda_M^+}{\pi^\intercal\lambda^+}(\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) = \lambda_M^+\bar{1}^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) \tag{133}$$

and, by assumption

$$\frac{\lambda_M^+}{\pi^\intercal\lambda^+} < 1. \tag{134}$$

This strategy is feasible since at $y = y_B^*$, if $\omega_M^* = \frac{1}{\lambda_M^+}$, from Eq. (126) we have

$$\omega^* = \left(\frac{1}{y_B^*} - \frac{\varpi}{\lambda_M^+}\right)(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}) = 0 \in \mathbb{R}_+^N. \tag{135}$$

Therefore, when the market index fund margin requirement is favorable compared to the margin requirements of the individual assets, the portfolio margin requirement starts binding at lower values of the lifetime relative risk aversion, y, and, at the value y_B^*, when the margin requirement starts binding, the investor only holds the market index fund in his portfolio.

Part 2: Analysis when the margin requirement is binding

We first derive the following Lemma.

Lemma J.1. *Choose $K < N$ assets and let J_K denote the $K \times N$ matrix whose first line is equal to e_k if asset k is among the K assets chosen and has the smallest index, second line is equal e_j if asset j is among the K assets chosen and has the second smallest index and so on. Let V_K be the covariance matrix formed by the set of the K chosen assets and the market index fund which has rank $K+1$ and is given by*

$$V_K = \begin{bmatrix} J_K\sigma\sigma^\intercal J_K^\intercal & J_K\sigma\sigma^\intercal\pi \\ \pi^\intercal\sigma\sigma^\intercal J_K^\intercal & \pi^\intercal(\sigma\sigma^\intercal)\pi \end{bmatrix}. \tag{136}$$

Then, we have that

$$V_K^{-1}\left[\left(J_K(\mu - r\overline{1}), \mu_M - r\right)\right] = (0, 0, .., 0, \varpi^{-1}). \tag{137}$$

Proof of Lemma J.1.

First, notice that

$$V_K = \begin{bmatrix} J_K\sigma\sigma^\intercal J_K^\intercal & \varpi J_K(\mu - r\overline{1}) \\ \varpi(\mu - r\overline{1})^\intercal J_K^\intercal & \varpi^2(\mu - r\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}) \end{bmatrix}. \tag{138}$$

Set

$$d = \varpi^2\left[(\mu - r\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}) - (\mu - r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu - r\overline{1})\right] > 0. \tag{139}$$

Inverting matrix V_K, we obtain that V_K^{-1} is equal to

$$\begin{bmatrix} \left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}\left(1+\frac{\varpi^2}{d}J_K(\mu-r\overline{1})(\mu-r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}\right) & -\frac{\varpi}{d}\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu-r\overline{1}) \\ -\frac{\varpi}{d}(\mu-r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1} & \frac{1}{d} \end{bmatrix}. \tag{140}$$

Let $I_{K,K+1}$ be the $K\times(K+1)$ matrix that consists of the first K rows of the $(K+1)\times(K+1)$ identity matrix. It follows that

$$\begin{aligned} &I_{K,K+1}V_K^{-1}[(J_K(\mu-r\overline{1}),\mu_M-r)] \\ &\quad=\frac{\varpi^2}{d}\left[\frac{d}{\varpi^2}+(\mu-r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu-r\overline{1})-(\mu-r\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\mu-r\overline{1})\right] \\ &\qquad\times\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu-r\overline{1}) \\ &\quad=0 \text{ (by definition of } d \text{ in Eq. 139).} \end{aligned} \tag{141}$$

It remains to check that the claim is true for the last component. Using the fact that $\mu_M-r=\pi^\intercal(\mu-r\overline{1})$, we have that

$$\begin{aligned} e_{K+1}^\intercal V_K^{-1}\left[\left(J_K(\mu-r\overline{1}),\mu_M-r\right)\right]=&-\frac{\varpi}{d}(\mu-r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu-r\overline{1})+\frac{1}{d}\pi^\intercal(\mu-r\overline{1}) \\ =&\frac{\varpi}{d}\left[-(\mu-r\overline{1})^\intercal J_K^\intercal\left(J_K\sigma\sigma^\intercal J_K^\intercal\right)^{-1}J_K(\mu-r\overline{1})\right. \\ &\left.\quad+(\mu-r\overline{1})^\intercal(\sigma\sigma^\intercal)^{-1}(\mu-r\overline{1})\right] \\ =&\varpi^{-1}. \end{aligned} \tag{142}$$

We now examine how asset selection takes place for values of the lifetime relative risk aversion, y, slightly below the value y_B^*, for which the margin requirement starts binding.

Special Case 1.1: $\lambda_M^+=\pi^\intercal\lambda^+$.

For y slightly below y_B^*, we have

$$\omega^*+\omega_M^*\pi=\frac{(\sigma\sigma^\intercal)^{-1}\left[\mu-r\overline{1}-\psi_N(\lambda^+,y)\lambda^+\right]}{y}, \tag{143}$$

where

$$\psi_N(\lambda^+, y) = \frac{(\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}) - y}{(\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}\lambda^+} > 0 \tag{144}$$

is decreasing in the lifetime relative risk aversion, y. As y decreases, eventually it reaches a threshold where exactly one component of vector $\omega^* + \omega_M^*\pi$ is equal to zero. Without loss of generality, we can assume that when $y = y_{N,D}^*$, then we have $\omega_N^* + \omega_M^*\pi_N = 0$. Since by assumption $\pi_N > 0$, we must have $\omega_M^* = 0$, i.e., the investor optimally chooses not to hold the market index fund as soon as dropping the first asset is optimal.

Special Case 1.2: $\lambda_M^+ > \pi^\intercal\lambda^+$.

For y slightly below y_B^*, only the N securities are held in a non-zero position in the portfolio. As argued above, it is never optimal to re-integrate the market index fund into the portfolio, since, if it were optimal to re-integrate the market index fund, it would be the next asset to be dropped as y decreases further, which leads to a contradiction.

Special Case 1.3: $\lambda_M^+ < \pi^\intercal\lambda^+$.

If

$$\frac{\mu_M - r}{\lambda_M^+} > \max_{k\in\{1,\ldots,N\}} \frac{\mu_k - r}{\lambda_k^+}, \tag{145}$$

then, for all $y \leq y_B^*$, the optimal portfolio is $\omega_M^* = 1/\lambda_M^+$ and $\omega^* = 0$, i.e., when the leveraged expected excess return of the market index fund is greater than the leveraged expected excess return of every risky asset, then, once the margin requirement binds, the investor only holds the market index fund in his portfolio.

