

This file was downloaded from BI Open Archive, the institutional repository (open access) at BI Norwegian Business School <http://brage.bibsys.no/bi>.

It contains the accepted and peer reviewed manuscript to the article cited below. It may contain minor differences from the journal's pdf version.

Roche, H., Tompaidis, S., & Yang, C. (2013). Why does junior put all his eggs in one basket? A potential rational explanation for holding concentrated portfolios. *Journal of Financial Economics*, 109(3), 775-796

Doi: <https://doi.org/10.1016/j.jfineco.2013.03.016>

Copyright policy of Elsevier, the publisher of this journal.

The author retains the right to post the accepted author manuscript on open web sites operated by author or author's institution for scholarly purposes, with an embargo period of 0-36 months after first view online.

<http://www.elsevier.com/journal-authors/sharing-your-article#>



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

Hervé Roche

Orfalea College of Business, California Polytechnic State University

Stathis Tompaidis

McCombs School of Business, University of Texas at Austin

Chunyu Yang

BI Norwegian Business School

Abstract

Empirical studies of household portfolios show that young households, with little financial wealth, hold underdiversified portfolios that are concentrated in a small number of assets, a fact often attributed to behavioral biases. We present a potential rational alternative: we show that investors with little financial wealth, who receive labor income, rationally limit the number of assets they invest in when faced with financial constraints such as margin requirements and restrictions on borrowing. We provide theoretical and numerical support for our results and identify the ratio of financial wealth to labor income as a useful control variable for household portfolio studies.

JEL classification: D81, D83, E21, G11.

Keywords: Asset Selection, Underdiversification, Labor Income, Financial Constraints, Household Portfolios.

We thank Carolina Fugazza, Michael Gallmeyer, Lorenzo Garlappi, Valery Polkovnichenko, Paul Willen, Robertson Williams III, and seminar participants at the 2006 North American Summer Meetings of the Econometric Society in Minneapolis, the sixth Annual UT - ITAM Economics Conference, the University of Texas at Austin, the University of British Columbia, Princeton University, the Fields Institute at the University of Toronto, the 2010 American Finance Association meeting in Atlanta, the Households, Risk and Insurance Workshop at Institut Louis Bachelier in Paris, the 2010 French Finance Association meetings in Paris, and an anonymous referee for helpful comments and suggestions.

Corresponding author: Stathis Tompaidis, IROM and Finance Departments, University of Texas at Austin, 1 University Station, B6500, Austin, TX 78712. Tel. 512-4715252, Fax 512-4710587, email: Stathis.Tompaidis@mcombs.utexas.edu

1. Introduction

Portfolio choice has been a topic studied extensively in the literature. Starting at least as early as Merton (1971) and Cass and Stiglitz (1970), theory suggests that the equity part of any investor's portfolio should include all the risky assets available held in the same proportions, with the mix between bonds and stocks determined by individual risk aversion. The prescription, often called *mutual fund separation theorem*, has partly been the reason for the explosive growth in the size of mutual funds that track the market portfolio over the last 40 years.

More recently, researchers have been able to empirically study household portfolios and have found deviations from the theoretical prescription: household portfolios are underdiversified and concentrated on a small number of stocks. As a sample of the empirical literature, Kelly (1995) studies the 1983 Survey of Consumer Finances and finds that diversification increases with portfolio size, investor age, and investor wealth. Polkovnichenko (2005) uses the 1983, 1989, 1992, 1995, 1998 and 2001 Survey of Consumer Finances and confirms that wealthier households hold more diversified portfolios, even though not all wealthy households are well diversified. He argues that investors are aware of the higher risk associated with undiversified portfolios and proposes preferences with rank dependency as a potential explanation. Ivković, Sialm, and Weisbenner (2008) use data from trades and monthly portfolio positions of retail investors at a large U.S. discount brokerage house for the 1991-1996 period and show that the number of stocks in the portfolio increases with the size of the account balance, and that concentrated portfolios have higher levels of risk and return and lower Sharpe ratios than diversified portfolios. Goetzmann and Kumar (2008) study the same data set and find that diversification increases with age and income, while households with only a retirement account hold less diversified portfolios than households with additional non-retirement investment accounts. They examine several potential explanations for the lack of diversification: small portfolio size and transaction costs; search and learning costs; investor demographics and financial sophistication; layered portfolio structure; preference to higher order moments; and behavioral biases such as illusion of control, investor over-confidence, local bias and trend-following behavior. Kumar (2009), using the same data set, finds that young investors have a strong preference for riskier stocks, and argues that the young are more likely to be heavy lottery players, and this

is reflected in their selection of stocks. Mitton and Vorking (2007), using a data set of 60,000 individual accounts find that investors hold underdiversified portfolios with positively skewed returns, a fact they attribute to heterogeneous preferences for skewness. Calvet, Campbell, and Sodini (2008) study a data set of the portfolios of the entire Swedish population and propose several measures to quantify the underdiversification of household portfolios. They show that increasing age, wealth and financial sophistication increase diversification, but also lead to investors taking more aggressive positions.

While behavioral biases have been considered the cause of the discrepancy between the theory and many of the empirical observations, especially for younger investors, in this paper we offer a potential rational alternative explanation. We extend the theoretical literature by considering an investor with constant relative risk aversion preferences who is able to invest in multiple risky assets and who receives labor income with a stochastic growth rate, and faces financial constraints in the form of margin requirements and borrowing restrictions: investment needs and margin requirements can be satisfied only out of the current financial wealth of the investor, effectively rendering future earnings nontradable.¹ Our main theoretical result is that investors facing binding financial constraints do not follow the theoretical prescription described earlier: rather than holding a diversified equity portfolio they optimally choose to concentrate their portfolio in a few assets. The extent to which investors limit their investments is captured by the ratio of their financial wealth to their income. We show that once the ratio of financial wealth to income drops below a threshold and the financial constraints bind, a sequence of thresholds follows, with the investor holding different combinations of the risky assets between thresholds. Progressively, as the ratio of financial wealth to income decreases, the investor tends to concentrate his portfolio into fewer assets. In the limit, when the financial wealth is negligible compared to human capital, the investor optimally holds a single risky asset, whose choice is based on the asset's leveraged expected return and its covariance with labor income.

While many types of financial frictions can rationally lead to portfolio selection and equity portfolios that appear underdiversified, it is the combination of financial constraints and labor income that is critical in generating dynamics that match the empir-

¹Having such restrictions impacts the investor's choices significantly. The restrictions can be attributed to adverse selection and moral hazard problems, as well as the inalienability of human capital. The cause for the constraints is beyond the scope of this paper and we will consider the restrictions as given.

ical observations. Expressed over the lifetime of an investor who starts out with little financial wealth, our results suggest that the investor's portfolio rationally includes only a few assets and appears underdiversified when the investor is young. As the investor ages, his financial wealth increases and his remaining labor income decreases, and progressively more assets are rationally incorporated in the portfolio. Close to retirement, when the accumulated financial wealth is large and remaining labor income relatively small, the financial constraints no longer bind and the investor's portfolio includes all the risky assets and appears diversified.

Our results can be intuitively understood as the combination of two effects whose relative contribution varies over the investor's lifetime: the increased demand for equity exposure when labor income is large compared to financial wealth; and the limited ability to satisfy this demand because of the margin requirements and borrowing constraint. Although our results are derived for the general case of stochastic labor income, it is simplest to explain the intuition in the case of an investor with deterministic labor income. In this case labor income can be thought of as a fixed investment in a risk-free bond. As shown in Merton (1971), without a margin requirement, a borrowing constraint, and other frictions, to find the optimal allocation in risky assets the investor should discount his lifetime income at the risk free rate, add it to his financial wealth, and choose an equity allocation based on the sum. Keeping financial wealth constant and increasing labor income implies increasing equity investment when measured as a fraction of financial wealth. In the absence of the margin requirements and the borrowing constraint the investor would borrow against his future income and increase expected return by leveraging his portfolio while keeping it diversified. Since margin requirements limit the extent that the portfolio can be leveraged, the only possible way to increase expected return is through shifting portfolio holdings towards assets with higher expected returns, sacrificing diversification. We show that as the demand for additional equity exposure increases; i.e., when the financial wealth to income ratio decreases, the demand for higher expected returns prevails. When labor income is deterministic, in the limit of zero financial wealth to income ratio, the investor holds a single risky asset based entirely on the asset's expected return. Since the demand for financial leverage is highest when the investor's financial wealth is smallest and the investor's lifetime labor income is greatest, young investors are most likely to sacrifice diversification and hold

concentrated portfolios. The demand for leverage eases as the investor ages and leads to more diversified portfolio choices.

An alternative interpretation of our selection result, where the investor trades diversification for higher expected returns, is that the investor can be thought of as becoming less risk averse the more the constraints bind. We show that in the limit when the ratio of his financial wealth to income ratio tends to zero the investor acts as if he were risk-neutral. Additionally, investor choices can be understood through adjusted asset characteristics: in the case of deterministic income, using duality, we show that the behavior of an investor that faces financial constraints corresponds to the behavior of an unconstrained investor whose opportunity set includes assets with adjusted Sharpe ratios. As the constraint binds, the adjusted Sharpe ratios of the assets decrease; at the point where an asset is dropped from the portfolio, its adjusted Sharpe ratio drops to zero.

We point out that while an investor whose financial wealth is a progressively smaller part of his lifetime labor income tends to choose progressively more concentrated portfolios, his overall risk, measured by the variance of returns of his overall portfolio, does not necessarily increase. Since lower values of the financial wealth to income ratio raise demand for equity, the unconstrained investor would have chosen a leveraged portfolio, trading additional variance for additional expected return. Unable to satisfy his risk appetite through leverage, the investor chooses a portfolio that achieves higher expected returns by concentrating his portfolio to stocks with high expected returns. Whether the variance of the returns of the overall portfolio of the constrained investor is greater or smaller than that of the portfolio of the unconstrained investor hinges on the balance of the two effects.

In addition to changes in the asset allocation, we show that the financial constraints induce changes in the investor's consumption behavior. In the case of an infinitely lived investor that receives an uninterrupted income stream, when the wealth of the investor tends to zero the investor's consumption rate tends to his income rate, preventing wealth accumulation. While this result depends on the assumption of infinite horizon, it does suggest that investors with relatively long horizons, little wealth, and large income, have little incentive to save, a result that we show in our numerical study. Another effect of the margin requirements and borrowing constraint is that investor consumption decreases, and, when income is deterministic, the volatility of the investor's consumption is lower

than the volatility of consumption for a similar investor that does not face financial constraints. The intuition behind this result is that the constrained investor leverages his portfolio less and therefore his overall wealth is less influenced by changes in the prices of the risky assets.

While the base case for our theoretical results is that of an infinitely lived investor with constant relative risk aversion preferences, we show that our results are robust. We show that portfolio concentration and underdiversification persist in the case of an investor with finite horizon; and also when the investor has general preferences and receives deterministic income.

We also consider special cases of our results that provide intuition on our asset selection results and shed light on the rich set of possibilities. Although our framework is a partial equilibrium one, it corresponds to a single factor Capital Asset Pricing Model (CAPM), where an unconstrained investor would only hold the market portfolio and the riskless asset. The behavior of the constrained investor is much more complicated. The simplest case is when asset returns are independent and uncorrelated with labor income growth: in that case a strict order exists and assets drop out of the investor's portfolio based on their leveraged excess return. A strict order also exists under assumptions other than independence, for example when the correlation among all asset pairs is the same and shorting or borrowing are not allowed. In the general case, when shorting is allowed and asset returns may be negatively correlated, while on average the portfolio becomes less diversified as the ratio of the investor's financial wealth to his income decreases, the behavior is not necessarily monotone and assets can both drop out and reenter the portfolio. We also show that, perhaps surprisingly, it is possible to have asset selection among assets with the same expected returns — and betas with the market portfolio — where the asset with the lowest idiosyncratic volatility is dropped from the portfolio; and also to have asset selection among assets with the same idiosyncratic volatility and different expected returns, where for a certain range of the ratio of the investor's financial wealth to income, the asset with the highest expected return drops out of the portfolio while assets with lower expected returns are still held. We also consider a case where, in addition to individual risky assets, the investor has access to an index fund. When the investor faces financial constraints asset selection occurs in this case as well, and the index fund is among the first assets to drop out of the investor's portfolio.

To the best of our knowledge, our paper is the first to link the combination of labor income and financial constraints to underdiversification of investor portfolios. Literature related to our paper includes the papers by Karatzas, Lehoczky, and Shreve (1987), and Cox and Huang (1989) who introduce martingale techniques that make it easier to deal with constraints on the investment strategies. Models with constraints on the portfolio policies are studied by Karatzas, Lehoczky, Shreve, and Xu (1991), Cvitanić and Karatzas (1992), Cvitanić and Karatzas (1993), He and Pearson (1991), Xu and Shreve (1992) and Tepla (2000). Cuoco (1997) is able to demonstrate existence of optimal strategies for the case of an investor that faces margin constraints and receives income but does not provide a characterization of the strategies. Koo (1998) and Koo (1999) solve the optimal investment and consumption problem for an investor that receives labor income and faces a short-sale constraint and describes properties of the optimal consumption plan, but does not discuss underdiversification. Cuoco and Liu (2000) discuss the case of an investor that is facing margin requirements but does not receive income, and provide a characterization of his optimal investment strategy. He and Pagès (1993), El Karoui and Jeanblanc-Picqué (1998) and Duffie, Fleming, Soner, and Zariphopoulou (1997) study the optimal asset selection problem of an investor who receives income and who is constrained to maintain nonnegative levels of current wealth, but do not address margin requirements. Cuoco and Liu (2004) find underdiversification in a study of the impact of VaR reporting rules in the portfolio choice of a financial institution that maximizes utility from terminal wealth. While very different, the setting in Cuoco and Liu (2004) can be shown to be a special case of our study, corresponding to an infinitely lived investor with constant relative risk aversion preferences that receives deterministic income.² Dybvig and Liu (2010) study the lifetime asset allocation problem with voluntary retirement, for the case of an investor that receives labor income and who is constrained from borrowing, but who can only invest in a single risky asset. Underdiversification results similar to the ones we obtain in this paper are also obtained independently and contemporaneously in the paper by Liu (2010). Liu (2010), describes a model where investors engage in asset selection due to the combination of a desire to guarantee a minimum level of wealth and constraints on their ability to borrow and to short-sell risky assets. The main difference

²Cuoco and Liu (2004) do not study the problem of optimal consumption and asset allocation for an investor that faces financial constraints and receives labor income — rather they consider the problem of asset allocation for a financial institution that faces a VaR constraint.

between the paper by Liu (2010) and our paper is that we consider a dynamic framework that allows us to quantify the degree of underdiversification in investor portfolios through their lifetime, as well as the impact of approaching retirement to the degree of portfolio diversification, while the framework in Liu (2010) is static and does not offer any prediction regarding the age of an investor and the degree of portfolio underdiversification.³ The implication from our paper for empirical studies is that wealth over lifetime income is a potential explanatory variable for observed underdiversification.