Next, assume that

$$\frac{\mu_M - r}{\lambda_M^+} < \max_{k\in\{1,\ldots,N\}} \frac{\mu_k - r}{\lambda_k^+}. \tag{146}$$

All assets cannot be re-integrated into the portfolio at $y = y_B^* - \varepsilon$, $\varepsilon > 0$, otherwise the condition $\psi(\lambda_M^+ - \pi^\intercal\lambda^+) - \frac{\varphi_M}{\lambda_M^+} + \pi^\intercal\frac{\varphi}{\lambda^+} = 0$ would be violated: at least one (possibly more) asset is not re-integrated into the portfolio for y slightly below y_B^*. For values of the lifetime relative risk aversion, y, slightly below y_B^*, by continuity of the optimal solution in parameter y, the market index fund must be held and we assume that it is optimal to hold in non-zero positions K securities. The optimal asset allocation is given by

$$
\begin{aligned}
(J_K\omega^*, \omega_M^*) =& \frac{V_K^{-1}\left[(J_K(\mu - r\bar{1}), \mu_M - r) - \psi_{K+1}(y)(J_K\lambda^+, \lambda_M^+)\right]}{y} \\
=& \frac{1}{y}\left(-\psi_{K+1}(y) I_{K,K+1} V_K^{-1}(J_K\lambda^+, \lambda_M^+), \varpi^{-1} - \psi_{K+1}(y) e_{K+1}^\intercal V_K^{-1}(J_K\lambda^+, \lambda_M^+)\right),
\end{aligned}
\tag{147}
$$

with

$$
\begin{aligned}
\psi_{K+1}(y) =& \frac{(J_K\lambda^+, \lambda_M^+)^\intercal V_K^{-1}(J_K(\mu - r\bar{1}), \mu_M - r) - y}{(J_K\lambda^+, \lambda_M^+)^\intercal V_K^{-1}(J_K\lambda^+, \lambda_M^+)} \\
=& \frac{y_B^* - y}{(J_K\lambda^+, \lambda_M^+)^\intercal V_K^{-1}(J_K\lambda^+, \lambda_M^+)}.
\end{aligned}
\tag{148}
$$

We note that since $I_{K,K+1} V_K^{-1} I_K^\intercal J_K(\mu - r\bar{1}) = 0$, the set of K securities optimally held must be such that $-I_{K,K+1} V_K^{-1}(J_K\lambda^+, \lambda_M^+) \in \mathbb{R}_{++}^K$. Recall that the Lagrange multiplier ψ_{K+1} increases as the lifetime relative risk aversion, y, decreases. This implies that allocations in the K securities must be increasing as y decreases whereas the position in the market index fund is decreasing. We conclude that:

- The market index fund is the next asset to be dropped out of the portfolio at threshold value $y_{M,D}^*$ such that $\varpi^{-1} - \psi_{K+1}(y_{M,D}^*) e_{K+1}^\intercal V_K^{-1}(J_K\lambda^+, \lambda_M^+) = 0$, which implies that
$$
y_{M,D}^* = -\frac{(J_K\lambda^+)^\intercal I_{K,K+1} V_K^{-1}(J_K\lambda^+, \lambda_M^+)}{\varpi} > 0. \tag{149}
$$
- The market index fund is never re-integrated into the portfolio, since should this happen, as the lifetime relative risk aversion, y, decreases further, the market index fund will again be the first asset to be dropped out, which contradicts the fact that a particular asset configuration can only occur once, when y belongs to a particular interval.

Special Case 2: Independent Assets, $\mu - r\bar{1} > 0$, No Short Sales, No Labor Income Correlation and $\lambda_M^+ = \pi^\intercal\lambda^+$.

We find that the level of the lifetime relative risk aversion for which the margin requirement starts binding, y_B^*, is given by

$$y_B^* = (\lambda^+)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\bar{1}), \tag{150}$$

and for $y > y_B^*$, the optimal allocation in asset k is such that

$$\omega_k^* + \omega_M^*\pi_k = \frac{\mu_k - r}{y\sigma_k^2}, \tag{151}$$

with $(\omega_k^*, \omega_M^*) \in \mathbb{R}_+^2$. Note that since $\lambda_M = \pi^\intercal\lambda^+$, we have that

$$(\lambda^+)^\intercal\omega^* + \lambda_M\omega_M^* = (\lambda^+)^\intercal(\omega^* + \omega_M^*\pi) \tag{152}$$

is strictly below one for $y > y_B^*$, and equal to one for $y = y_B^*$. For y slightly below y_B^*, we have

$$\omega_k^* + \omega_M^*\pi_k = \frac{1}{y\sigma_k^2}(\mu_k - r - \psi_N(\lambda^+, y)\lambda_k^+) \tag{153}$$

and the Lagrange multiplier $\psi_N(\lambda^+, y)$ is equal to $(\alpha^\intercal\xi - y)/\alpha^\intercal\bar{1}$, where $\xi_k = (\mu_k - r)/\lambda_k^+$ and $\alpha_k = \left(\lambda_k^+/\sigma_k\right)^2$, $k \in \{1, \ldots, N\}$. It is possible to rewrite the optimal aggregate asset holding for security k as

$$\omega_k^* + \omega_M^*\pi_k = \frac{\alpha_k}{y\lambda_k^+\alpha^\intercal\bar{1}}(y - \alpha^\intercal(\xi - \xi_k\bar{1})). \tag{154}$$

At $y = y_B^*$, we must have $\psi_N(\lambda^+, y_B^*) = 0$, which leads to

$$y_B^* = \alpha^\intercal\xi. \tag{155}$$

Next, without loss of generality, assume that $0 < \xi_N < \xi_{N-1} < \cdots < \xi_1$. From Eq. (154), since $\xi \geq 0$ and $\alpha \geq 0$, it is easy to see that as the lifetime relative risk aversion, y, decreases, asset allocation $\omega_N^* + \omega_M^*\pi_N$ is the first allocation to hit zero at

$$y_{N,D}^* = \alpha^\intercal(\xi - \xi_N\bar{1}). \tag{156}$$

Since $(\omega_N^*, \omega_M^*) \in \mathbb{R}_+^2$, $\pi_N > 0$, it must be the case that at $y = y_{N,D}^*$, we have $\omega_N^* = \omega_M^* = 0$, i.e., the aggregate position in asset N, as well as the position in the market index fund, are equal to zero.

More generally, for $K \in \{1, \ldots, N\}$ define the cutoff where asset K is dropped from the investor's portfolio

$$y^*_{K,D} = (I_K \alpha)^\intercal \left[I_K(\xi - \xi_K \overline{1}) \right], \tag{157}$$

and by convention, set $y^*_{N+1,D} = y^*_B$; observe that $0 = y^*_{1,D} < y^*_{2,D} < \cdots < y^*_{N+1,D}$. For $k = 1, \ldots, N-1$, we have $K \in \{1, \ldots, N\}$ with

$$\omega^*_k = \frac{\alpha_k}{y \lambda^+_k} \frac{\left[y - y^*_{K,D} \right]^+}{(I_K \alpha)^\intercal \overline{1}}. \tag{158}$$

Thus, there are exactly $N+1$ regions: if $K = 1, \ldots, N-1$, when the value of the lifetime relative risk aversion, y, is such that $y^*_{K,D} < y \leq y^*_{K+1,D}$, only the first K assets are held in the portfolio, and the market index fund is not held. When $y^*_{N,D} < y \leq y^*_B$, the margin requirement is binding: all the N securities are held long in the portfolio and the investor may have a long position in the market index fund. Finally, when $y > y^*_B$, the margin requirement is not binding: all the N securities are held long in the portfolio and the investor may have a long position in the market index fund. Observe that, given our assumptions, asset $N-1$ in general may have a beta below one and/or a larger volatility than the market index fund and still, for all values of the lifetime relative risk aversion, $y \in \left[y^*_{N-1,D}, y^*_{N,D} \right]$, the investor optimally chooses to hold asset $N-1$ and to not hold the market index fund.

General Case, No Labor Income Correlation, $\Sigma = 0$

As long as the margin requirement is not binding, the optimal allocations satisfy

$$\omega^* = \left(\frac{1}{y} - \varpi \omega^*_M \right) \frac{(\sigma \sigma^\intercal)^{-1}(\mu - r\overline{1})}{y}. \tag{159}$$

As argued before, it is optimal to let the constraint bind at the lowest possible values y^*_B. Since investing nothing into the securities and holding a long position to the maximum allowed by the market index fund margin coefficient is a feasible strategy, we conclude that $y^*_B \leq \lambda^+_M / \varpi$. The key thing to observe is that y^*_B is such that $(y^*_B)^{-1} - \varpi \omega^*_M$ must be non-negative as we must have $\omega^*_M \leq \lambda^+_M$ and $y^*_B \leq \lambda^+_M / \varpi$. This implies that at $y = y^*_B$, the sign of the position in asset i is the same as the sign of $e_i^\intercal (\sigma \sigma^\intercal)^{-1}(\mu - r\overline{1})$, which pins down the value of the margin coefficients for the securities. Since, by assumption, $\pi > 0$, this implies that at $y = y^*_B$, on the aggregate all securities must be held in a long position. Then as argued before, it is optimal to choose the position in the market index fund, ω_M^*, such that $(\lambda_M - \pi^\intercal\lambda)\omega_M^*$ achieves the lowest possible value. Therefore, it is never optimal to short the index fund, so we must have $\lambda_M = \lambda_M^+$. It is then easy to see that

$$y_B^* = \begin{cases} \lambda^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}), \text{ if } \lambda_M^+ - \pi^\intercal\lambda > 0, \text{ when } \omega_M^* = 0 \text{ is optimal,} \\ \lambda_M^+\overline{1}^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}), \text{ if } \lambda_M^+ - \pi^\intercal\lambda = 0, \text{ when } \omega_M^* = \left[0, \frac{1}{\lambda_M^+}\right] \text{ is optimal} \\ \lambda_M^+\overline{1}^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1}), \text{ if } \lambda_M^+ - \pi^\intercal\lambda < 0, \text{ when } \omega_M^* = \frac{1}{\lambda_M^+} \text{ is optimal and } \omega^* = 0. \end{cases} \tag{160}$$

We now examine how asset selection takes place for $y < y_B^*$. If

$$\frac{\mu_M - r}{\lambda_M^+} > \max_{k\in\{1,\ldots,N\}} \max_{\lambda_k\in\{-\lambda_k^-,\lambda_k^+\}} \frac{\mu_k - r}{\lambda_k}, \tag{161}$$

then no securities are held at $y = y_B^*$ and are never re-integrated into the portfolio: for all $y \leq y_B^*$, $\omega^* = 0$ and $\omega_M^* = 1/\lambda_M^+$, i.e., only the market index fund is held when its leveraged expected excess return is greater than the leveraged expected excess returns of every risky asset.

Next, we assume that

$$\frac{\mu_M - r}{\lambda_M^+} < \max_{k\in\{1,\ldots,N\}} \max_{\lambda_k\in\{-\lambda_k^-,\lambda_k^+\}} \frac{\mu_k - r}{\lambda_k}. \tag{162}$$

For values of the lifetime relative risk aversion, y, slightly below y_B^*, the analysis conducted for the no short sale case still applies. In particular, if not already dropped from the portfolio, the market index fund is the first asset to be dropped from the portfolio, possibly at the same time as another security, as soon as y reaches a low enough level. The only case that remains to be investigated is the case where if at $y = y_K^*$, exactly K securities are held in the portfolio, some possibly in a short position, is it optimal to re-integrate the market index fund into the portfolio? The answer is no: should the market index fund be re-integrated into the portfolio at $y_{R,M}^* < y_K^*$, using Lemma J.1, we obtain that the set of K securities must be such that $-I_{K,K+1}V_K^{-1}(J_K\lambda, \lambda_M^+)$ has the same sign as the vector of margin coefficients $J_K\lambda$. Recall that the Lagrange multiplier ψ_{K+1} increases as the value of the lifetime relative risk aversion, y, decreases. As $J_K\omega^* = -y^{-1}\psi_{K+1}(y)I_{K,K+1}V_K^{-1}(J_K\lambda^+, \lambda_M^+)$, the allocations in the K securities must be increasing, in absolute value, as y decreases, and because the margin requirement is binding, the position in the market index fund has to be decreasing, in absolute value. Eventually, the market index fund drops out from the portfolio, which leads to a contradiction. We conclude that once dropped from the portfolio, the market index fund is never re-integrated into the portfolio at a lower level of the lifetime relative risk aversion, y.