To quantify the magnitude of our theoretical results, we also present numerical results for the case of an investor that receives income until age 65 and then retires with an expected remaining lifetime of 20 years. The investor has access to five risky assets, calibrated to match the risk-return characteristics of five industry portfolios based on data from 1927-2004. Our calibration implies that the expected return of the risky assets is proportional to the assets' covariance with the market portfolio, in line with the CAPM. To solve the problem, we employ a new numerical algorithm, originally introduced in Yang (2010), which is an extension of the algorithm developed by Brandt, Goyal, Santa-Clara, and Stroud (2005).⁴ The algorithm determines optimal asset allocations by solving the first order Karush-Kuhn-Tucker conditions using functional approximation of conditional expectations and projection of the value function on a set of radial basis functions to address the curse of dimensionality problem when facing a large number of state and choice variables. The extension allows for a more accurate estimation of conditional expectations by limiting the region where test solutions are generated iteratively, a process called "Test Region Iterative Contraction (TRIC)". Our numerical results are in line with the theoretical intuition: young investors hold concentrated portfolios, engage in asset selection, and save a smaller fraction of their income compared to older investors, who hold portfolios that are close to diversified. It is interesting to note that our results indicate that when investors are severely constrained they only choose high-tech stocks for their equity portfolio, increasing the expected return

³An additional difference is that Liu (2010) attributes the underdiversification to solvency constraints, while we show that it is margin requirements that are fundamental. Indeed, imposing a solvency constraint without a margin requirement does not lead to underdiversified portfolios as shown, for example, in He and Pagès (1993), El Karoui and Jeanblanc-Picqué (1998) and Duffie et al. (1997). Beyond the differences in the setup, our results also differ from the results in Liu (2010) in that we show that the portfolios held by the investor when the constraint binds are a complicated sequence depending on asset and labor income characteristics, rather than the simple sequence obtained in Liu (2010).

⁴See also Carroll (2006), and Garlappi and Skoulakis (2010).

of their portfolio but lowering its Sharpe ratio. We provide values for the underdiversification measures developed by Calvet et al. (2008) and show that, in line with our theoretical result, the investor's effective risk aversion tends to zero when the value of the financial wealth to income ratio tends to zero.

The remainder of the paper is organized as follows: in Section 2 we present our theoretical model and results. Section 3 discusses the numerical algorithm used to solve the finite horizon problem, and presents the numerical results for the optimal allocations and diversification measures for our calibrated model of five industry indexes and associated comparative statics. Section 4 concludes. The proofs and a detailed description of the numerical algorithm are contained in the Appendix.

2. Theoretical Analysis

For our theoretical results, we consider a continuous time economic setting with an infinitely lived investor who derives utility from consumption and who is able to invest in a riskless money-market account and N risky securities that evolve according to geometric Brownian motion with constant coefficients. We have chosen geometric Brownian motion for tractability reasons. We assume that the investor's utility is of the constant relative risk aversion (CRRA) type. Our results also hold for general preferences when the labor income received by the investor is deterministic.

2.1. *The Economic Setting*

The Financial Market and the Labor Income process.

Uncertainty is modeled by a probability space (Ω, \mathcal{F}, P) on which an $N + M$ dimensional, standard, Brownian motion $(w, w^Y) = ((w_1, w_2, \dots, w_N), (w_1^Y, w_2^Y, \dots, w_M^Y))$ is defined. A state of nature ω is an element of Ω . \mathcal{F} denotes the tribe of subsets of Ω that are events over which the probability measure P is assigned. At time t , the investor's information set is \mathcal{F}_t , where \mathcal{F}_t is the σ -algebra generated by the observations of (w, w^Y) , $\{(w_s, w_s^Y); 0 \leq s \leq t\}$. The filtration $\mathbb{F} = \{\mathcal{F}_t, t \in \mathbb{R}_+\}$ is the information structure and satisfies the usual conditions (increasing, right-continuous). In our setting

only a money-market account that pays a constant interest rate and N risky securities are available. The additional M factors are associated with the growth rate of the labor income. The market is incomplete: to represent the risk present in the economy an additional M securities would be required.

The value of the money-market account, B evolves according to

$$dB_t = rB_t dt, \quad (1)$$

where r is the constant interest rate. Let $S = (S_1, S_2, \dots, S_N)$ be the vector stock price process whose dynamics are given by

$$dS_t = I_{S_t} \mu dt + I_{S_t} \sigma dw_t, \quad (2)$$

where I_S is a diagonal $N \times N$ matrix with diagonal elements S , μ is an $N \times 1$ matrix, $\sigma = [\sigma_{ij}]$ is an $N \times N$ matrix and dw_t is the increment of the N dimensional Wiener process w , with $[dw_{it}, dw_{jt}] = \delta_{ij}$ where $\delta_{ij} = 0, i \neq j$ and $\delta_{ii} = 1$. The instantaneous covariance matrix $\sigma \sigma^\top$ is assumed to be nonsingular.

We assume that the investor receives a nonnegative income stream at a rate Y_t

$$dY_t = Y_t (m dt + \Sigma^\top dw_t + \Theta^\top dw_t^Y), \quad (3)$$

where m is the growth rate of income, $\Sigma^\top = (\Sigma_1, \Sigma_2, \dots, \Sigma_N) \in \mathbb{R}^N$, $\Theta^\top = (\Theta_1, \Theta_2, \dots, \Theta_M) \in \mathbb{R}^M$, and dw_t^Y is the increment of the M -dimensional Wiener process w^Y , with $[dw_t, dw_t^Y] = 0$. All the coefficients are assumed to be constant.

Trading Strategies and Margin Requirements.

We assume that consumption c and trading strategies (x, z) are adapted processes to the filtration \mathbb{F} , where x is the dollar amount invested in the money-market account and $z^\top = (z_1, z_2, \dots, z_N)$, are the dollar amounts invested in the N risky assets.

To trade in risky assets, U.S. investors must hold sufficient wealth in a margin account. This wealth can be held in securities or cash. The Federal Reserve Board's Regulation T sets the initial margin requirement for stock positions undertaken through

brokers. The values for the initial margin requirement are 50 percent for a long equity position, and 150 percent for a short equity position.⁵

For our model, we impose the following margin constraint on an investor that holds $z_i, i = 1, \dots, N$ dollar amounts in the risky assets

$$(\lambda^+)^T z^+ + (\lambda^-)^T z^- \leq W, \quad (4)$$

with $\lambda_i^+ = 1 - \kappa_i^+, \lambda_i^- = \kappa_i^- - 1$ with $\kappa_i^+ \leq 1 \leq \kappa_i^-$, for $i = 1, \dots, N$, where, for any real number x , we have $x = x^+ - x^-$, with $x^+ = \max(x, 0)$ and $x^- = \max(-x, 0)$. The regulation T initial margin requirements correspond to $\kappa_i^+ = 0.5$ and $\kappa_i^- = 1.5$, for $i = 1, \dots, N$.

We define the set of all possible margin coefficients, Λ :

$$\Lambda = \{\lambda \in \mathbb{R}^N, \lambda_i \in \{\lambda_i^+, -\lambda_i^-\}, i = 1, \dots, N\}, \quad (5)$$

and the set of all feasible allocations in the risky assets, Q :

$$Q = \{x \in \mathbb{R}^N, \lambda^T x \leq 1, \lambda \in \Lambda\}. \quad (6)$$

Q is a convex set prescribed by 2^N linear constraints of which at most N are binding at the same time. Risky investment z satisfies the margin constraint in Eq. (4) if and only if z/W is in Q . We note that the margin constraint is more stringent than the constraint of nonnegative wealth $W \geq 0$.

Preferences.

⁵See Fortune (2000) as well as the Federal Reserve Board's Regulation T for institutional details. In addition to initial margin requirements, there are also maintenance margin requirements that correspond to the level in the margin account at which collateral needs to be added to the account to avoid liquidation of the position. Including a maintenance margin would make the problem path dependent and we do not consider it in this paper. Further complications regarding margin accounts include the fact that 102% of the collateral held in the margin account needs to be held in liquid assets, and that the collateral does not, in general, earn the risk free rate of interest, see Geczy, Musto, and Reed (2002). While it would be interesting to consider these additional features, our main results regarding asset selection and underdiversification of an investor's portfolio when the investor's labor income is large compared to his financial wealth, would not change.

There is a single perishable good available for consumption, the numéraire. Preferences are represented by a time additive utility function

$$U(c) = E \left[\int_0^\infty u(c_t) e^{-\theta t} dt \right], \quad (7)$$

where the time discount factor, θ , is constant. The utility function u is of the CRRA type, with risk aversion coefficient γ

$$u(c) = \begin{cases} \frac{c^{1-\gamma}}{1-\gamma}, & \gamma \neq 1 \\ \ln c, & \gamma = 1 \end{cases} \quad (8)$$

Optimization Problem. The investor's problem is to maximize his expected, cumulative, discounted utility of consumption

$$F(W_t, Y_t) = \max_{(c, \frac{z}{W} \in Q)} E_t \left[\int_t^\infty u(c_s) e^{-\theta(s-t)} ds \right], \quad (9)$$

under the budget constraint

$$dW_s = (rW_s - c_s + Y_s + z_s^\top(\mu - r\bar{1})) ds + \sigma z_s^\top dw_s, \quad (10)$$

when labor income follows

$$dY_s = Y_s (m ds + \Sigma^\top dw_s + \Theta^\top dw_s^Y), \quad (11)$$

where $W_t > 0, Y_t > 0$ are the initial conditions for the investor's wealth and income rate.

Transversality Condition.

The transversality condition for this problem is given by

$$\lim_{T \rightarrow \infty} E_t [F(W_{t+T}, Y_{t+T}) e^{-\theta(T+t)}] = 0. \quad (12)$$

Properties of the Primal Value Function F .

The value function F satisfies the following propositions whose proof can be found in Sections A, and B of the Appendix.

Proposition 1. F is homogeneous of degree $1 - \gamma$ in (W, Y) .

Proposition 2. *F is nondecreasing in W and Y and jointly concave in (W, Y) .*

The homogeneity of F allows us to rewrite the value function in terms of the ratio of wealth over income, as

$$F(W, Y) = Y^{1-\gamma} f\left(\frac{W}{Y}\right), \quad (13)$$

for a nondecreasing smooth function f . We will refer to f as the reduced value function. Below, we denote by $v = W/Y$ the ratio of wealth over income. Notice that the concavity of F in (W, Y) implies that f is concave in v .

Conditions on Parameters.

To get a well defined problem, we impose that the following parameters A, B, C are positive. These parameters appear in the characterization of the value function, the optimal consumption plan and the optimal investment strategy in both the optimization problem of an unconstrained as well as that of a constrained investor

$$\begin{aligned} A^{-1} &= \frac{\theta}{\gamma} + \frac{\gamma - 1}{\gamma} \left(r + \frac{1}{2\gamma} (\mu - r\bar{1})^\top (\sigma\sigma^\top)^{-1} (\mu - r\bar{1}) \right) > 0 \\ B^{-1} &= r - m + (\sigma\Sigma)^\top (\sigma\sigma^\top)^{-1} (\mu - r\bar{1}) > 0 \\ C^{-1} &= \theta + (\gamma - 1) \left(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} \right) > 0. \end{aligned} \quad (14)$$

Under the financial market described above Duffie et al. (1997) study the Merton problem for a HARA preference investor who receives labor income that follows geometric Brownian motion that is not perfectly correlated with the returns of the risky assets under the constraint of nonnegative wealth. A formal analysis of the existence and uniqueness of the solution of this problem under margin requirements requires the use of viscosity solutions as in Duffie et al. (1997) and is beyond the scope of our paper. However, a heuristic derivation of the Hamilton-Jacobi-Bellman equation can provide some insight regarding portfolio selection.

$$\begin{aligned}
\left(\theta + (\gamma - 1)\left(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2}\right)\right) f(v) = \max_{\omega \in Q} & \frac{\gamma (f'(v))^{\frac{\gamma-1}{\gamma}}}{1 - \gamma} + f'(v) \\
& + (r - m + \gamma (\Sigma^\top \Sigma + \Theta^\top \Theta)) v f'(v) \\
& + \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} v^2 f''(v) \\
& + \omega^\top ((\mu - r\bar{1} - \gamma \sigma \Sigma) v f'(v) - \sigma \Sigma v^2 f''(v)) \\
& + v^2 f''(v) \omega^\top \frac{\sigma \sigma^\top}{2} \omega,
\end{aligned} \tag{15}$$

where $\omega = z/W$ is the vector of the percentage of wealth invested in each of the risky assets. This problem is equivalent to

$$\max_{\omega \in Q} \left[\omega^\top (\mu - r\bar{1} + (y - \gamma) \sigma \Sigma) - \frac{y}{2} \omega^\top \sigma \sigma^\top \omega \right], \tag{16}$$

where y is the investor's lifetime relative risk aversion,

$$y = -\frac{W F_{11}}{F_1}. \tag{17}$$

As shown in Section L of the Appendix, the boundary condition at $v = W/Y = 0$, is given by

$$\frac{1}{f(0)} = (1 - \gamma) \left(\theta + (\gamma - 1) \left(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} \right) \right). \tag{18}$$

2.2. Benchmark Case: No Margin Requirements, Labor Income Spanned by the Risky Assets

Following Merton (1971), when the income growth is spanned by the returns of the risky assets, $\Theta = 0$, and the investor does not face margin requirements, markets are complete. The optimal asset allocations z^f , and optimal consumption c^f are given by

$$\begin{aligned}
z^f &= \frac{(\sigma \sigma^\top)^{-1} (\mu - r\bar{1})}{\gamma} W - B \frac{(\sigma \sigma^\top)^{-1} \eta}{\gamma} Y \\
c^f &= \frac{W + BY}{A},
\end{aligned} \tag{19}$$

where

$$\eta = \mu - r\bar{1} - \gamma\sigma\Sigma, \quad (20)$$

is the vector of expected excess returns adjusted for labor income correlation with the risky assets.

2.3. Case with Margin Requirements

Cuoco and Liu (2000) characterize the optimal consumption and portfolio choices for an investor that is subject to margin requirements and that does not receive income; i.e., $Y \equiv 0$. The case of an investor that receives an income stream and is subject to margin requirements is considerably more complicated than the case without income. Intuitively, from the work of Merton (1971), we know that when income growth is spanned by the returns of the risky assets and without margin requirements the investor should discount his future earnings using the market discount factor, add the discounted value to his current wealth, and make an investment choice based on the sum, provided he can borrow against his future earnings. Since discounted future earnings may be a significant portion of the sum, and possibly many times the current wealth, the allocation may violate the margin constraint. The extent to which the margin constraint binds depends on the ratio between the current wealth and the discounted value of future earnings.

Before addressing the general case, we first consider the case where the adjusted excess return vector η , defined in Eq. (20), is identically equal to zero. In this case, we show in Section E of the Appendix that labor income has no impact on portfolio holdings and the fraction of wealth invested in each asset is constant. If $(\sigma\sigma^\top)^{-1}\sigma\Sigma \in Q$, the margin constraint is never binding and $z^*/W = (\sigma\sigma^\top)^{-1}\sigma\Sigma$. Otherwise, the margin constraint is always binding. Depending on the parameters of the model, the number of assets in the portfolio can range from N (full diversification) to 1 (full selection). A condition under which exactly K assets are optimally held in the portfolio is reported in Section E of the Appendix.