Appendix K. Proof of Proposition 9

Investment inside the nonbinding region.

Recall that we assume $\Theta = 0$. We start with some properties of the optimal allocations inside the nonbinding region. Consumption, wealth and income are linked by the following relationship $W + BY = Ac + Kc^{1-\beta}Y^{\beta}$ or, equivalently, using reduced variables

$$v + B = Af'(v)^{-\frac{1}{\gamma}} + Kf'(v)^{\frac{\beta-1}{\gamma}}. \tag{163}$$

Applying Itô's lemma and identifying the coefficients with the wealth dynamics, the optimal portfolio allocations are given by

$$z^* = z^f - \beta K \frac{(\sigma\sigma^{\intercal})^{-1}\eta}{\gamma} f'(v)^{\frac{\beta-1}{\gamma}} Y. \tag{164}$$

When $e_i^{\intercal}(\sigma\sigma^{\top})^{-1}\eta > 0 (< 0)$, the constrained asset allocation z_i^* is lower (higher) than its unconstrained counterpart z_i^f. Next, we show that, inside the nonbinding region, income has the same effect on the constrained risky allocations as it has on the unconstrained ones. Differentiating Eq. (163) yields

$$\frac{f'(v)}{f''(v)} = -\frac{A}{\gamma} f'(v)^{-\frac{1}{\gamma}} + \frac{\beta - 1}{\gamma} K f'(v)^{\frac{\beta-1}{\gamma}} < 0. \tag{165}$$

From Eqs. (165) and (163) it is easy to check that the margin requirement is not binding for $f'(v)^{\frac{1}{\gamma}} \le Z^*$, for some $0 < Z^* < \widehat{Z}$ where $\widehat{Z} = \beta A/\left((\beta-1)B\right)$. Then, we have

$$\begin{aligned} \frac{\partial z^*}{\partial Y} &= (\sigma\sigma^\intercal)^{-1}\eta\left(B - \beta K f'(v)^{\frac{\beta-1}{\gamma}}(1 - \frac{\beta-1}{\gamma}\frac{vf''(v)}{f'(v)})\right) \\ &= \frac{(\sigma\sigma^\intercal)^{-1}\eta}{A-(\beta-1)Kf'(v)^{\frac{\beta}{\gamma}}}\left(AB + (\beta-1)^2 BKf'(v)^{\frac{\beta}{\gamma}} - \beta^2 AKf'(v)^{\frac{\beta-1}{\gamma}}\right). \end{aligned} \tag{166}$$

Set $Z = f'(v)^{\frac{1}{\gamma}}$ and for Z in $[0, Z^*]$, define the auxiliary function h with

$$h(Z) = AB + (\beta-1)^2 BKZ^\beta - \beta^2 AKZ^{\beta-1}. \tag{167}$$

h is a smooth function with $h'(Z) = \beta(\beta-1)^2 KBZ^{\beta-2}(Z - \widehat{Z}) < 0$, so it is decreasing on $[0, Z^*]$, since $Z^* < \widehat{Z}$. We want to show that h is positive on $[0, Z^*]$. First, note that $h(0) = AB > 0$. Then, for $Z = Z^*$, the margin constraint is binding and for $Z \le Z^*$ we have $(\lambda_B^*)^\top z^* \le W$ or, equivalently, using the expression of z^*

$$v(1 - \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1})}{\gamma}) \ge \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\eta}{\gamma}(B - \beta Kf'(v)^{\frac{\beta-1}{\gamma}}). \tag{168}$$

Using Eq. (163) we obtain that for all Z in $[0, Z^*]$

$$\begin{aligned} KZ^\beta &\ge \vartheta(Z - \overline{Z}) \\ K(Z^*)^\beta &= \vartheta(Z^* - \overline{Z}), \end{aligned} \tag{169}$$

where

$$\begin{aligned} \overline{Z} &= \frac{1 - \frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}(\mu - r\overline{1})}{\gamma}}{1 - (\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma}\frac{A}{B} > 0 \\ \vartheta &= \frac{B\left(1 - (\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma\right)}{1 - (\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\sigma\Sigma + (\beta-1)\frac{(\lambda_B^*)^\intercal(\sigma\sigma^\intercal)^{-1}\eta}{\gamma}} > 0. \end{aligned} \tag{170}$$

Finally, we have

$$h(Z^*) = \frac{B}{Z^*}(\beta\overline{Z} - (\beta-1)Z^*). \tag{171}$$

It remains to show that

$$Z^* \le \beta\overline{Z}/(\beta-1). \tag{172}$$

Set $x = Z/Z^*$ and $x^* = \overline{Z}/Z^* < 1$, so that for all $0 \le x \le 1$, we have

$$x^\beta \ge \frac{x - x^*}{1 - x^*}. \tag{173}$$

We want to show that this is the case if and only if

$$x^* \ge (\beta - 1)/\beta, \tag{174}$$

or, equivalently,

$$\beta \le \frac{1}{(1 - x^*)}. \tag{175}$$

For $x \in [0, 1]$, define the auxiliary function f with

$$f(x) = x^\beta - \frac{x - x^*}{1 - x^*}. \tag{176}$$

Observe that

$$\begin{aligned} f(0) =& \frac{x^*}{1 - x^*} > 0 \\ f(1) =& 0 \\ f'(x) =& \beta x^{\beta-1} - (1 - x^*)^{-1}. \end{aligned} \tag{177}$$

If $\beta > 1/(1-x^*)$, then $f'(1) > 0$ and since $f(1) = 0$, it must be the case that $f(1-\varepsilon) < 0$, for some $\varepsilon > 0$ small enough. This leads to a contradiction since by the condition in Eq. (169) f is non-negative on $[0, 1]$. Thus, we must have $\beta \le 1/(1-x^*)$ or, equivalently,

$$Z^* \le \frac{\beta \overline{Z}}{\beta - 1}. \tag{178}$$

It follows that $h(Z^*) \ge 0$ and for all Z in $[0, Z^*)$, $h(Z) > 0$. We can conclude that z_i^* is increasing (decreasing) with income exactly when $e_i^\top (\sigma\sigma^\top)^{-1}\eta > 0 (< 0)$. Finally, since

$$\frac{z^*}{W} = (\sigma\sigma^\intercal)^{-1}\sigma\Sigma + \frac{(\sigma\sigma^\intercal)^{-1}\eta}{y}, \tag{179}$$

we deduce that

$$\frac{\partial}{\partial Y}\left(\frac{1}{y}\right) \ge 0, \tag{180}$$

which implies that

$$\frac{\partial y}{\partial Y} \leq 0. \tag{181}$$

Furthermore, since

$$\frac{\partial y}{\partial Y} = -v\frac{\partial y}{\partial v}, \tag{182}$$

we find that

$$\frac{\partial y}{\partial v} \geq 0. \tag{183}$$

At $Y = 0$; i.e., when v is infinite, $y = \gamma$, so we deduce that for all v inside the non-binding region, $y < \gamma$. Finally, note that z^*/W rises as v and W decrease.

Global properties of the optimal consumption c^*.

Recall that

$$c^* = Yf'(v)^{-\frac{1}{\gamma}}, \tag{184}$$

so

$$\frac{\partial c^*}{\partial W} = -\frac{f''(v)f'(v)^{-\frac{1}{\gamma}-1}}{\gamma} > 0. \tag{185}$$

Then

$$\frac{\partial c^*}{\partial Y} = \frac{f'(v)^{-\frac{1}{\gamma}}}{\gamma}\left(\gamma - y\right). \tag{186}$$

Inside the nonbinding region, we have seen that $y < \gamma$, and inside the binding region, we must have $y < y_B^* < \gamma$. Hence, we always have $y < \gamma$ and we conclude that $\partial c^*/\partial Y > 0$.

Appendix L. Proof of Proposition 10

For $y < y_B^*$, the Hamilton-Jacobi-Bellman equation is such that the coefficient of the term $v^2 f''(v)$ is positive and the coefficient of the term $-(f'(v))^2/f''(v)$ is nonnegative. This is exactly the same type of ODE studied by Duffie, Fleming, Soner, and Zariphopoulou (1997). In Proposition 1 of their paper, these authors establish that

$$\lim_{v\downarrow 0} f'(v) \tag{187}$$

exists, is positive and finite. They also show that

$$\lim_{v\downarrow 0}\sup -vf''(v) = 0. \tag{188}$$

Since

$$0 < -vf''(v) \leq \sup_{0<x\leq v} - xf''(x), \tag{189}$$

it follows that

$$\lim_{v\downarrow 0} - vf''(v) = 0. \tag{190}$$

Hence, we have

$$\lim_{v\downarrow 0} - \frac{vf''(v)}{f'(v)} = 0. \tag{191}$$

Around $v = 0$, we postulate the following asymptotic expansion

$$f(v) \underset{0}{\sim} d_0 + v - d_1 v^{\frac{3}{2}} + d_2 v^2 + o(v^2), \tag{192}$$

for some constants d_0, $d_1 > 0$ and d_2 to be determined. Our choice for $f'(0) = 1$ is justified because if $f'(0) = 1$, the quantity

$$\frac{\gamma}{1-\gamma}\left(f'(v)\right)^{\frac{\gamma-1}{\gamma}} + f'(v) \tag{193}$$

achieves its maximum value for $v = 0$. Using the Hamilton-Jacobi-Bellman Eq. (69) for $K = 1$ and identifying coefficients, we obtain

$$f(0) = d_0 = \frac{1}{(1-\gamma)\left(\theta + (\gamma-1)(m - \gamma\frac{\Sigma^\top\Sigma + \Theta^\top\Theta}{2})\right)} > 0, \tag{194}$$

and

$$\theta + (\gamma-1)(m - \gamma\frac{\Sigma^\intercal\Sigma + \Theta^\top\Theta}{2}) = \frac{9}{8\gamma}d_1^2 + (r - m + \gamma(\Sigma^\intercal\Sigma + \Theta^\top\Theta) + \frac{\eta_1}{\lambda_1^*}). \tag{195}$$

It follows that

$$d_1 = \frac{2\sqrt{2\gamma(\theta + \gamma(m - (\gamma+1)\frac{\Sigma^\intercal\Sigma + \Theta^\top\Theta}{2})) - (r + \eta_1/\lambda_1^*)}}{3} > 0. \tag{196}$$

This implies that

$$c^* \underset{0}{\sim} Y \tag{197}$$

and,

$$y = -\frac{vf''(v)}{f'(v)} \underset{0}{\sim} \frac{3d_1 v^{\frac{1}{2}}}{4}. \tag{198}$$

Appendix M. Proof of Proposition 11

When income is deterministic, we have, in the dual formulation, $\kappa^* \equiv 0$. Notice that $u'(c_t^*) = X_t^{a^*,b^*,0}$ with

$$dX_t^{a^*,b^*,0} = X_t^{a^*,b^*,0}(-(r + a_t^*)dt + (\zeta_t^{a^*,b^*})^\intercal dw_t), \tag{199}$$

where

$$\zeta_t^{a^*,b^*} = -\sigma^{-1}\left(b_t^* - a_t^*\overline{1} + \mu - r\overline{1}\right). \tag{200}$$

Using Itô's lemma, we find that the consumption growth rate is given by

$$\frac{dc_t^*}{c_t^*} = \left(\frac{r + a_t^* - \theta}{RR(c_t^*)} + \frac{1}{2}\frac{RP(c_t^*)}{(RR(c_t^*))^2}\left\|\zeta_t^{a^*,b^*}\right\|^2\right)dt + \frac{(\zeta_t^{a^*,b^*})^\intercal}{RR(c_t^*)}dw_t, \tag{201}$$

where

$$RR(c) = -\frac{cu''(c)}{u'(c)} \tag{202}$$

is the relative risk aversion ratio and

$$RP(c) = -\frac{cu'''(c)}{u''(c)} \tag{203}$$

is the relative risk prudence ratio. The instantaneous volatility of consumption is given by $\left\|\zeta_t^{a^*,b^*}\right\|^2 / (RR(c_t^*))^2$. We now show that for all $t \geq 0$, $\left\|\zeta_t^{a^*,b^*}\right\|^2 \leq \left\|\zeta^{0,0}\right\|^2$. Inside the nonbinding region, we have $\zeta_t^{a^*,b^*} = \zeta^{0,0}$. Inside the binding region when K assets are held, for some $\lambda \in \Lambda$, we have

$$\begin{aligned} b^* &= (\overline{1} - \lambda)a^* \\ a^* &= \frac{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K(\mu - r\overline{1}) - y}{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}(I_K\lambda)} > 0, \end{aligned} \tag{204}$$

so that

$$\left\|\zeta^{a^*,b^*}\right\|^2 = (I_K(\mu - r\bar{1} - \lambda a^*))^\intercal (I_K \sigma\sigma^\intercal I_K^\intercal)^{-1} I_K(\mu - r\bar{1} - \lambda a^*). \tag{205}$$

Hence for all $\lambda \in \Lambda$

$$\frac{\partial}{\partial y}\left\|\zeta^{a^*,b^*}\right\|^2 = -2\frac{\partial a^*}{\partial y}(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1} I_K(\mu - r\bar{1} - \lambda a^*) = \frac{2}{(I_K\lambda)^\intercal (I_K\sigma\sigma^\intercal I_K^\intercal)^{-1}(I_K\lambda)} > 0. \tag{206}$$

Given what precedes, since at $y = y_B^*$ we have $\left\|\zeta^{a^*,b^*}\right\| = \|\zeta^{0,0}\|$, we deduce that for all $y \leq y_B^*$, $\left\|\zeta^{a^*,b^*}\right\| \leq \|\zeta^{0,0}\|$.