In the remainder of the paper we assume that the adjusted excess return vector is not equal to zero, $\eta \neq 0$. To study the impact of the income stream on the asset allocations, we choose parameters such that the margin requirement is not binding for

low income levels. Our motivation behind this assumption is that it generates dynamics for the investor's asset allocation that are richer compared to the case when the margin requirement binds for low, or zero, levels of income. This assumption is satisfied if the following inequalities hold

$$\begin{aligned} \max_{\lambda \in \Lambda} \left(\frac{\lambda^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1})}{\gamma} \right) &< 1, \\ \max_{\lambda \in \Lambda} \left(\lambda^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K \sigma \Sigma \right) &< 1, \end{aligned} \tag{21}$$

where, for every combination of $K \leq N$ assets among all N available risky assets, we denote by J_K 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.

2.3.1. Positive correlation, no shorting or borrowing

A setting of practical interest is when the investor's opportunity set includes assets whose returns are positively correlated and when the investor is unable to short an asset or borrow against his future income. We are able to show, under an additional assumption, that the investor's portfolio exhibits monotonic asset selection, with assets dropping out of the portfolio at thresholds in the investor's financial wealth to income ratio.

Proposition 3. *Assume that labor income growth is uncorrelated with the returns of the risky assets, i.e., $\eta = \mu - r\bar{1}$, that $(\sigma \sigma^\top)^{-1} \eta \in \mathbb{R}_+^N$, and that the off-diagonal elements of the inverse covariance matrix, $(\sigma \sigma^\top)^{-1}$ are non-positive. Then, if shorting and borrowing are not allowed, there exist $N + 1$ regions in the investor's risky portfolio composition, separated by N thresholds, in the values of the ratio of financial wealth to income. For large values of the wealth to income ratio the investor holds all the risky assets. As the ratio decreases, the margin requirement binds when the first threshold is crossed and, as each new threshold is crossed, an asset is dropped from the portfolio until only the risky asset with the highest expected return remains in the portfolio.*

The proof of Proposition 3 is provided in Appendix C. We note that the assumption that $(\sigma \sigma^\top)^{-1} \eta \in \mathbb{R}_+^N$ ensures that all assets are held long in the portfolio when the margin

requirement is not binding, or, equivalently, that all the market portfolio weights are positive. In addition, assuming that $(\sigma\sigma^\top)^{-1}\eta \in \mathbb{R}_+^N$ implies that the risky assets are substitute goods. In Appendix C we show that the assumption that all the off-diagonal entries of the inverse covariance matrix $(\sigma\sigma^\top)^{-1}$ are non-positive implies that all the entries of the covariance matrix $\sigma\sigma^\top$ 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 $\rho \geq 0$.

Proposition 3 indicates that the investor engages in asset substitution as the margin constraint becomes binding. Intuitively, the investor tries to improve his return, within the bounds of the shorting and borrowing constraints imposed on him, and in doing so shifts his portfolio composition toward fewer assets. In the limit, when the investor's current wealth is negligible compared to his future earnings, the investor acts as if he is risk-neutral: he holds a single risky asset to the maximum extent allowed by the constraints, chosen based on the asset's expected excess return.

2.3.2. Special Case: Risky Assets with Independent Returns and Uncorrelated Labor Income Growth

In the special case where the returns of the risky assets are independent and the income growth is uncorrelated with the asset returns, we are able to obtain a result stronger than the result in Proposition 3, and completely characterize the thresholds at which the risky assets are dropped from the investor's portfolio, even when shorting and borrowing are allowed.

Proposition 4. *Given the evolution of the price of the risky assets, money-market account, and income, described by Eqs. (1), (2), and (3), and under the assumptions on the parameters given by Eq. (14), and assuming that the vector of adjusted excess returns, η , is not equal to zero, and that returns of the risky assets are independent and the growth of the labor income is uncorrelated with the returns of the risky assets, $\Sigma = 0$, we have that*

(i) The excess return $\mu_k - r$ of risky asset k , and the corresponding margin requirement coefficient λ_k^* have the same sign, and risky securities can be ranked according to their leveraged excess expected return

$$0 < \frac{\mu_N - r}{\lambda_N^*} < \frac{\mu_{N-1} - r}{\lambda_{N-1}^*} < \dots < \frac{\mu_1 - r}{\lambda_1^*}. \quad (22)$$

(ii) The optimal asset allocation can be described by $N + 1$ distinct regions defined by the values of the lifetime relative risk aversion y , or, equivalently the values of the current wealth to income ratio v . These regions are characterized in terms of the thresholds $0 < y_{2,D}^* < \dots < y_{N,D}^* < y_B^*$, with

$$\begin{aligned} y_B^* &= \alpha^\top \xi \\ y_{N,D}^* &= \alpha^\top (\xi - \xi_N \bar{1}) \\ y_{K,D}^* &= (I_K \alpha)^\top [I_K (\xi - \xi_K \bar{1})], \quad K = 2, \dots, N - 1 \end{aligned} \quad (23)$$

where λ^* , the vector of margin coefficients when the constraint binds, $\xi_k = (\mu_k - r) / \lambda_k^*$, and $\alpha_k = (\lambda_k^* / \sigma_k)^2$. As the lifetime relative risk aversion crosses threshold $y_{i,D}^*$, $i = 2, \dots, N$, asset i is dropped from the investor's portfolio. Between thresholds $y_{i,D}^*$ and $y_{i+1,D}^*$ the portfolio contains assets $1, \dots, i$.

(iii) The optimal asset allocations are given by

$$\begin{aligned} \omega_k^* &= \frac{\mu_k - r}{y \sigma_k^2}, \quad \text{for } k = 1, \dots, N, \quad y > y_B^* \\ \omega_k^* &= \frac{\alpha_k [y - y_{K,D}^*]^+}{y \lambda_k^* (I_K \alpha)^\top \bar{1}}, \quad \text{for } k = 1, \dots, N, \quad y_{K,D}^* < y < y_{K+1,D}^* \end{aligned} \quad (24)$$

where, by convention, we assume that $y_{N+1,D}^* = y_B^*$.

The proof of the proposition is provided in Section D of the Appendix. The intuition behind Proposition 4 is illustrated in Fig. 1. The figure considers an investor that has access to three risky assets. The parameters are chosen such that when financial wealth is very large relative to income the investor holds a portfolio that includes all three risky assets. A similar investor with relatively less financial wealth increases his allocation to the risky assets. The larger income is, relative to financial wealth, the higher a percentage of financial wealth the investor places in risky assets. As long as the constraint is not binding the investor maintains the relative proportions of the risky assets in his equity

portfolio, keeping the portfolio diversified. At some point, for an investor with low enough financial wealth relative to his income, the size of the equity portfolio is large enough for the margin constraint to bind. In the figure this happens when the chosen allocation reaches the shaded plane. For an investor with even lower financial wealth to income ratio the margin constraint restricts him in choosing portfolios on the shaded plane in Fig. 1. As long as the allocation does not reach the edge of the shaded plane the investor maintains all the risky assets in his portfolio but in proportions that vary with the ratio of financial wealth to income. Once the allocation reaches an edge of the shaded plane the investor is restricted to choose portfolios along that edge, and is no longer holding all the risky assets. Further decreases in the financial wealth to income ratio eventually lead the investor to hold an equity portfolio consisting of a single asset, represented in the figure by a vertex. Given our results, the asset eventually held is the one with the highest expected leveraged return, irrespective of the asset's volatility.

We point out that the geometrical intuition developed in Fig. 1 also applies to other problems. For example, the framework described in Liu (2010) corresponds to the margin requirements becoming more stringent as the investor approaches the subsistence level of wealth. In that case it is the shaded plane in Fig. 1 that approaches the origin, forcing the optimal asset allocation towards a concentrated portfolio, and eventually a single asset. This intuition makes clear that asset selection is a general feature of margin requirements.

Relying on duality techniques developed by Cvitanić and Karatzas (1992) and Cuoco (1997), it is possible to interpret the asset allocation and consumption problem with margin requirements as an inter-temporal consumption-investment problem for an investor facing no financial constraints by adjusting the risky assets' returns and the risk-free interest rate. This approach can be used to quantify the impact of the margin requirements: as the margin constraint becomes more stringent, adjusted Sharpe ratios for the risky assets shrink (in absolute value), which makes risky assets less attractive to the investor. We provide further details in Sections H and I of the Appendix.

2.3.3. *General Case*

When income growth and the returns of the risky assets are correlated to each other, the returns of the risky assets are negatively correlated, and shorting and borrowing are

allowed subject to a margin requirement, the general intuition of moving towards more concentrated portfolios as the ratio of financial wealth to income increases remains the same but with two caveats: a) in the limit when the ratio tends to zero, the investor holds a single asset chosen based on its leveraged return, *adjusted for correlation with income*; b) it is possible for an asset to reenter the portfolio after it has been dropped.

Proposition 5 below characterizes the optimal portfolio allocations according to the investor's lifetime relative risk aversion, y , or, equivalently, in terms of the ratio, $v = W/Y$, of current financial wealth to income. To state the proposition, we define $e_i^\top = (0, 0, \dots, 1, \dots, 0)$, where 1 is in the i th position, and I_K to be the $K \times N$ matrix that consists of the first K rows of the $N \times N$ identity matrix.

Proposition 5. *Given the evolution of the price of the risky assets, money-market account, and income, described by Eqs. (1), (2), and (3), and under the assumptions on the parameters given by Eq. (14), and assuming that the vector of adjusted excess returns, η , is not equal to zero, let*

$$y_B^* = \frac{(\lambda_B^*)^\top (\sigma \sigma^\top)^{-1} \eta}{1 - (\lambda_B^*)^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma} \quad (25)$$

where

$$\lambda_B^* = \operatorname{argmax}_{\lambda \in \Lambda} \frac{\lambda^\top (\sigma \sigma^\top)^{-1} \eta}{1 - \lambda^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma}$$

- (i) *When the current wealth to income ratio v is large enough such that the lifetime relative risk aversion, y , is greater than y_B^* , $y > y_B^*$, the margin constraint is not binding and the investor holds all N assets in his portfolio. As the ratio v decreases, y decreases and at $y = y_B^*$, the margin constraint starts binding. The margin coefficients λ_i are determined by whether the position in asset i is long or short at levels of the lifetime relative risk aversion above the level where the constraint binds.*
- (ii) *For all values of lifetime relative risk aversion y lower than y_B^* , $y < y_B^*$, the margin constraint is binding.*
- (iii) *As the lifetime relative risk aversion y decreases, the portfolio composition changes at thresholds where risky assets may drop out or reenter the portfolio. Each portfolio configuration can be encountered at most once, between two thresholds in the values of the lifetime relative risk aversion.*

(iv) For small enough values of the lifetime relative risk aversion, the investor holds only one risky asset to the maximum extent allowed by the margin. The ultimately selected asset i , satisfies

$$i = \operatorname{argmax}_{k \in (1, \dots, N)} \frac{\eta_k}{\lambda_k}, \quad (26)$$

where, given a vector x , we use the notation x_i to denote the vector's i -th component. The threshold in the lifetime relative risk aversion where the next to last asset, j , is dropped out of the portfolio is given by $y_{j,D}^*$

$$y_{j,D}^* = \min_{k \neq i}^* \{y_{k,D}^L, y_{k,D}^S\}, \quad (27)$$

where, given a set of nonnegative numbers, x_1, x_2, \dots, x_k , where at least one number is strictly positive, we define $\min^*(x_1, x_2, \dots, x_k)$ to be the smallest, nonzero, number, and where

$$y_{k,D}^L = \left[\frac{\frac{\eta_i}{\lambda_i} - \frac{\eta_k}{\lambda^+}}{\left(\frac{\sigma_i}{\lambda_i}\right)^2 - \frac{(\sigma\Sigma)_i}{\lambda_i} - \rho_{ik} \frac{\sigma_i}{\lambda_i} \frac{\sigma_k}{\lambda^+} + \frac{(\sigma\Sigma)_k}{\lambda^+}} \right]^+ \quad (28)$$

$$y_{k,D}^S = \left[\frac{\frac{\eta_i}{\lambda_i} + \frac{\eta_k}{\lambda^-}}{\left(\frac{\sigma_i}{\lambda_i}\right)^2 - \frac{(\sigma\Sigma)_i}{\lambda_i} + \rho_{ik} \frac{\sigma_i}{\lambda_i} \frac{\sigma_k}{\lambda^-} - \frac{(\sigma\Sigma)_k}{\lambda^-}} \right]^+.$$

The proof of Proposition 5 is provided in Section E of the Appendix. Proposition 5 indicates that, even with correlated asset returns and the possibility of shorting and borrowing, the investor engages in asset substitution as the margin constraint becomes binding. An alternative intuition can be described in terms of the investor's lifetime relative risk aversion y : lifetime relative risk aversion is high when income is relatively small compared to the investor's wealth, while it is low when discounted future earnings are much larger than current wealth. When lifetime relative risk aversion is low, the investor is willing to increase his exposure in risky assets, resulting in the margin constraint binding and leading the investor to hold fewer assets. Compared to the case where the inverse covariance matrix has nonpositive entries and shorting and borrowing are not allowed, in the general case assets can both drop out and reenter the investor's portfolio as the ratio of financial wealth to income declines, while, in the limit when the investor's current wealth is negligible compared to his future earnings, the investor

chooses the single risky asset to hold based not only on the asset's expected excess return, but the asset's expected excess return adjusted for covariance with the labor income $\mu_1 - r - \gamma\sigma_1\Sigma_1$. Yet, qualitatively, the behavior is similar. When the financial constraints are not binding, in both cases all assets are held. As the financial wealth to income ratio decreases the constraints bind, and a series of thresholds exists, where portfolio composition changes, with the portfolio becoming less diversified.

In Section E of the Appendix we show that when income is deterministic our results hold beyond the case of investors with CRRA preferences: asset selection occurs as the lifetime relative risk aversion varies for preferences represented by any smooth and concave utility function.⁶ Moreover, in the nonbinding region the two fund separation theorem applies with the risky assets being held in the same proportions as in the unconstrained case. However, the investor's risky asset allocation is smaller (in absolute value) than the allocation of an investor with the same wealth and income that does not face a margin requirement.

2.3.4. *Special Case: Two Risky Assets with Uncorrelated Labor Income Growth*

The case with two risky assets is simple enough to be tractable, yet rich enough to illustrate the complexity that may arise as the margin requirements and the borrowing constraint become binding.

Proposition 6. *Under the same assumptions as Proposition 5, when the number of risky assets available to the investor is two and labor income growth is uncorrelated with asset returns, and assuming that the first asset has a higher leveraged expected return compared to the second asset, the following cases are possible:*

- (i) *there is a single threshold, y_B^* , in the lifetime relative risk aversion, y , and two regions in the portfolio composition. In the first region, when y is greater than the threshold y_B^* the margin requirement is not binding, but only the first asset is*

⁶Cuoco and Liu (2004) find risk shifting behavior in the context of a financial institution that needs to follow VaR reporting rules and that tries to optimize its asset selection. This behavior leads the financial institution to invest in underdiversified portfolios as the VaR constraint becomes binding. Through a transformation, their setting can be shown to be a special case of ours, corresponding to the case of deterministic income. We note that our general result is strikingly different from the result in the special case studied by Cuoco and Liu (2004), especially regarding the possibility of asset reintegration after an asset has been dropped from the investor's portfolio.

held in the portfolio. In the second region the margin requirement binds at values of $y \leq y_B^*$, and the portfolio consists of holding the first asset at the maximum amount allowed.