Appendix N. Numerical Algorithm

N.1. Model Setup

Market

The continuous-time dynamics of the asset values and income changes are given by Eqs. (1), (2), and (3) in the paper. We approximate the continuous-time dynamics by a discrete-time Markov chain using the discretization described in He (1990). In this discretization an N dimensional multivariate normal distribution is described by $N+1$ nodes. Discretizing returns in this fashion preserves market completeness in discrete time.

Optimization problem

We consider the optimization problem described in Eq. (9) of Section 2 of the paper in a discrete-time setting, where the investor starts working at time 0 and retires at time T. From the discussion of homogeneity in Section 2 of the paper, we can reduce the number of state variables after scaling by income Y_t and obtain the following Bellman equation at $t = 0, \ldots, T-1$:

$$\begin{array}{ll} f_t(v_t) = & \max_{q_t,\omega_t} u(q_t) + \beta E_t\left[g_t^{1-\gamma} f_{t+1}(v_{t+1})\right] \\ \text{s.t.} & v_{t+1} = g_t^{-1}(v_t + 1 - q_t)\left(\sum_{i=1}^N \omega_{i,t} R_{i,t}^e + R^f\right) \\ & \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- \leq 1 \\ & f_T = \phi_\tau \frac{(v_T+1)^{1-\gamma}}{1-\gamma} \end{array} \tag{207}$$

where $v_t = W_t/Y_t$ is the wealth over income ratio; $q_t = c_t/Y_t$ is the consumption over income ratio; $\omega_t = z_t/W_t$ is the portfolio weight; $g_t = Y_{t+1}/Y_t$ is the income growth rate over period t; R^e is the expected one period excess asset return; R^f is the one period return of the money-market account; $f_t(v_t) = Y_t^{-(1-\gamma)} F_t(W_t, Y_t)$ is the reduced value function; and the factor ϕ_τ captures the effect of the investor's remaining lifetime. If the investor's remaining life is τ years, and the opportunity set remains constant, then the factor ϕ_τ is given by

$$\begin{aligned} \phi_\tau &= \left[\frac{1-(\beta\alpha)^{1/\gamma}}{1-(\beta\alpha)^{(\tau+1)/\gamma}}\right]^{-\gamma}, \\ \alpha &= E\left[\left(\sum_{i=1}^{N}\omega_i^* R_i^e + R^f\right)^{1-\gamma}\right] \end{aligned} \tag{208}$$

where ω^* are the optimal portfolio weights after retirement — see Ingersoll (1987).

N.2. Solution Methodology

To solve the problem described in Eq. (207), we extend the method proposed by Brandt, Goyal, Santa-Clara, and Stroud (2005) to incorporate endogenous state variables and constraints on portfolio weights. We also use an iterative method to find the solution to the Karush-Kuhn-Tucker (KKT) conditions; i.e., the first order conditions with constraints. The idea is to approximate the conditional expectations in the KKT conditions locally within a region that contains the solution to the KKT conditions and iteratively contract the size of the region.

As suggested by Carroll (2006), we separate consumption optimization from portfolio optimization in Eq. (207) by defining a new variable, total investment I_t:

$$I_t = v_t - q_t + 1. \tag{209}$$

At the optimal value of consumption, q_t^*, Eq. (209) defines an one-to-one correspondence between wealth v_t and total investment I_t. Therefore we can specify a particular grid, G, either through wealth, $v_t(G)$, or, equivalently, through investment, $I_t(G)$. Specifying $I_t(G)$ instead of $v_t(G)$ allows splitting the problem in Eq. (207) into two subproblems:

[Portfolio Optimization]

$$
\begin{array}{lll}
f_t^p\left(I_t\right) = & \max\limits_{\omega_t} \beta E_t\left[g_t^{1-\gamma} f_{t+1}\left(v_{t+1}\right)\right], & t = 0, \ldots, T-1 \\
\text{s.t.} & v_{t+1} = g_t^{-1} I_t \left(\sum_{i=1}^N \omega_{i,t} R_{i,t}^e + R^f\right) & \\
& \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- \leq 1 &
\end{array}
\tag{210}
$$

[Consumption Optimization]

$$
f_t\left(v_t\right) = \max_{q_t} u\left(q_t\right) + f_t^p\left(v_t - q_t + 1\right), t = 0, \ldots, T-1, \tag{211}
$$

where $f^p(\cdot)$ is the value function of the portfolio optimization problem in Eq. (210). Given the separation of consumption and portfolio optimization, we use the following algorithm to solve the problem in Eq. (207):

Algorithm

Step 1: Set the terminal condition at time T.

Step 2: Find the optimal portfolio and consumption backwards at $t = T-1, T-2, \cdots, 0$:

- Step 2.1: Construct a grid for total investment I_t with n_g grid points $\{I_t^i\}_{i=1}^{n_g}$.
- Step 2.2: Find the optimal portfolio and consumption at each grid point $I_t^i, i = 1, \cdots, n_g$:
  - Step 2.2.1: [Portfolio optimization] given I_t^i, find $\omega_t^*\left(I_t^i\right)$ by solving Eq. (210).
  - Step 2.2.2: [Consumption optimization] given $\{I_t^i, \omega_t^*\left(I_t^i\right)\}$, find $q_t^*\left(I_t^i\right)$ by solving Eq. (211).
  - Step 2.2.3: Recover state variable v_t at grid point i by $v_t^i = I_t^i + q_t^*\left(I_t^i\right) - 1$.