- (ii) there are two thresholds, $y_B^* > y_{2,D}^*$, in the lifetime relative risk aversion and three regions in portfolio composition. In the first region, when the lifetime relative risk aversion, y , is greater than the larger threshold, y_B^* , the margin requirement is not binding and both assets are held in the portfolio. In the second region, for values of the lifetime relative risk aversion $y_B^* \geq y > y_{2,D}^*$, the margin requirement binds and both assets are still held in the portfolio. In the third region, for values $y \leq y_{2,D}^*$, the second asset is dropped from the portfolio, and the portfolio consists of holding the first asset at the maximum amount allowed.
- (iii) there are four thresholds $y_B^* > y_{1,D}^* > y_{1,R}^* > y_{2,D}^*$ in the lifetime relative risk aversion and five regions in portfolio composition. The subscripts D, R , in $y_{i,D}, y_{i,R}$ indicate whether asset i is dropped from (D), or reenters (R) the portfolio. In the first region, when $y > y_B^*$, the margin requirement does not bind and both assets are held in the portfolio, with the first asset held short — note that, given our assumption that the first asset has a higher leveraged expected return, for small enough values of lifetime relative risk aversion, y , the first asset is held long — in the second region, when $y_B^* \geq y > y_{1,D}^*$, the margin requirement binds and both assets are held; in the third region, when $y_{1,D}^* \geq y > y_{1,R}^*$, only the second asset is held in the portfolio; in the fourth region, when $y_{1,R}^* \geq y > y_{2,D}^*$, the first asset reenters the portfolio and both assets are held in the portfolio with the first asset held long; finally, in the fifth region, when $y_{2,D}^* \geq y$, the second asset is dropped, and only the first asset is held long in the portfolio to the maximum amount allowed by the margin requirement.

The proof of Proposition 6 is provided in Section F of the Appendix. The proposition illustrates the complications that may arise in selecting a portfolio in the face of constraints. In particular, the case with the four thresholds and the five regions indicates that while the number of assets held in the portfolio tends to decrease as the lifetime relative risk aversion decreases, it does not decrease monotonically, and there can be regions where the asset with the highest leveraged expected return is not held in the portfolio. We note that, by inspection, five regions is the maximum number of regions

that can be encountered when only two risky assets are available, regardless of the labor income correlation with the returns of the risky assets.

2.3.5. *Special Case: Risky Assets with Equal Betas*

An additional case that can be useful to understand our asset selection result is when leveraged expected asset returns adjusted for income correlation are equal among several assets. Since our setting corresponds to that of a single factor CAPM, if the risky assets have the same expected return they also have the same covariance with the market portfolio and the same beta. The following proposition illustrates that asset selection is possible even in the simplest case, when all the assets have the same leveraged expected returns.

Proposition 7. *Assume that there exists a vector of margin coefficients $\lambda^* \in \Lambda$, such that all the leveraged expected returns adjusted for income correlation are equal; i.e.,*

$$\frac{\eta_i}{\lambda_i^*} = \frac{\eta_j}{\lambda_j^*} > 0, \text{ for all } i, j = 1, \dots, N, \quad (29)$$

Then, it is possible that multiple regions of asset selection exist. In addition, if y_K is the largest value of the lifetime relative risk aversion such that only K assets are held in the portfolio with the same signs as the signs of the vector of margin coefficients λ^ , then, for all values of the lifetime relative risk aversion, y , smaller than y_K , $y \leq y_K$, the same K assets are optimally held in the portfolio in constant proportions that do not depend on the asset expected returns.*

The proof of Proposition 7 is provided in Section G of the Appendix. Proposition 7 indicates that, in addition to multiple regions where asset selection occurs, there exists a region where the investor’s portfolio is “trapped”, and no further selection is possible. We point out that, even though the leveraged expected returns of the assets, adjusted for income correlation, are equal across assets, the assets’ volatility does not determine whether an asset is ultimately held or dropped.

2.3.6. Presence of an Index Fund

We consider a case that is of practical interest: the availability, in addition to individual risky assets, of a diversified index fund that is a combination of the individual risky assets. We assume that the investor can invest into N securities and into a market index fund (asset M) spanned by the N securities that corresponds to the market portfolio.

Let π denote the vector of the market portfolio weights

$$\pi = \frac{(\sigma\sigma^\top)^{-1}(\mu - r\bar{\mathbf{1}})}{\bar{\mathbf{1}}^\top(\sigma\sigma^\top)^{-1}(\mu - r\bar{\mathbf{1}})} \in \mathbb{R}_{++}^N. \quad (30)$$

For convenience, set $\varpi = (\bar{\mathbf{1}}^\top(\sigma\sigma^\top)^{-1}(\mu - r\bar{\mathbf{1}}))^{-1}$, so that $\pi = \varpi(\sigma\sigma^\top)^{-1}(\mu - r\bar{\mathbf{1}})$. In other words, the index fund is a linear combination of the N securities with market weights π . The $(N + 1) \times (N + 1)$ covariance matrix V of the N securities and the market index fund is singular with rank N and is given by

$$V = \begin{bmatrix} \sigma\sigma^\top & (\sigma\sigma^\top)\pi \\ \pi^\top(\sigma\sigma^\top) & \pi^\top(\sigma\sigma^\top)\pi \end{bmatrix}. \quad (31)$$

The expected excess return on the index fund is $\mu_M - r = \pi^\top(\mu - r\bar{\mathbf{1}})$ and the expected excess return adjusted for labor correlation is $\eta_M = \pi^\top\eta$.

The following proposition shows that asset selection still occurs in this special case. The proof of the proposition is provided in Appendix J.

Proposition 8. *Consider an investor who faces margin requirements and whose opportunity set includes a riskless asset, N risky assets and an additional, redundant, market index fund, and who receives labor income that is uncorrelated with the returns of the risky assets. We consider two cases:*

- a) *If the margin requirement of the index fund is greater than, or equal to, the weighted average margin requirement of the individual securities, then the margin constraint starts binding at the same level of the financial wealth to income ratio as in the absence of the index fund. We have that:*
 - i) *The optimal portfolio is no longer unique and all securities may be included in the portfolio.*

- ii) *As the ratio of financial wealth over income declines below the level where the margin requirement starts binding, there exists a threshold at which the index fund is dropped from the portfolio and is never reintegrated at lower levels of the ratio.*
- b) *If the margin requirement for the market index fund is smaller than the weighted average margin requirement of the individual securities, with weights determined by the market portfolio, then the margin constraint starts binding at a level of the wealth over income ratio that is lower compared to the case when no index fund is available. We have that:*
- i) *When the margin constraint starts binding, the investor holds only the index fund to the maximum allowed by the margin requirement for the index fund.*
 - ii) *If the leveraged expected return of the index fund is greater than the leveraged expected return of every individual asset, then the index fund is the only asset held in the investor's portfolio for all values of the financial wealth to income ratio once the margin requirement binds. Otherwise, as the wealth to income ratio declines, some individual securities are reintegrated into the portfolio, and there exists a threshold in the investor's financial wealth over income ratio, such that the index fund is dropped from the portfolio and is never reintegrated into the portfolio at lower levels of the ratio.*

The intuition behind Proposition 8 is that, since the index fund is a linear combination of the risky assets, it is subject to the same tradeoffs as the risky assets. In particular, if the margin requirements are the same for both the risky assets and the index fund, then the existence of the index fund does not influence the aggregate holdings of the investor's portfolio, implying that the index fund drops from the portfolio at the same threshold of financial wealth to income ratio as the first risky asset would in the absence of the index fund. Differences between the case of an opportunity set with and without the index fund occur only when the margin requirement for the index fund is lower compared to the margin requirements of individual risky assets. In that case holding the index fund is preferable compared to holding the individual assets. Yet, even in that case, unless the leveraged expected return of the index fund is higher than the leveraged expected return of every risky asset, there is a threshold in the investor's financial wealth to income ratio where the index fund drops from the investor's portfolio, and is not re-integrated for lower values of the ratio.

We note that our result of asset selection would also hold in the case where multiple index funds are available to investor, as long as the funds are weighted linear combinations of assets already in the investor's opportunity set.

2.3.7. Properties of the Consumption and Investment Plans when Labor Income Growth is Spanned

In addition to our result on asset selection, in the case when labor income growth is spanned by the returns of the risky assets, we can characterize optimal consumption using the following propositions, whose proofs are provided in Sections K, L, and M of the Appendix.

Proposition 9. *The optimal consumption is increasing in current wealth and current income and is lower than its unconstrained counterpart. Inside the nonbinding region, z_k^*/W the fraction of wealth invested in risky asset k is lower (higher) than its unconstrained counterpart z_k^f/W whenever $e_k^\top(\sigma\sigma^\top)^{-1}\eta \geq 0(\leq 0)$. Under the same condition, it is increasing (decreasing) in income (wealth).*

Proposition 10. *In the limit of zero current wealth, the lifetime relative risk aversion y goes to zero and the optimal consumption rate is equal to the income rate.*

We note that Proposition 10 implies that an investor with zero current wealth will never be able to accumulate wealth and will always consume his income. This result indicates the extent to which the margin constraint renders holding risky assets unattractive. While the result holds for the infinite horizon setting, in the numerical section we also consider the life-cycle problem, where the investor receives income for only part of his life and then retires and consumes his accumulated wealth.

Proposition 11. *When income is deterministic, for any level of consumption c , the optimal consumption process has a lower volatility than its unconstrained counterpart. These results hold for every strictly concave utility function.*

The intuition behind Proposition 11 is that a margin-constrained investor that receives deterministic income faces less uncertainty than a similar investor that does not face a margin requirement: even though the constrained investor holds an underdiversified portfolio, the magnitude of the portfolio is relatively small compared to the portfolio

of the unconstrained investor. Given the smaller portfolio size, random fluctuations in the stock prices have a smaller impact on the sum of the investor's wealth and discounted future earnings, resulting in a smoother consumption pattern. We point out the result does not necessarily hold in the case of stochastic income growth: if the stochastic income growth is highly correlated with the risky assets, investment in the risky assets can act as a hedge, smoothing out income shocks, rather than magnifying asset price shocks due to the larger size of the investment portfolio.

2.3.8. Finite Horizon

Our analysis has focused so far on the infinite horizon case. To accommodate the case of life-cycle consumption and investment we now consider a case where the investor receives an income stream Y only over the period $[0, T]$ with $T > 0$. At time T , the investor retires and no longer receives any income. After date T , death occurs after an additional τ years. We assume that the investor does not have a bequest motive. Since we assume that the margin constraint is not binding when there is no income, after time T the margin constraint can be ignored — the investor is still subject to margin requirements after retirement, but given the range of parameters we study, the margin requirements do not bind when income is equal to zero. At time T the value function B is given by

$$B(W_T, \tau) = \frac{A^\gamma (1 - e^{-\tau/A})^\gamma}{1 - \gamma} W_T^{1-\gamma}, \quad (32)$$

where A is defined in Eq. (14).

For time $t \leq T$ the value function F satisfies

$$F(W_t, Y_t, t) = \max_{(c, \frac{z}{W} \in Q)} E_t \left[\int_t^T \frac{c_s^{1-\gamma}}{1-\gamma} e^{-\theta(s-t)} ds + B(W_T, \tau) e^{-\theta(T-t)} \right], \quad (33)$$

under the budget constraint

$$dW_s = (rW_s - c_s + Y_s + z_s^\top (\mu - r\bar{\mathbf{1}})) ds + \sigma z_s^\top dw_s, \quad (34)$$

when labor income follows

$$dY_s = Y_s (m ds + \Sigma^\top dw_s + \Theta^\top dw_s^Y), \quad (35)$$

with $W_t > 0, Y_t > 0$ given. Note that F is still homogeneous of degree $1 - \gamma$ and can be written as $F(W_t, Y_t, t) = Y_t^{1-\gamma} f(v_t, t)$, with $v = W/Y$. Over $[0, T]$, the reduced value function f satisfies the following Hamilton-Jacobi-Bellman equation

$$\begin{aligned} \left(\theta + (\gamma - 1)(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2}) \right) f(v, t) &= f_2(v, t) + \frac{\gamma (f_1(v, t))^{\frac{\gamma-1}{\gamma}}}{1 - \gamma} + f_1(v, t) \\ &+ (r - m + \gamma (\Sigma^\top \Sigma + \Theta^\top \Theta)) v f_1(v, t) + \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} v^2 f_{11}(v, t) \\ &+ \max_{\omega \in Q} \left[\omega^\top (\eta v f_1(v, t) - \sigma \Sigma Y v^2 f_{11}(v, t)) + v^2 f_{11}(v, t) \omega^\top \frac{\sigma \sigma^\top}{2} \omega \right], \end{aligned} \quad (36)$$

with $\omega = z/W$ and boundary condition $f(v_T, T) = B(v_T, \tau)$. The analysis conducted in the infinite horizon case still applies, so asset selection still takes place. However, the lifetime relative risk aversion $y_t = -v f_{11}(v, t)/f_1(v, t)$ depends on both the wealth to income ratio, v , and on time to retirement, which implies that the thresholds in the wealth to income ratio where the investor changes his portfolio change as the investor approaches retirement.

3. Numerical Algorithm and Results

To quantitatively illustrate our theoretical results we consider a discrete-time example of an investor who receives income over his working life and retires at a prespecified age. The investor has access to a set of risky assets that we calibrate to U.S. industry portfolios. To numerically solve this optimal asset allocation and consumption problem we extend the numerical algorithm proposed by Brandt et al. (2005) to allow for endogenous state variables and margin constraints. The algorithm is designed to solve optimal control problems using a functional approximation of conditional expectations and is particularly suitable for problems with a large number of state and choice variables. The algorithm proceeds in a dynamic programming fashion, solving the optimal consumption and asset allocation problem backward in time. At each time step the value function is approximated using functional interpolation. The optimal allocation and consumption are computed as solutions of the first order conditions for the prob-

lem's value function, augmented by the constraints multiplied by Lagrange multipliers. One key difference in the algorithm, compared to the algorithm by Brandt et al. (2005), is the introduction of an iterative step to solve the first order conditions: rather than relying on approximating the first order conditions over a large region, we focus our approximation in a neighborhood of a potential solution. Once the solution is computed, we further restrict the neighborhood of approximation and refine the solution, until a desired accuracy is achieved. This improvement in the algorithm by Brandt et al. (2005) was introduced in Yang (2010). We outline the steps of the algorithm below and describe it in detail in Section N of the Appendix.

Algorithm

Step 1: Dynamic Programming

- a. For each time step, starting at the terminal time and working backward, construct a grid in the state space and compute the value function and optimal consumption and portfolio decisions for each point in the grid.
- b. Approximate the value function on the corresponding grid points. This approximation will be used in earlier time steps to compute conditional expectations of the value function.

Step 2: Karush-Kuhn-Tucker Conditions

To solve the Bellman equation for each point on the grid perform the following steps

- a. Combine the constraints in the portfolio positions and the evolution of the state variables with the value function in a Lagrangian function with Lagrange multipliers.
- b. Make a change of variables that allows the consumption optimization problem to be solved independently of the asset allocation optimization problem.
- c. Construct the system of first order conditions (KKT conditions) for the consumption and asset allocation optimization problems.
- d. Find the solution of the Karush-Kuhn-Tucker conditions for the asset allocation optimization problem using an iterative process:
 - i. Start by choosing a region in the choice space that includes the optimal portfolio. This choice can be informed by knowledge of the optimal port-

folio at nearby grid points at the same time step, or for the same grid point at a later date.

- ii. Find an approximately optimal portfolio by solving the system of KKT conditions. To solve the system of KKT conditions, approximate the conditional expectations in the derivatives of the Lagrangian function using cross-test-solution regression: choose a quasi-random set of feasible allocations; calculate the required conditional expectations — interpolating the value of the value function in the following time step from the values at the grid points, estimated from Step 1 in the algorithm — for each feasible choice, and project each on a set of basis functions of the choice variables; solve the resulting system of equations.
- iii. Test Region Contraction: Repeat step (ii) until a convergence criterion is satisfied, using a smaller region in which feasible portfolio choices are drawn, chosen based on the location of the previously computed approximately optimal portfolio.
- iv. Given the optimal portfolio choice, compute the optimal consumption choice using the appropriate KKT condition.

3.1. Calibration

To apply the numerical algorithm, we consider the case of an investor that receives income until age 65, at which point he retires. After retirement the investor has an expected lifetime of 20 years, which matches the data in the 2004 Mortality Table for the Social Security Administration for a 65-year-old female, see Social Security Administration (2004). For the base case we assume that income grows stochastically, independently of the returns of the risky assets, at a growth rate of 3% per year with annualized volatility of 10%, in line with the assumptions in Viceira (2001). We also assume that the investor is not able to either borrow, or short any of the assets, corresponding to parameter values $\lambda^+ = 1$, $\lambda^- = \infty$.

The opportunity set available to the investor includes five risky assets corresponding to the indexes of five industries: Consumer, Manufacturing, High Tech, Health, and Other. To calculate the covariance matrix for each industry we constructed real returns for each industry using the inflation data provided in Robert Shiller's website, see Shiller

(2003), to deflate the annual returns of the five industry portfolios between 1927 and 2004, provided in Ken French's website, see French (2008). The expected returns for each industry are computed using the methodology proposed by Black and Litterman (1992), by matching the market capitalization weights for each industry in July 2008, provided in Ken French's website, to the relative weights that a CRRA investor who receives no income would allocate to each industry within his equity portfolio. The risk-free interest rate was computed from the data in Robert Shiller's website to match the realized one year real interest rate between 1927 and 2004.

The calibration implies that the asset returns satisfy a single factor CAPM. The return of each individual asset includes a systematic and an idiosyncratic component, with the expected return determined by the asset's beta with respect to the market portfolio. Any portfolio of the risky assets that is not proportional to the market portfolio incurs uncompensated idiosyncratic risk.

The calibrated parameters are given in Table 1. The table presents the expected returns, volatilities, and correlations for the five industry indexes, as well as the values for their betas with the market portfolio, Sharpe ratios, systematic risk, and idiosyncratic risk for each index. From the table we notice that the High Tech index has the highest beta and expected return but neither the highest nor the lowest Sharpe ratio or idiosyncratic volatility. The three indexes with the highest Sharpe ratio are Consumer, Manufacturing, and Other, while the Health index has the lowest Sharpe ratio. The Health index has the highest idiosyncratic volatility and the Manufacturing and Other indexes the lowest.

3.2. Diversification Measures

Calvet et al. (2008) present an empirical analysis of diversification of household portfolios in Sweden, and describe several measures that quantify the degree that investors deviate from mean-variance optimal portfolios. We use the same measures to determine the potential magnitude of the impact of the financial constraints on diversification. We present the measures below, following the description and notation in Calvet et al. (2008).

Denoting by $r_{h,t}$ the returns of the risky asset portfolio of an investor that receives labor income and who is subject to margin requirements that can only be satisfied through financial wealth, and by $r_{B,t}$ the returns of the risky portfolio of an investor that does not receive labor income and who is not subject to margin requirements, we have the following variance decomposition

$$r_{h,t} = \alpha_h + \beta_h r_{B,t} + \epsilon_{h,t}, \quad (37)$$

Notice that, given our calibration, the assumptions underlying the CAPM are satisfied for the unconstrained investor who does not receive labor income. In line with the CAPM, this investor holds a risky portfolio that includes systematic risk only.

If we denote by σ_B, σ_h the standard deviation of the returns of the equity portfolio of the unconstrained investor that does not receive labor income, and of the constrained investor that receives labor income, respectively, we have

$$\sigma_h^2 = \beta_h^2 \sigma_B^2 + \sigma_{i,h}^2. \quad (38)$$

The interpretation of this decomposition is that the portfolio of the constrained investor that receives labor income has *systematic risk* $|\beta_h| \sigma_B$ and *idiosyncratic risk* $\sigma_{i,h}$. The *idiosyncratic variance share* is given by

$$\frac{\sigma_{i,h}^2}{\sigma_h^2} = \frac{\sigma_{i,h}^2}{\beta_h^2 \sigma_B^2 + \sigma_{i,h}^2}. \quad (39)$$

Another measure of portfolio diversification is the Sharpe ratio of the risky portion of the portfolio. We denote the Sharpe ratio of the portfolio of an investor that does not receive income and does not face financial constraints S_B , and the Sharpe ratio of a constrained investor that receives labor income S_h . These ratios are defined by the ratio of the excess return of the respective portfolio to the standard deviation of excess returns

$$S_h = \frac{\mu_h}{\sigma_h}, \quad (40)$$

where μ_h, σ_h , are the excess return and the standard deviation of excess return for the portfolio of the constrained investor that receives labor income. The *relative Sharpe ratio loss* is defined by

$$RSRL_h = 1 - \frac{S_h}{S_B}. \quad (41)$$

While the relative Sharpe ratio loss is a measure of the diversification loss in the risky asset portion of the portfolio, it does not necessarily reflect the overall efficiency loss in the portfolio. To capture this loss, we define the *return loss* as the average return loss by the investor for choosing a portfolio other than the one suggested by the CAPM

$$RL_h = w_h(S_B\sigma_h - \mu_h), \quad (42)$$

where w_h is the portion of the portfolio invested in risky assets.

Finally, we define a measure associated with utility losses for the constrained investor who receives labor income, compared to the unconstrained investor who does not receive labor income. It is defined as the increase in the risk-free rate that would make the constrained investor indifferent between being constrained with the higher risk-free rate and being unconstrained. In the case of a risk-averse investor with CRRA preferences with risk aversion coefficient γ , Calvet et al. (2008) calculate the utility loss from the relationship

$$UL_h = \frac{S_B^2 - S_h^2}{2\gamma}. \quad (43)$$

3.3. Base Case

The optimal asset allocations for the base case parameters, listed in Table 1, are presented in Fig. 2 for investors 30 and 60 years old over a range of wealth to income ratios.

From Fig. 2 we notice that as the financial wealth of the investor decreases compared to his income, the investor allocates a larger proportion of his wealth to the risky assets. For the younger, 30-year-old, investor the margin constraint binds if the investor's financial wealth is smaller than 10.4 times his annual income. While the proportion in which each risky asset is held within the equity portfolio does not change when the margin constraint is not binding, once the constraint binds the investor shifts his portfolio to

increase the portfolio's expected return, and holds an underdiversified portfolio. When the financial wealth reaches a level of 6.9, 5.6, 3.4, and 0.98 times the investor's annual income, the investor drops the Health, Manufacturing, Consumer, and Other industry indexes from his portfolio, respectively. For financial wealth levels below 98% of the investor's annual income, the investor's equity portfolio consists only of the index of the High Tech industry. A similar pattern is observed for the 60-year-old investor. In that case, since the remaining income spans a smaller number of years; i.e., the discounted value of future earnings is smaller than that of the younger investor, the constraint binds at a lower level of the financial wealth equal to 2.3 times annual income. For lower levels of the financial wealth to income ratio the older investor also shifts his equity portfolio, and holds an underdiversified portfolio, dropping the Health, Manufacturing, Consumer, and Other industry indexes at ratios of 1.8, 1.6, 1.1, and 0.4 respectively.

Table 2 presents further details of the optimal allocations for different levels of the financial wealth to annual income ratio, as well as values for the various diversification measures and the investor's lifetime relative risk aversion. From the table we notice that when the margin constraint is not binding and the ratio of financial wealth to income decreases, the investor increases the portfolio's expected return by increasing the percentage of his wealth invested in risky assets while maintaining a diversified portfolio. Once the constraint binds, further reductions in the financial wealth to annual income ratio result in a deterioration of the portfolio diversification measures. As an example, a younger, 30-year-old, investor whose financial wealth is equal to one year of his labor income holds a portfolio that has 11.4% idiosyncratic volatility — which corresponds to 11.8% of the portfolio's variance — Sharpe ratio of 25.8% compared to 27.3% achieved when the portfolio is diversified, a return loss of 51 basis points per year, and a utility loss of 14 basis points per year. We point out that the higher idiosyncratic volatility, lower Sharpe ratio and higher expected return compared to the diversified portfolio, is in line with the empirical evidence in Ivković et al. (2008). Even though the equity portfolio of the constrained investor has higher volatility, its size is smaller than the equity portfolio of the unconstrained investor. Due to the smaller size of the equity portfolio, shocks to the prices of the risky assets have a smaller effect to the wealth of the constrained investor, leading to a smoother consumption choice. The lifetime relative risk aversion of the investor is 0.27, close to that of a risk-neutral investor. We point out that even though the 30-year-old investor holds a single asset in the portfolio

only when his financial wealth is below 98% of his annual income, it is much more likely to observe an underdiversified portfolio, since it corresponds to a ratio of the investor's wealth that is less than 10.4 times his annual income.

Panel B of Table 2 presents allocations and diversification measures for the 60-year-old investor. The results are qualitatively similar to the results in Panel A, with the main difference being that the margin constraint binds at lower levels of the financial wealth to annual income ratio.

Table 3 presents results obtained by simulating the evolution of the portfolio of an investor starting at age 20. From Panel A we notice that the investor whose financial wealth at age 20 was twice his annual income, holds, at age 30, a portfolio that is, almost always, underdiversified. At the same time the investor consumes close to his annual labor income. At age 45 an investor that has done badly starts consuming considerably less than his annual income and saves the remainder, while a consumer that has done well consumes more than his annual income. The portfolio is still mostly underdiversified and the margin requirements bind. At age 60 the investor has accelerated his saving behavior and is mostly unconstrained in his financial portfolio.⁷

Panel B of Table 3 presents the simulation results for an investor whose financial wealth at age 20 is equal to ten times his annual income. Even though this investor is relatively richer than the investor in Panel A, the margin constraint still largely binds at age 30, leading to the investor holding an underdiversified equity portfolio. Given his large financial wealth, this investor saves less, and starts saving later in life, compared to the investor in Panel A. Overall, the results in both panels indicate that younger investors, even if they have significant amounts of financial wealth, are holding portfolios far less diversified than those held by older, unconstrained, investors.

3.4. Comparative Statics

Regulation T Margin Constraints

⁷Since consumption is measured with respect to current annual labor income and since in this example income has a 3% annual drift, the reduction in consumption relative to income observed in the table does not necessarily imply a reduction in the actual amount consumed by the investor. Nevertheless, consumption to income ratios above 1.0 imply that the investor consumes part of his financial wealth while ratios below 1.0 imply that the investor saves part of his labor income.

Fig. 3 presents the optimal asset allocations for an investor facing a margin constraint in line with the requirements in Regulation T of 50% for long positions and 150% for short positions. For the calibrated parameter values from Table 1, the investor never shorts any of the risky assets. Compared to Fig. 2, the investor is unconstrained for a greater range of his financial wealth to income ratio, with the allocations being identical when the margin constraint does not bind for either investor. The margin constraint for the investor that faces the Regulation T margin requirements binds at a level of financial wealth equal to 2.6 times his annual income at age 30 and 94% of his annual income at age 60. Below that level, the investor's portfolio is underdiversified. The order that assets drop out of the portfolio is the same as in the base case, and the last asset held in the portfolio is the High Tech industry index, which is exclusively held at levels of financial wealth below 30% of annual income at age 30 and below 19% of annual income at age 60.

Deterministic Income

To understand the impact of labor income risk on the investor's allocation to risky assets and on the degree of underdiversification, we consider a comparative static where the investor receives labor income with a deterministic growth rate. Fig. 4 presents the optimal asset allocations. The volatility of the growth of the labor income is set to zero, while the remaining parameters are the same as in the base case, given in Table 1. From the figure we notice that, when labor growth is deterministic, allocations in the industry indexes are increased compared to the base case of stochastic income for investors of both age 30 and age 60. This increased allocation to risky assets is intuitively expected due to the lower risk implied by the deterministic nature of income growth.

In line with our theoretical results, the order in which assets are dropped from the equity portfolio when the ratio of financial wealth to income decreases is the same as in the base case, described in Fig. 2. Due to the higher allocations to the risky assets, compared to the case with stochastic income, the threshold when the margin constraint binds is higher when income is deterministic. For a 30-year-old investor who receives income with deterministic growth the margin binds at a financial wealth level equal to 13.0 times his annual income, while for the investor who receives income with stochastic growth the margin constraint binds at a level of financial wealth equal to 10.4 times his annual income. An alternative explanation for this difference is that since income growth is uncorrelated with asset returns, income helps in diversifying the investor's

portfolio, implying that the margin requirement is less onerous. Below these levels, the investor holds an underdiversified portfolio.

Overall, Fig. 4 illustrates that the differences between the case of deterministic and stochastic income growth are quantitative rather than qualitative: in both cases an investor with low levels of financial wealth compared to labor income holds underdiversified portfolios consisting of only a few out of many possible risky assets. However, the range of financial wealth to annual income ratios where an investor's portfolio is underdiversified is greater for an investor whose labor income grows deterministically compared to an investor whose labor income grows with a stochastic rate that is independent of the returns of the risky assets.

Nonnegative Wealth

A case of constrained choice previously studied in the literature is the case when the investor's wealth is required to remain greater or equal to zero but where the investor does not face a margin requirement, see He and Pagès (1993), El Karoui and Jeanblanc-Picqué (1998), Duffie et al. (1997), and Dybvig and Liu (2010). Compared to a requirement of nonnegative wealth, the margin requirement is a stricter constraint, since it automatically guarantees nonnegative wealth. To quantify the difference in asset allocations, Fig. 5 presents the optimal asset allocation for an investor facing a nonnegative wealth constraint, but who is otherwise identical to our base case investor. From the figure we notice that in both the cases of a nonnegative wealth constraint and of a margin requirement, investment in risky assets increases as the wealth to income ratio decreases. On the other hand, there are significant qualitative differences: unlike the case of a margin requirement, an investor that faces a nonnegative wealth constraint maintains a diversified portfolio, even when his income is much greater than his wealth; in addition, the size of the risky asset portfolio of an investor that faces a nonnegative wealth constraint is much larger than that of an investor that faces a margin requirement. To finance this larger investment in risky assets, the results in the figure indicate that the investor that is constrained to maintain nonnegative wealth borrows amounts up to 10 times his financial wealth or more, using his income as collateral.