After specifying the factor ϕ_τ, Step 1 is trivial. Step 2.1 requires constructing a grid in an one-dimensional space. To account for the nonlinearity of the value function at lower wealth levels we place more grid points toward the lower investment values in a double exponential manner as suggested by Carroll (2006).

N.3. Portfolio Optimization

Given a grid point I_t^i, $i = 1, \cdots, n_g$, we want to optimize over ω_t by solving Eq. (210). To simplify the problem, and slightly abusing notation, we consider ω_t^+, ω_t^- as choice variables, such that $\omega_t^+ \geq 0, \omega_t^- \geq 0, \omega_t = \omega_t^+ - \omega_t^-$ and solve the following problem:

$$\begin{aligned} f_t^p(I_t) = \quad & \max_{\omega_t^+, \omega_t^-} \beta E_t \left[g_t^{1-\gamma} f_{t+1}(v_{t+1}) \right] \\ s.t. \quad & v_{t+1} = g_t^{-1} I_t \left[\sum_{i=1}^N \left(\omega_{i,t}^+ - \omega_{i,t^-} \right) R_{i,t}^e + R^f \right] \\ & \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- \leq 1 \\ & \omega_{i,t}^+, \omega_{i,t}^- \geq 0, i = 1, \cdots, N \end{aligned} \tag{212}$$

Notice that to maintain equivalence between Eqs. (210) and (212) we also need the constraints $\omega_{i,t}^+ \omega_{i,t}^- = 0$ for $i = 1, \cdots, N$, in Eq. (212). However, one can show that dropping these constraints will expand the feasible region but will not introduce new optimal solutions which are non-trivially different.

The Lagrangian and KKT conditions of the problem in Eq. (212) are given by:

Lagrangian

$$\begin{aligned} \mathcal{L}^p \left(\omega_t^+, \omega_t^-, l_t^+, l_t^-, l_t^m \right) = \quad & \beta E_t \left[g_t^{1-\gamma} f_{t+1}(v_{t+1}) \right] + \sum_{i=1}^N l_{i,t}^+ \omega_{i,t}^+ + \sum_{i=1}^N l_{i,t}^- \omega_{i,t}^- \\ & + l_t^m \left(1 - \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ - \lambda^- \sum_{i=1}^N \omega_{i,t}^- \right) \end{aligned} \tag{213}$$

KKT conditions

$$\begin{aligned} & 0 = \beta I_t E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} R_{i,t}^e \right\} + l_{i,t}^+ - l_t^m \lambda^+, i = 1, \ldots, N && \text{FOCs} \\ & 0 = -\beta I_t E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} R_{i,t}^e \right\} + l_{i,t}^- - l_t^m \lambda^-, i = 1, \ldots, N && \text{FOCs} \\ & 0 = l_{i,t}^+ \omega_{i,t}^+, i = 1, \cdots, N && \text{Complementarity} \\ & 0 = l_{i,t}^- \omega_{i,t}^-, i = 1, \cdots, N && \text{Complementarity} \\ & 0 = l_t^m \left(1 - \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ - \lambda^- \sum_{i=1}^N \omega_{i,t}^- \right) && \text{Complementarity} \\ & 1 \geq \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- && \text{Feasibility} \\ & 0 \leq \omega_{i,t}^+, \omega_{i,t}^-, l_{i,t}^+, l_{i,t}^-, l_t^m, i = 1, \cdots, N && \text{Feasibility} \end{aligned} \tag{214}$$

where l_t^m is the Lagrange multipliers of the margin constraint; l_t^+ and l_t^- are the Lagrange multipliers of the non-negativity constraints. While in general the KKT conditions are only necessary for optimality, for the problem in Eq. (212) the KKT conditions are both necessary and sufficient since the objective function is concave in $\left(\omega_t^+, \omega_t^-\right)$ and all constraints are linear in $\left(\omega_t^+, \omega_t^-\right)$.

Solving the KKT conditions requires enumeration of all the possibilities for the complementary conditions. In general, the $2N+1$ Lagrange multipliers $(l_t^m, l_{i,t}^+, l_{i,t}^-, i = 1, \cdots, N)$ give 2^{2N+1} possible specifications of the complementary conditions. However many of these specifications can be combined or ignored: if the margin constraint is not binding ($l_t^m = 0$) we only need to solve the FOCs without splitting ω_t as $\omega_t^+ - \omega_t^-$; if the margin constraint is binding ($l_t^m > 0$) we can ignore all the specifications with $\omega_{i,t}^+ \omega_t^- > 0, i = 1, \cdots, N$, since these specifications are not optimal. Overall there are $3^N + 1$ specifications that need to be checked. Once a solution to the KKT conditions under any of these specifications is found we can stop since the sufficiency of the KKT conditions guarantees optimality.

Approximation of conditional expectations

We use functional approximation to approximate conditional expectations in the KKT conditions as a linear combination of basis functions:

$$E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}\left(v_{t+1}\right)}{\partial v_{t+1}} R_{i,t}^e \middle| I_t, \omega_t^+, \omega_t^- \right\} \approx \sum_{j=1}^{n_b} \alpha_{ij}\left(I_t\right) b_j\left(\omega_t\right), i = 1, \cdots, N, \tag{215}$$

where n_b is the number of basis functions and $\{b_j(\cdot)\}_{j=1}^{n_b}$ are the basis functions on portfolio weights $\omega_t = \omega_t^+ - \omega_t^-$. The coefficients $\alpha_{ij}\left(I_t\right)$ at each investment grid point $\{I_t^i\}_{i=1}^{n_g}$ are estimated through cross-test-solution regression in the following way: we randomly generate n_s test solutions $\left\{\omega_t^{(k)}\right\}_{k=1}^{n_s}$ within a set called the test region. To guarantee that all the test solutions are feasible we assume that the test region is included in the set of all feasible solutions Q. For each test solution $\omega_t^{(k)}$ we evaluate the basis functions at the test solution $\left\{b_j(\omega_t^{(k)})\right\}_{j=1}^{n_b}$; given the test solution $\omega_t^{(k)}$ and the investment level I_t, we generate returns for the risky assets following the discretization procedure described in He (1990) and compute the expectation of the left-hand-side of Eq. (215); the weights $\alpha_{ij}(I_t)$ are estimated by OLS regression across the n_s test solutions. The basis functions we use are powers of the choice variables up to third order. We use the multidimensional root-finding solver of the GSL library to solve the KKT conditions. We use 300 grid points and 300 test solutions after checking that the results do not change if 500 grid points and 500 test solutions are used.