Correlated Labor Income

Fig. 6 presents the asset allocations for the case when the growth in the labor income is correlated with the returns of the High Tech index, keeping overall volatility of labor

income growth the same as in the base case. The parameters for this case approximate the labor income received by an employee in a technology firm. From the figure we note that, although the High Tech index has the highest expected return, it is the first asset to drop from the portfolio, at a ratio of financial wealth to annual income of 118.9 for the 30-year-old investor. In addition, the portfolio becomes underdiversified before the margin requirement binds, at a ratio of financial wealth to annual income of 11.4, with the proportion of the High Tech index decreasing as labor income increases relative to financial wealth. These results are expected, since, intuitively, as the investor's labor income increases relative to his financial wealth, his labor income creates an implicit exposure to the shocks in the High Tech index and a corresponding hedging demand, resulting in a decrease in the portfolio allocation in the High Tech index until it is dropped from the portfolio. The remaining assets are uncorrelated with the growth labor income and drop out of the portfolio as the ratio of financial wealth to income decreases, with the index Other being the last asset held — for the 30-year-old investor the assets Health, Manufacturing, and Consumer drop at thresholds of financial wealth to annual income ratios of 4.6, 4.5, and 2.3 respectively. This case illustrates that correlation between income growth and asset returns alters portfolio composition, while still resulting in asset selection and concentrated portfolios.

4. Conclusion

The results we have presented indicate that financial constraints can be a significant determinant of individual portfolios, and can, to some extent, account for empirical findings. The ratio of current wealth to income is instrumental in the determination of the portfolios, and the extent to which investors deviate from diversified portfolios. For large values of this ratio the investor is unconstrained, while the constraint has the largest effect at low values of the ratio. This result, which holds even when fully diversified index funds are available, implies that young investors are most likely to be affected while older investors are more likely to hold diversified portfolios. This prediction is in line with several empirical papers. For example, Goetzmann and Kumar (2008) and Calvet et al. (2008) report that age is a significant determinant of underdiversification. Kumar (2009) reports that young investors are more likely to hold stocks with lottery-like payoffs that seemingly expose them to uncompensated risk. Goetzmann and Kumar (2008) report

that households that only have a retirement investment account, which presumably includes households that do not have enough wealth for an additional investment account, hold more underdiversified portfolios. Our findings also provide a rational explanation for the empirical finding that investors only hold a small number of stocks in their portfolio: similar to Black (1972), constrained investors try to increase their expected return at the cost of holding less diversified portfolios.⁸ Ivković et al. (2008) show that, while investors hold relatively few stocks in their portfolios, the number rises with an increase in account balance, which can be thought of as a proxy of current wealth.

Beyond the existing empirical literature, our theoretical and numerical results also provide several empirically testable predictions. For example, the calibration predicts that severely constrained investors; i.e., those with a very low wealth to income ratio, will hold only the asset with the highest leveraged expected return adjusted for covariance with labor income, which is not necessarily the one with the highest Sharpe ratio. As we already mentioned, Ivković et al. (2008) report that concentrated portfolios have lower Sharpe ratios and higher expected returns. Another example would be to test whether investors that borrow on margin — a possible indication that the investor is financially constrained — hold less diversified portfolios than investors that do not.

An additional empirical prediction involves the dynamics of underdiversification: given that our model predicts that the degree of underdiversification of an investor's portfolio depends on the ratio of his labor income to his financial wealth, we would expect underdiversification to decrease following a negative shock to an investor's labor income, such as the loss of a job.

While our findings reveal a clear link between the combination of labor income and financial constraints and underdiversified portfolios, several hurdles remain before a rational model can explain all the available empirical findings. One challenge is the conflict between the theoretical prediction that investors will not hold the riskless asset and an undiversified equity portfolio simultaneously, and the empirical results reported in Polkovnichenko (2005) and Calvet et al. (2008). While our model cannot address this issue, a possible resolution could be a model with an additional cost imposed on trading

⁸An interesting question is whether the inclusion of put options, with their higher leverage, would alleviate the financial constraint. While options have not empirically been a significant component of individual portfolios, the reason why investors shun them is unclear, and can be, for example, due to their high prices, see Kubler and Willen (2006). This question is outside the scope of our paper.

risky assets. Such a cost could be due, for example, to transaction costs for buying and selling assets, the cost of learning about asset characteristics, or capital gain taxes.⁹

While our framework implies that holding individual stocks in investor portfolios is optimal even when mutual funds or exchange-traded funds (ETFs) are available, there may be other, quantitatively different, explanations. Beyond forced holdings, such as company stock granted to employees, individual stock selection could be predicted in a rational framework by including competition for scarce local resources, creating a home-bias effect in areas where local companies grant stocks to their employees they cannot trade out of — see DeMarzo, Kaniel, and Kremer (2004). Van Nieuwerburgh and Veldkamp (2010) propose another possible explanation. Assuming that investors have limited resources to learn about individual stocks, they show that it is optimal to focus on a small subset of stocks, and hold portfolios that simultaneously include a diversified fund and a concentrated set of assets.

In addition to holding underdiversified portfolios, the empirical literature shows that investors tend to focus on assets with high idiosyncratic volatilities, a preference often attributed to behavioral biases. We have shown that rational alternatives exist: the covariance structure of the risky assets and the asset characteristics affect which stocks are held and can lead to portfolios concentrated on assets with higher idiosyncratic risk; another possibility is that labor income growth is positively correlated with assets with low idiosyncratic volatility. For investors that cannot short a risky asset, with little financial wealth, and relatively many years to retirement, the positive correlation between labor income growth and the returns of a risky asset generates hedging demand leading to dropping the correlated asset from the investor’s portfolio. Across such investors one would observe concentrated portfolios on assets whose correlation with labor income growth is low, which are likely to have high idiosyncratic volatility. A final possibility is that assets with high betas and expected returns have high idiosyncratic volatilities: since constrained investors tend to trade excess return for diversification, concentrated portfolios are likely to include assets with high expected returns and high idiosyncratic volatility. In our numerical base case, calibrated to five industry portfolios, we have seen that this relationship is not monotone since the asset with the highest expected return, the High Tech index, is not the asset with the highest idiosyncratic volatility.

⁹In Gallmeyer, Kaniel, and Tompaidis (2006) it is shown that capital gain taxation can also induce an investor to hold an underdiversified portfolio, while simultaneously holding the riskless asset.

In addition to rational explanations, it is likely that behavioral based explanations have a significant effect. Our contribution in this paper is to offer the ratio of wealth over income as a variable that can be used to understand underdiversification in investor portfolios. It would be important to find additional variables that can distinguish between the rational and behavioral explanations.

Beyond offering a potential explanation for the empirically observed concentration and underdiversification of household portfolios, our framework can also be applied to the case of mutual fund and hedge fund managers that face leverage constraints. Similar to the investors in our framework, the constraints would lead the managers to trade diversification for higher excess returns, leading them to hold portfolios concentrated in a few assets with high expected returns and betas.

An interesting extension of our work would be to consider assets with different margin requirements. In this case, we expect that the assets that have the highest expected return, when leveraged to the greatest extent possible, would appear most attractive to constrained investors. Such behavior would be in line with the preference of individual investors for residential real estate investments over financial investments, due to the lower margin requirements for residential real estate.

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

$$dW_{\alpha s} = \alpha W_s ds - \alpha c_s ds + \alpha Y_s ds + \alpha z_s^\top (\mu - r\bar{1}) ds + \alpha z_s^\top \sigma dw_s = \alpha dW_s, \quad (44)$$

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

$$\lambda^\top(\alpha z) = \alpha \lambda^\top z \leq \alpha W. \quad (45)$$

It follows that

$$F(\alpha W, \alpha Y) \leq \alpha^{1-\gamma} F(W, Y), \quad (46)$$

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

$$F(W, Y) = F(\alpha^{-1}\alpha W, \alpha^{-1}\alpha Y) \leq \alpha^{\gamma-1} F(\alpha W, \alpha Y), \quad (47)$$

so given Eq. (46) in fact we have

$$F(\alpha W, \alpha Y) = \alpha^{1-\gamma} F(W, Y). \quad (48)$$

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

$$S : (\alpha c + (1 - \alpha)c', \alpha x + (1 - \alpha)x', \alpha z + (1 - \alpha)z'), \quad (49)$$

is admissible for the initial condition

$$I : (\alpha W_t + (1 - \alpha)W'_t, \alpha Y_t + (1 - \alpha)Y'_t). \quad (50)$$

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

$$W_s^\alpha = \alpha W_s + (1 - \alpha)W'_s, \quad (51)$$

and similarly for the income process

$$Y_s^\alpha = \alpha Y_s + (1 - \alpha)Y'_s. \quad (52)$$

The margin constraint is satisfied since

$$\lambda^\top(\alpha z + (1 - \alpha)z') = \alpha \lambda^\top z + (1 - \alpha) \lambda^\top z' \leq \alpha W + (1 - \alpha)W \leq W, \quad (53)$$

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

$$E_t \left[\int_t^\infty u(\alpha c_s + (1 - \alpha)c'_s) e^{-\theta s} ds \right] > E_t \left[\int_t^\infty (\alpha u(c_s) + (1 - \alpha)u(c'_s)) e^{-\theta s} ds \right], \quad (54)$$

which implies that

$$F(\alpha W_t + (1 - \alpha)W'_t, \alpha Y_t + (1 - \alpha)Y'_t) > \alpha F(W_t, Y_t) + (1 - \alpha)F(W'_t, Y'_t). \quad (55)$$

Appendix C. Proof of Proposition 3

We note that the assumption that $(\sigma\sigma^\top)^{-1}\eta \in \mathbb{R}_+^N$ 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 $(\sigma\sigma^\top)^{-1}$ are non-positive implies that $(\sigma\sigma^\top)^{-1} = \alpha I_N - P$, where $\alpha > 0$ and P is a matrix with non-negative elements. Since $(\sigma\sigma^\top)^{-1}$ 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

$$\sigma\sigma^\top = \frac{1}{\alpha}(I_N - \frac{P}{\alpha})^{-1} = \frac{1}{\alpha} \sum_{n=0}^{\infty} \left(\frac{P}{\alpha}\right)^n < \infty \quad (56)$$

as the spectral radius of P/α , is strictly less than one. We conclude that all the entries of the covariance matrix $\sigma\sigma^\top$ 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 $M = (\sigma\sigma^\top)^{-1} = [m_{ij}]$. It is easy to check that

$$\begin{aligned} m_{ii} &= \frac{1 + (N-2)\rho}{(1-\rho)(1+(N-1)\rho)} \frac{1}{\sigma_i^2} > 0 \\ m_{ij} &= -\frac{\rho}{(1-\rho)(1+(N-1)\rho)} \frac{1}{\sigma_i\sigma_j} < 0, \quad i \neq j. \end{aligned} \quad (57)$$

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

$$\max_{\omega \in \mathbb{R}_+^N} \omega^\top \eta - \frac{y}{2} \omega^\top (\sigma\sigma^\top) \omega, \quad \text{s.t. } \omega^\top \lambda^+ \leq 1, \quad (58)$$

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

$$\max_{\hat{\omega} \in \mathbb{R}_+^N} \hat{\omega}^\top \hat{\eta} - \frac{y}{2} (\hat{\omega})^\top (\hat{\sigma}\hat{\sigma}^\top) \hat{\omega}, \quad \text{s.t. } \hat{\omega}^\top \bar{1} \leq 1. \quad (59)$$

Observe that $(\sigma\sigma^\top)^{-1}\eta \in \mathbb{R}_+^N$ if and only if $(\widehat{\sigma}\widehat{\sigma}^\top)^{-1}\widehat{\eta} \in \mathbb{R}_+^N$. 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

$$\mathcal{N}_{a,b} = \{(a, b) \in \mathbb{R}_+ \times \mathbb{R}_+^N, (1 - \lambda_i^+)a \leq b_i \leq (1 + \lambda_i^-)a\}. \quad (60)$$

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

$$\min_{(a,b) \in \mathbb{R}_+ \times \mathbb{R}_+^N} ay + \frac{1}{2}(\eta + b - a\bar{\mathbf{1}})^\top (\sigma\sigma^\top)^{-1}(\eta + b - a\bar{\mathbf{1}}). \quad (61)$$

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

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

$$\frac{1}{2}(\eta + b - a^*(y)\bar{\mathbf{1}})^\top (\sigma\sigma^\top)^{-1}(\eta + b - a^*(y)\bar{\mathbf{1}}) = \frac{(a^*(y))^2}{2}(\widehat{\eta}^*(y) + b' - \bar{\mathbf{1}})^\top (\sigma\sigma^\top)^{-1}(\widehat{\eta}^*(y) + b' - \bar{\mathbf{1}}), \quad (62)$$

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

$$\pi_i(b'_i, b_i'^{-1}; y) = -\frac{1}{2}(\widehat{\eta}^*(y) + b' - \bar{\mathbf{1}})^\top (\sigma\sigma^\top)^{-1}(\widehat{\eta}^*(y) + b' - \bar{\mathbf{1}}). \quad (63)$$

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

$$\begin{aligned} \frac{\partial^2 \pi_i}{\partial b'_i \partial b'_j} &= -e_i^\top (\sigma\sigma^\top)^{-1} e_j \geq 0 \\ \frac{\partial^2 \pi_i}{\partial b_i \partial y} &= \frac{\partial a^*(y)}{\partial y} \frac{1}{(a^*(y))^2} e_i^\top (\sigma\sigma^\top)^{-1} \eta \leq 0. \end{aligned} \quad (64)$$

From Eq. (64), we have that this game satisfies the supermodularity conditions, so its unique Nash equilibrium $(b_1^*, b_2^*, \dots, b_N^*)$ 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

$$\omega_k^* = \frac{\mu_k - r}{y\sigma_k^2}. \quad (65)$$

For y slightly below y_B^* , we have

$$\omega_k^* = \frac{1}{y\sigma_k^2} (\mu_k - r - \psi_N(\lambda^*, y)), \quad (66)$$

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

$$\psi_N(\lambda^*, y) = \frac{\alpha^\top \xi - y}{\alpha^\top \mathbf{1}}, \quad (67)$$

where

$$\xi_k = \frac{\mu_k - r}{\lambda_k^*}, \quad (68)$$

and

$$\alpha_k = \left(\frac{\lambda_k^*}{\sigma_k} \right)^2, k \in \{1, \dots, N\}. \quad (69)$$

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

$$\frac{\omega_k^*}{\lambda_k^*} \geq 0, \quad (70)$$

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

$$\omega_k^* = \frac{\alpha_k}{y\lambda_k^*\alpha^\top\bar{\mathbf{1}}}(y - \alpha^\top(\xi - \xi_k\bar{\mathbf{1}})). \quad (71)$$

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

$$\psi_N(\lambda^*, y_B^*) = 0, \quad (72)$$

which leads to

$$y_B^* = \alpha^\top\xi. \quad (73)$$

Next, without loss of generality, assume that $0 < \xi_N < \xi_{N-1} < \dots < \xi_1$. 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

$$y_{N,D}^* = \alpha^\top(\xi - \xi_N\bar{\mathbf{1}}). \quad (74)$$

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

$$y_{K,D}^* = (I_K\alpha)^\top [I_K(\xi - \xi_K\bar{\mathbf{1}})], \quad (75)$$

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

$$\omega_k^* = \frac{\mu_k - r - \psi_K(\lambda, y)\lambda_k}{y\sigma_k^2}. \quad (76)$$

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, $K \in \{1, \dots, N\}$ with

$$\omega_k^* = \frac{\alpha_k}{y\lambda_k^*} \frac{[y - y_{K,D}^*]^+}{(I_K\alpha)^\top\bar{\mathbf{1}}}, \quad k = 1, \dots, N. \quad (77)$$

Appendix E. Proof of Proposition 5

The margin constraint is equivalent to 2^N linear constraints of the form $\lambda^\top z \leq W$, 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 k th 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 $\{\lambda^{(i)}\}_{i=1}^{K+1}$ of the binding constraints must have $N - K$ components in common. The Hamilton-Jacobi-Bellman (HJB) equation for the primal value function F is

$$\begin{aligned} \theta F = \max_{\frac{z}{W} \in Q} & \frac{\gamma(F_1)^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + (rW + Y)F_1 + mYF_2 + \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} Y^2 F_{22} \\ & + z^\top ((\mu - r\bar{1})F_1 + \sigma \Sigma Y F_{12}) + \frac{z^\top \sigma \sigma^\top z}{2} F_{11}. \end{aligned} \quad (78)$$

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

$$\max_{\omega \in Q} \omega^\top (\eta + y \sigma \Sigma) - \frac{y}{2} \omega^\top \sigma \sigma^\top \omega, \quad (79)$$

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

$$y = -\frac{W F_{11}}{F_1} = -\frac{v f''(v)}{f'(v)}, \quad (80)$$

the program defined in Eq. (79) 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. (79) is independent of the parameter y , so the fraction of wealth invested in each asset is constant. The unconstrained allocation is

$$\frac{z}{W} = (\sigma \sigma^\top)^{-1} \sigma \Sigma. \quad (81)$$

If

$$\max_{\lambda \in \Lambda} (\lambda^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma) \leq 1, \quad (82)$$

the margin constraint is never binding, so

$$\frac{z^*}{W} = (\sigma \sigma^\top)^{-1} \sigma \Sigma. \quad (83)$$

If, on the other hand,

$$\max_{\lambda \in \Lambda} (\lambda^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma) > 1, \quad (84)$$

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

$$\max_{\lambda \in \Lambda} (\lambda_k e_k^\top I_K \omega^*) > 0, \quad k = 1, \dots, K, \quad (85)$$

with

$$I_K \omega^* = \frac{(I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma + (1 - \lambda^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma) I_K \lambda}{\lambda^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda} \quad (86)$$

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

$$\omega^* = \frac{(\sigma \sigma^\top)^{-1}}{y} (\eta + y \sigma \Sigma). \quad (87)$$

To satisfy the margin constraint, we must have

$$\max_{\lambda \in \Lambda} (\omega^*)^\top \lambda \leq 1. \quad (88)$$

First, we characterize the binding cutoff y_B^* . As long as the constraint is not binding, the optimal asset allocation is given by Eq. (87). Define

$$y_B^* = \max_{\lambda \in \Lambda} \frac{\lambda^\top (\sigma \sigma^\top)^{-1} \eta}{1 - \lambda^\top (\sigma \sigma^\top)^{-1} \sigma \Sigma}. \quad (89)$$

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

$$(\lambda_B^*)^\top \frac{(\sigma \sigma^\top)^{-1}}{y_B^*} (\eta + y_B^* \sigma \Sigma) = 1, \quad (90)$$

so the constraint is binding at $y = y_B^*$. Using Eq. (87), and the condition on the matrix J_K in Eq. (21), 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. (87) is the same for all $i = 1, \dots, N$.

Case $\Theta = 0$. Using Eq. (87), we obtain the following reduced HJB equation

$$\begin{aligned} \left(\theta + (\gamma - 1)(m - \gamma \frac{\Sigma^\top \Sigma}{2}) \right) f(v) &= \frac{\gamma}{1 - \gamma} (f'(v))^{\frac{\gamma-1}{\gamma}} + f'(v) + B^{-1} v f'(v) \\ &\quad - \frac{1}{2} \eta^\top (\sigma \sigma^\top)^{-1} \eta \frac{(f'(v))^2}{f''(v)}. \end{aligned} \quad (91)$$

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

$$\begin{aligned} \left(\theta + (\gamma - 1)(m - \gamma \frac{\Sigma^\top \Sigma}{2}) \right) J(x) &= \frac{\gamma}{1 - \gamma} x^{\frac{\gamma-1}{\gamma}} + x \\ &\quad + (\theta - B^{-1} + (\gamma - 1)(m - \gamma \frac{\Sigma^\top \Sigma}{2})) x J'(x) \\ &\quad + \frac{1}{2} \eta^\top (\sigma \sigma^\top)^{-1} \eta x^2 J''(x). \end{aligned} \quad (92)$$

The general solution is

$$J(x) = \frac{\gamma A x^{\frac{\gamma-1}{\gamma}}}{1 - \gamma} + Bx + \frac{\gamma K}{\beta - 1 + \gamma} x^{\frac{\beta-1+\gamma}{\gamma}} + \frac{\gamma L}{\delta - 1 + \gamma} x^{\frac{\delta-1+\gamma}{\gamma}}, \quad (93)$$

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

$$\frac{1}{2\gamma^2} (\eta^\top (\sigma\sigma^\top)^{-1} \eta) x^2 + \left(A^{-1} - B^{-1} - \frac{1}{2\gamma^2} \eta^\top (\sigma\sigma^\top)^{-1} \eta \right) x = A^{-1}. \quad (94)$$

We note that if x is a root of the quadratic

$$\left(\theta + (\gamma - 1) \left(m - \gamma \frac{\Sigma^\top \Sigma}{2} \right) \right) = (\theta - B^{-1} + (\gamma - 1) \left(m - \gamma \frac{\Sigma^\top \Sigma}{2} \right)) x + \frac{1}{2} \eta^\top (\sigma\sigma^\top)^{-1} \eta x^2, \quad (95)$$

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

$$\frac{1}{2} (\eta^\top (\sigma\sigma^\top)^{-1} \eta) x^2 + \left(A^{-1} - B^{-1} - \frac{1}{2} \eta^\top (\sigma\sigma^\top)^{-1} \eta \right) x = A^{-1}. \quad (96)$$

Differentiating Eq. (93) with respect to x and using the fact that $x = f'(v)$ and $v = -J'(x)$ leads to

$$v + B = A f'(v)^{-\frac{1}{\gamma}} + K f'(v)^{\frac{\beta-1}{\gamma}} + L f'(v)^{\frac{\delta-1}{\gamma}}. \quad (97)$$

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

$$f_0(v) < f(v) + f_0(M) - f(M). \quad (98)$$

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

$$L = \omega^\top (\eta + y\sigma\Sigma) - \frac{1}{2} y \omega^\top \sigma \sigma^\top \omega - \psi(\omega^\top \lambda - 1), \quad (99)$$

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

$$\omega^* = \frac{(\sigma\sigma^\top)^{-1}}{y}(\eta + y\sigma\Sigma - \psi_N(\lambda_B^*, y)\lambda_B^*). \quad (100)$$

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

$$\psi_N(\lambda_B^*, y) = \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \eta - (1 - (\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \sigma \Sigma) y}{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \lambda_B^*}. \quad (101)$$

This derivation is valid as long as for all $i = 1, \dots, N$, $\omega_i^*/\lambda_{B,i}^* \geq 0$. 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. (100) and (101), it is easy to verify that $z_i^* = 0$ exactly when $y = y_{i,N}$ with

$$y_{i,N} = \frac{\left(\lambda_B^* - \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \lambda_B^*}{e_i^\top (\sigma\sigma^\top)^{-1} \lambda_B^*} e_i \right)^\top (\sigma\sigma^\top)^{-1} \eta}{1 - \left(\lambda_B^* - \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \lambda_B^*}{e_i^\top (\sigma\sigma^\top)^{-1} \lambda_B^*} e_i \right)^\top (\sigma\sigma^\top)^{-1} \sigma \Sigma}. \quad (102)$$

We can assume that $y_{N,N} = \max_{i=1, \dots, N} \{y_{i,N}\}$ 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 $\bar{\lambda}_B^*$ 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 $\psi_N(\lambda_B^*, y_{N,N}) \neq \psi_N(\bar{\lambda}_B^*, y_{N,N})$, 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. (79) 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 simplify 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 $[y_{N,K}^-, y_{N,K}^+]$ with $0 < y_{N,K}^- < y_{N,K}^+ \leq y_B^*$, the first K assets are optimally held in nonzero positions. Then, for all $y \in [y_{N,K}^-, y_{N,K}^+]$, there is a vector $\lambda \in \Lambda$ such that the optimal asset allocation is given by*

$$I_K \omega^* = \frac{I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K (\eta - \psi_K(\lambda, y) \lambda)}{y}, \quad (103)$$

and satisfies that

(i) for all $i \in \{1, \dots, K\}$, $\omega_i^*/\lambda_i \geq 0$

(ii) the Lagrange multiplier ψ_K associated with the optimization problem is positive and given by

$$\psi_K(\lambda, y) = \frac{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \eta - (1 - (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma) y}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda}. \quad (104)$$

(iii) Risky asset allocations are given by

$$I_K \omega^* = \frac{(I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \eta}{y} + \left(M_K - \frac{L_K}{y} \right) (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda, \quad (105)$$

with

$$\begin{aligned} L_K &= \frac{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \eta}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda} \\ M_K &= \frac{1}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda}. \end{aligned} \quad (106)$$

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

Proof of Lemma E.1: By assumption for all $\lambda \in \Lambda$, $(1 - \lambda^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma) > 0$, so as y decreases, ψ_K remains positive. Consider the optimization problem $P_{N,K}$ 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. (103) where the Lagrange multiplier is given by Eq. (104) 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. (103) for $K = J$ unless asset $K + 1$ is held in a zero position. Whenever the asset position is given by Eq. (103), 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. (104) 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 $[y_{N,K}^-, y_{N,K}^+]$ 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 $I_K^T \lambda$ 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. (104) is a linear function that decreases with the second argument y , which implies that the components of the vector $y I_K \omega^*$ is also a linear function of y , where $I_K \omega^*$ is given by Eq. (103). 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

$$\omega_{K+1}^* = \frac{A_{K+1} - B_{K+1}y}{y}, \quad (107)$$

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 $[y_{N,K}^-, y_{N,K}^+]$. 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 if $e_{J+1}^\top \omega^* \neq 0$ at all $y = y_{N,J}^- - \varepsilon$, with $\varepsilon > 0$ small, where allocation ω^* is given by Eq. (103) 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. (79) is $\omega^* = (0, \dots, \omega_i^*, \dots, 0)$, with $\omega_i^* = 1/\lambda_i$, where asset i is such that $\eta_i/\lambda_i = \max_{k=1, \dots, N} \eta_k/\lambda_k$ 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. (79) is equivalent to the following program

$$\max_{\omega \in Q} \omega^\top \left(\frac{\eta}{y} + \sigma \Sigma \right) - \frac{1}{2} \omega^\top \sigma \sigma^\top \omega, \quad (108)$$

where $\omega = \frac{z}{W}$. Without loss of generality, assume that at $y = y_1^*$ the solution of the optimization problem defined in Eq. (108) is $\omega_1^* = \frac{1}{\lambda_1}$ with $\lambda_1 \in \{-\lambda^-, \lambda^+\}$, and $\omega_k^* = 0$ for $k = 2, \dots, N$. We want to show that this is also the optimal solution for all $y < y_1^*$. The key is to observe that $y < y_1^*$

$$\max_{\omega \in Q} \omega^\top \left(\frac{\eta}{y} + \sigma \Sigma \right) - \frac{1}{2} \omega^\top \sigma \sigma^\top \omega \leq \max_{\omega \in Q} \left[\omega^\top \left(\frac{\eta}{y_1^*} + \sigma \Sigma \right) - \frac{1}{2} \omega^\top \sigma \sigma^\top \omega \right] + \max_{\omega \in Q} \omega^\top \left(\frac{\eta}{y} - \frac{\eta}{y_1^*} \right). \quad (109)$$

By assumption, the optimal solution to the optimization problem

$$\max_{\omega \in Q} \left[\omega^\top \left(\frac{\eta}{y_1^*} + \sigma \Sigma \right) - \frac{1}{2} \omega^\top \sigma \sigma^\top \omega \right], \quad (110)$$

is $(\frac{1}{\lambda_1}, 0, \dots, 0)$ and it turns out that for $y < y_1^*$ the optimal solution to the optimization problem

$$\max_{\omega \in Q} \omega^\top \left(\frac{\eta}{y} - \frac{\eta}{y_1^*} \right), \quad (111)$$

is also $(\frac{1}{\lambda_1}, 0, \dots, 0)$. 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

$$\begin{aligned}
\left(\theta + (\gamma - 1)(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2})\right) f(v) &= \frac{\gamma}{1 - \gamma} (f'(v))^{\frac{\gamma-1}{\gamma}} + f'(v) \\
&+ (B_K^{-1} + \gamma(\Sigma^\top \Sigma + \Theta^\top \Theta) - \gamma(I_K \Sigma)^\top I_K \Sigma + L_K) v f'(v) \\
&+ \frac{1}{2} (\Sigma^\top \Sigma + \Theta^\top \Theta + M_K^2 (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda \\
&\quad - 2M_K (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma) v^2 f''(v) \\
&- \frac{1}{2} ((I_K \eta)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \eta - L_K^2 (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda) \frac{(f'(v))^2}{f''(v)}
\end{aligned} \tag{112}$$

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

$$\Sigma^\top \Sigma + \Theta^\top \Theta - (I_K \Sigma)^\top I_K \Sigma + (M_K (I_K \sigma I_K^\top)^{-1} I_K \lambda + I_K \sigma \Sigma)^\top (M_K (I_K \sigma I_K^\top)^{-1} I_K \lambda + I_K \sigma \Sigma), \tag{113}$$

which is positive.

Deterministic income and general preferences.

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

$$\theta F = \max_{\frac{z}{W} \in Q} \tilde{u}(F_1) + (rW + Y)F_1 + mYF_2 + z^\top (\mu - r\bar{\mathbb{I}})F_1 + \frac{1}{2} z^\top \sigma \sigma^\top z W^2 F_{11}, \tag{114}$$

where \tilde{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 $e_i^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{\mathbb{I}})$.