Test region iterative contraction (TRIC)

TRIC is a method introduced in Yang (2010) to improve the accuracy of the functional approximation approach for solving the dynamic portfolio choice problem. When we approximate the conditional expectation in Eq. (215) through cross-test-solution regressions, the quality of the approximation is affected by the number of basis functions n_b, the number of test solutions n_s, and the size of the test region: keeping n_b and n_s constant, the smaller the test region, the more accurate the approximation. This motivates the method of contracting the test region in an iterative manner: at each iteration i, we estimate the approximation in Eq. (215) with test solutions generated within $Q^{(i)}$; using this approximation we solve the KKT conditions to find $\omega^{(i)}$; if $\omega^{(i)} \in Q^{(i)}$ we contract the test region of the next iteration to $Q^{(i+1)} \subset Q^{(i)}$; if the new solution is outside the test region, $\omega^{(i)} \notin Q^{(i)}$, we enlarge the test region of the next iteration to $Q^{(i)} \subset Q^{(i+1)} \subset Q^{(i-1)}$; after each iteration, we check convergence by computing the relative change in portfolio weights $\left\|\omega^{(i)} - \omega^{(i-1)}\right\| / \left\|\omega^{(i-1)}\right\|$, where $\|x\|$ is the norm of x, defined by $\|x\|^2 = \text{Trace}(x^\intercal x)$, and comparing it with a threshold ε. In our numerical tests we contracted the test region by 50%. If the test region did not contain the solution, we expanded the test region by 150%. In the results we report the algorithm converged within two to three iterations for most grid points.

To start the procedure we need an initial test region $Q^{(0)}$ that contains the optimal solution. If no further information is available we can set $Q^{(0)} = Q$, the feasible region of the problem, defined in Eq. (212). However, it is possible to obtain a smaller $Q^{(0)}$ if we know the solution for similar parameter values, called a reference solution. We have used our knowledge of the asymptotic behavior of the solutions to construct reference solutions: for each time period we always solve from the grid point with the highest investment level down to the grid point with the lowest investment level; the solution at the higher level grid point serves as the reference solution for the adjacent lower level grid point; when we change between time periods the reference solution at the highest level grid point is set by linearly interpolating the solutions at the next period; at the last time period, $t = T - 1$, the reference solution at the highest level grid point, where the margin constraint is not binding, is set to the analytical solution.

N.4. Consumption Optimization and Value Function Sensitivity

After the optimal portfolio at an investment grid point has been found, we find the optimal level of consumption at that grid point by solving the consumption optimization problem defined in Eq. (211). The first order condition leads to

$$q_t^{-\gamma} = \frac{\partial f_t^p (I_t)}{\partial I_t}. \tag{216}$$

To evaluate the term $\partial f_t^p (I_t) / \partial I_t$, we apply the envelope theorem to the Lagrangian $\mathcal{L}^p$ in Eq. (213) and obtain

$$\frac{\partial f_t^p (I_t)}{\partial I_t} = \left. \frac{\partial \mathcal{L}^p}{\partial I_t} \right|_{\omega_t^* (I_t)} = \beta E_t \left[g_t^{-\gamma} \frac{\partial f_{t+1} (v_{t+1})}{\partial v_{t+1}} \left(\sum_{i=1}^{n} \omega_t^* (I_t) R_{i,t}^e + R^f \right) \right], \tag{217}$$

where the conditional expectation is estimated using the discretization scheme for the returns of the risky assets.

In both the portfolio optimization step and the consumption optimization step at time t, we need to evaluate the value function sensitivity $\partial f_{t+1} (v_{t+1}) / \partial v_{t+1}$. To evaluate this sensitivity without knowing the functional form of $f_{t+1} (v_{t+1})$, we apply the envelope theorem to the Lagrangian, $\mathcal{L} (q_{t+1}, v_{t+1}) = u (q_{t+1}) + f_{t+1}^p (v_{t+1} - q_{t+1} + 1)$, and get

$$\frac{\partial f_{t+1} (v_{t+1})}{\partial v_{t+1}} = \left. \frac{\partial \mathcal{L} (q_{t+1}, v_{t+1})}{\partial v_{t+1}} \right|_{q_{t+1}^* (v_{t+1})} = \left. \frac{\partial f_{t+1}^p (I_{t+1})}{\partial I_{t+1}} \right|_{q_{t+1}^* (v_{t+1})} = q_{t+1}^{*-\gamma} (v_{t+1}). \tag{218}$$

Thus, due to the form of the Lagrangian, the value function sensitivity of the problem in Eq. (207) is completely specified by the optimal consumption as

$$\frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} = \begin{cases} q_{t+1}^{*-\gamma} (v_{t+1}) & \text{if } t < T-1 \\ \phi_\tau (v_T + 1)^{-\gamma} & \text{if } t = T-1 \end{cases} \tag{219}$$

To evaluate the value function sensitivity at values of v between grid points, we linearly interpolate the optimal consumption results on grid points.

References

Brandt, M., Goyal, A., Santa-Clara, P., Stroud, J., 2005. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. Review of Financial Studies 18, 831–873.

Carroll, C. D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics Letters 91, 312–320.

Cuoco, D., 1997. Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. Journal of Economic Theory 72, 33–73.

Duffie, D., Fleming, W., Soner, H. M., Zariphopoulou, T., 1997. Hedging in incomplete markets with hara utility. Journal of Economic Dynamics and Control 21, 753–782.

He, H., 1990. Convergence from discrete- to continuous-time contingent claims prices. Review of Financial Studies 3, 523–546.

He, H., Pagès, H. F., 1993. Labor income, borrowing constraints and equilibrium asset prices. Economic Theory 3, 663–696.

Ingersoll, J. E., 1987. Theory of Financial Decision Making. Rowman & Littlefield Publishers, Lanham, Maryland.

Ivković, Z., Sialm, C., Weisbenner, S., 2008. Portfolio concentration and the performance of individual investors. Journal of Financial and Quantitative Analysis 43, 613–656.

Levin, J., 2006. Supermodular games. Stanford University. Available at `http://www.stanford.edu/%7Ejdlevin/Econ%20286/Supermodular%20Games.pdf`.

Topkis, D., 1998. Supermodularity and Complementarity. Princeton University Press, Princeton, New Jersey.

Yang, C., 2010. Functional Approximation Methods for Solving Stochastic Control Problems in Finance. Ph.D. thesis, University of Texas at Austin.