Appendix F. Proof of Proposition 6

When $N = 2$ and $\Sigma = 0$, the program defined by Eq. (79) becomes

$$\max_{\omega \in Q} \omega^\top (\mu - r\bar{\mathbf{1}}) - \frac{y}{2} \omega^\top \sigma \sigma^\top \omega - \psi_2 (\omega^\top \lambda - 1), \quad (115)$$

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

$$\frac{z^*}{W} = \frac{(\sigma \sigma^\top)^{-1}}{y} (\mu - r\bar{\mathbf{1}} - \psi_2 \lambda), \quad (116)$$

and

$$\psi_2 = \frac{\lambda^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{\mathbf{1}}) - y}{\lambda^\top (\sigma \sigma^\top)^{-1} \lambda}. \quad (117)$$

The constraint starts binding at

$$y = y_B^* = \max_{\lambda \in \Lambda} \lambda^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{\mathbf{1}}), \quad (118)$$

so that

$$\lambda_B^* = \arg \max_{\lambda \in \Lambda} \lambda^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{\mathbf{1}}). \quad (119)$$

The covariance matrix is

$$\sigma \sigma^\top = \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}, \quad (120)$$

so that

$$(\sigma \sigma^\top)^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho \sigma_1 \sigma_2 \\ -\rho \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix}. \quad (121)$$

For $y \geq y_B^*$

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{y(1 - \rho^2)\sigma_1^2\sigma_2^2} [\sigma_2^2(\mu_1 - r) - \rho\sigma_1\sigma_2(\mu_2 - r)] \\ \frac{z_2^*}{W} &= \frac{1}{y(1 - \rho^2)\sigma_1^2\sigma_2^2} [\sigma_1^2(\mu_2 - r) - \rho\sigma_1\sigma_2(\mu_1 - r)], \end{aligned} \quad (122)$$

For $y \leq y_B^*$

$$\begin{aligned}\frac{z_1^*}{W} &= \frac{1}{\lambda_1 \left[\left(\frac{\sigma_1}{\lambda_1} \right)^2 + \left(\frac{\sigma_2}{\lambda_2} \right)^2 - 2\rho \frac{\sigma_1 \sigma_2}{\lambda_1 \lambda_2} \right]} \left[\left(\frac{\mu_1 - r}{\lambda_1} - \frac{\mu_2 - r}{\lambda_2} \right) \frac{1}{y} + \left(\frac{\sigma_2}{\lambda_2} \right)^2 - \rho \frac{\sigma_1 \sigma_2}{\lambda_1 \lambda_2} \right] \\ \frac{z_2^*}{W} &= \frac{1}{\lambda_2 \left[\left(\frac{\sigma_1}{\lambda_1} \right)^2 + \left(\frac{\sigma_2}{\lambda_2} \right)^2 - 2\rho \frac{\sigma_1 \sigma_2}{\lambda_1 \lambda_2} \right]} \left[\left(\frac{\mu_2 - r}{\lambda_2} - \frac{\mu_1 - r}{\lambda_1} \right) \frac{1}{y} + \left(\frac{\sigma_1}{\lambda_1} \right)^2 - \rho \frac{\sigma_1 \sigma_2}{\lambda_1 \lambda_2} \right].\end{aligned}\tag{123}$$

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 $\bar{\lambda}_1^* = \lambda^+$ if $\lambda_1^* = -\lambda^-$, $\bar{\lambda}_1^* = -\lambda^-$ if $\lambda_1^* = \lambda^+$.

General properties. For $y \leq y_B^*$, asset allocations (z_1, z_2) are given by Eq. (123) *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 $y = \bar{y}$, then it is optimal to hold only asset 1 for all $y \leq \bar{y}$ (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

$$\begin{aligned}y_{2,D}^L &= \left[\frac{\frac{\mu_1 - r}{\lambda_1^*} - \frac{\mu_2 - r}{\lambda^+}}{\left(\frac{\sigma_1}{\lambda_1^*} \right)^2 - \rho \frac{\sigma_1 \sigma_2}{\lambda_1^* \lambda^+}} \right]^+ \quad \text{and} \quad y_{2,D}^S = \left[\frac{\frac{\mu_1 - r}{\lambda_1^*} + \frac{\mu_2 - r}{\lambda^-}}{\left(\frac{\sigma_1}{\lambda_1^*} \right)^2 + \rho \frac{\sigma_1 \sigma_2}{\lambda_1^* \lambda^-}} \right]^+ \\ y_{1,R}^* &= \left[\frac{\frac{\mu_1 - r}{\lambda_1^*} - \frac{\mu_2 - r}{\lambda_2}}{-\left(\frac{\sigma_2}{\lambda_2} \right)^2 + \rho \frac{\sigma_1 \sigma_2}{\lambda_1^* \lambda_2}} \right]^+ \quad \text{and} \quad y_{1,D}^* = \left[\frac{\frac{\mu_1 - r}{\lambda_1^*} - \frac{\mu_2 - r}{\lambda_2}}{-\left(\frac{\sigma_2}{\lambda_2} \right)^2 + \rho \frac{\sigma_1 \sigma_2}{\lambda_1^* \lambda_2}} \right]^+, \end{aligned}\tag{124}$$

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. (122). 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. (122) it must be the case that $y_{1,R}^* = y_{1,D}^* = y_B^* = \frac{\mu_2 - r}{\lambda_2 \sigma_2^2}$. On $[0, y_{2,D}^*]$ asset 1 is held alone and on $[y_{2,D}^*, y_B^*]$ 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, $y \leq y_B^*$, 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 $(\lambda_{B,1}^*, \lambda_{B,2}^*) = (-\lambda^-, \lambda^+)$ so

$$y_B^* = (-\lambda^-, \lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}) > 0. \quad (125)$$

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

$$y_{1,D}^* = \frac{\frac{\mu_1 - r}{\lambda^-} + \frac{\mu_2 - r}{\lambda^+}}{\left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho \frac{\sigma_1}{\lambda^-} \frac{\sigma_2}{\lambda^+}}, \quad y_{1,R}^* = \frac{\frac{\mu_1 - r}{\lambda^+} - \frac{\mu_2 - r}{\lambda^+}}{-\left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho \frac{\sigma_1}{\lambda^+} \frac{\sigma_2}{\lambda^+}}, \quad y_{2,D}^* = \frac{\frac{\mu_1 - r}{\lambda^+} - \frac{\mu_2 - r}{\lambda^+}}{\left(\frac{\sigma_1}{\lambda^+}\right)^2 - \rho \frac{\sigma_1}{\lambda^+} \frac{\sigma_2}{\lambda^+}}. \quad (126)$$

and one can check that indeed, $0 < y_{2,D}^* < y_{1,R}^* < y_{1,D}^* < y_B^*$.

On $[y_{1,D}^*, y_B^*]$ asset 1 is held (short) and asset 2 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\lambda^- \left[\left(\frac{\sigma_1}{\lambda^-}\right)^2 + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + 2\rho \frac{\sigma_1}{\lambda^-} \frac{\sigma_2}{\lambda^+} \right]} \left[\left(-\frac{\mu_1 - r}{\lambda^-} - \frac{\mu_2 - r}{\lambda^+} \right) \frac{1}{y} + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + \rho \frac{\sigma_1}{\lambda^-} \frac{\sigma_2}{\lambda^+} \right] < 0 \\ \frac{z_2^*}{W} &= \frac{1}{\lambda^- \left[\left(\frac{\sigma_1}{\lambda^-}\right)^2 + \left(\frac{\sigma_2}{\lambda^+}\right)^2 + 2\rho \frac{\sigma_1}{\lambda^-} \frac{\sigma_2}{\lambda^+} \right]} \left[\left(\frac{\mu_2 - r}{\lambda^+} + \frac{\mu_1 - r}{\lambda^-} \right) \frac{1}{y} + \left(\frac{\sigma_1}{\lambda^-}\right)^2 + \rho \frac{\sigma_1}{\lambda^-} \frac{\sigma_2}{\lambda^+} \right] > 0, \end{aligned} \quad (127)$$

on $[y_{1,R}^*, y_{1,D}^*]$ only asset 2 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= 0 \\ \frac{z_2^*}{W} &= \frac{1}{\lambda^+}, \end{aligned} \quad (128)$$

on $[y_{2,D}^*, y_{1,R}^*]$ asset 1 is held (long) and asset 2 is held (long)

$$\begin{aligned} \frac{z_1^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \left[\frac{\mu_1 - \mu_2}{y} + \frac{\sigma_2}{\lambda^+} (\sigma_2 - \rho\sigma_1) \right] > 0 \\ \frac{z_2^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \left[\frac{\mu_2 - \mu_1}{y} + \frac{\sigma_1}{\lambda^+} (\sigma_1 - \rho\sigma_2) y \right] > 0, \end{aligned} \quad (129)$$

and finally on $[0, y_{2,D}^*]$ only asset 1 is held (long)

$$\begin{aligned}\frac{z_1^*}{W} &= \frac{1}{\lambda^+} \\ \frac{z_2^*}{W} &= 0.\end{aligned}\tag{130}$$

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

$$y_B^* = (\lambda^+, \lambda^+)^\top (\sigma\sigma^\top)^{-1} (\mu - r\mathbf{1}) > 0.\tag{131}$$

Margin coefficient $\bar{\lambda}_1^*$ is irrelevant. We have $0 < y_{2,D}^* < y_B^*$ with

$$y_{2,D}^* = \frac{\mu_1 - \mu_2}{\frac{\sigma_1}{\lambda^+} (\sigma_1 - \rho\sigma_2)}.\tag{132}$$

On $[y_{2,D}^*, y_B^*]$ asset 1 is held (long) and asset 2 is held (long)

$$\begin{aligned}\frac{z_1^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \left[\frac{\mu_1 - \mu_2}{y} + \frac{\sigma_2}{\lambda^+} (\sigma_2 - \rho\sigma_1) \right] > 0 \\ \frac{z_2^*}{W} &= \frac{1}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \left[\frac{\mu_2 - \mu_1}{y} + \frac{\sigma_1}{\lambda^+} (\sigma_1 - \rho\sigma_2) \right] > 0,\end{aligned}\tag{133}$$

and on $[0, y_{2,D}^*]$ only asset 1 is held (long)

$$\begin{aligned}\frac{z_1^*}{W} &= \frac{1}{\lambda^+} \\ \frac{z_2^*}{W} &= 0.\end{aligned}\tag{134}$$

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

$$y_{2,D}^* = y_B^* = \lambda^+ \frac{\mu_1 - r}{\sigma_1^2} > 0.\tag{135}$$

Only asset 1 is held (long)

$$\begin{aligned}\frac{z_1^*}{W} &= \frac{1}{\max\{y, y_B^*\}} \frac{\mu_1 - r}{\sigma_1^2} > 0 \\ \frac{z_2^*}{W} &= 0.\end{aligned}\tag{136}$$

Appendix G. Proof of Proposition 7

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

$$\frac{\eta_i}{\lambda_i^*} = \frac{\eta_j}{\lambda_j^*} > 0, \quad (137)$$

Assume that, for $y \in [y_{N,K}^-, y_{N,K}^+]$, it is optimal to hold the first K assets in such a way that the condition in Eq. (137) is satisfied for $k = 1, \dots, K$. The Lagrange multiplier given by Eq. (104) can be written as

$$\psi_K(y, \lambda^*) = \frac{\eta_1}{\lambda_1^*} - \frac{1 - (\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda^*} y, \quad (138)$$

which leads to the following optimal portfolio allocation

$$I_K \omega^* = (I_K \sigma \sigma^\top I_K^\top)^{-1} \left(\sigma \Sigma + \frac{1 - (\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda^*} \lambda^* \right). \quad (139)$$

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 $\{1, \dots, K\}^2$ we obtain that

$$I_K \omega^* = \frac{(I_K \sigma \sigma^\top I_K^\top)^{-1} \lambda^*}{(\lambda^*)^\top I_K^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda^*}, \quad (140)$$

which 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. (137) holds for $K = N$, then if at $y \in [y_{N,K}^-, y_{N,K}^+]$ 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 $K \in \{1, \dots, N\}$.

Step 2: If no security among the K assets held is first dropped out of the portfolio, reintegrating asset $i \in \{K + 1, \dots, N\}$ 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. (139) 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. (139), the expression for the asset allocation is, by assumption, optimal for values of the lifetime relative risk aversion y in $[y_{N,K}^-, y_{N,K}^+]$ 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. (139), 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

$$d\widehat{B}_t = (r + a)\widehat{B}_t dt, \quad (141)$$

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

$$d\widehat{S}_t = I_{\widehat{S}_t}(\mu + b)dt + I_{\widehat{S}_t}\sigma dw_t, \quad (142)$$

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

$$d\widehat{P}_t = I_{\widehat{P}_t}\widehat{\mu}dt + I_{\widehat{P}_t}\widehat{\sigma}dw_t^Y, \quad (143)$$

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 $\kappa = -\widehat{\sigma}^{-1}(\widehat{\mu} - r\mathbf{1})$.

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

$$d\pi_t^{a,b,\kappa} = \pi_t^{a,b,\kappa} \left(-(r + a_t)dt - (\sigma^{-1} (b_t - a_t\mathbf{1} + \mu - r\mathbf{1}))^\top dw_t + \kappa_t^\top dw_t^Y \right), \quad (144)$$

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

$$e(a, b, \kappa) = \sup_{\frac{z}{z+x} \in Q} -ax - b^\top z. \quad (145)$$

The effective domain \mathcal{N} is defined by

$$\mathcal{N} = \{(a, b, \kappa) \in \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^M, e(a, b, \kappa) < \infty\}. \quad (146)$$

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

$$\mathcal{N} = \{(a, b, \kappa) \in \mathbb{R}_+ \times \mathbb{R}_+^N \times \mathbb{R}^M, \kappa^+ a \leq b_i \leq \kappa^- a, i = 1, \dots, N\}, \quad (147)$$

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 $a \geq 0$, $b_i \geq 0, i = 1, \dots, N$. If $z_i \geq 0, i = 1, \dots, N$ we have

$$-ax - b^\top z = -a \left(x + (1 - \lambda^+) \sum_{i=1}^N z_i \right) - \sum_{i=1}^N (b_i - (1 - \lambda^+)a) z_i. \quad (148)$$

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

$$-ax - b^\top z = -a \left(x + (1 + \lambda^-) \sum_{i=1}^N z_i \right) - \sum_{i=1}^N (b_i - (1 + \lambda^-)a) z_i. \quad (149)$$

Since $z_i \leq 0, i = 1, \dots, N$, we must have $b_i - (1 + \lambda^-)a \leq 0, i = 1, \dots, N$. 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), is equivalent to

$$\begin{aligned} F(W_0, Y_0) &= \max_c E_0 \left[\int_0^\infty u(c_s) e^{-\theta s} ds \right] \\ \text{such that } E_0 \left[\int_0^\infty \pi_s^* c_s ds \right] &= W_0 + E_0 \left[\int_0^\infty \pi_s^* Y_s ds \right], \end{aligned} \quad (150)$$

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

Appendix I. Dual Approach

To ensure that the optimization problem, given by Eqs. (9), and (150) are equivalent, it is enough to determine the saddle point $(c^*, \phi^*, (a^*, b^*, \kappa^*))$ of the functional

$$\mathcal{L}(c, \psi, (a, b, \kappa)) = E_0 \left[\int_0^\infty u(c_s) e^{-\theta s} ds \right] - \phi \left(E_0 \left[\int_0^\infty \pi_s^{a, b, \kappa} (c_s - Y_s) ds \right] - W_0 \right). \quad (151)$$

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

$$E_0 \left[\int_0^\infty \pi_s^{a, b, \kappa} (I(\phi^* \pi_s^{a, b, \kappa} e^{\theta s}) - Y_s) ds \right] = W_0, \quad (152)$$

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

$$X_t^{a, b, \kappa} = \phi^* \pi_t^{a, b, \kappa} e^{\theta t}. \quad (153)$$

The dual value function J is given by

$$J(X_0, Y_0) = \min_{(a,b,\kappa) \in \mathcal{N}} E_0 \left[\int_0^\infty (\tilde{u}(X_s^{a,b,\kappa}) + X_s^{a,b,\kappa} Y_s) e^{-\theta s} ds \right], \quad (154)$$

where $\tilde{u}(X) = \max_{c \geq 0} u(c) - Xc$ 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

$$\tilde{u}(X) = \begin{cases} \frac{\gamma X^{\frac{\gamma-1}{1-\gamma}}}{1-\gamma} & , \quad \gamma \neq 1, \\ -\ln X - 1 & , \quad \gamma = 1. \end{cases} \quad (155)$$

Properties of the dual value function

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

$$W = -J_1(X, Y) \text{ and } X = F_1(W, Y). \quad (156)$$

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:

$$\begin{aligned} \theta J = & \frac{\gamma X^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + XY + (\theta - r)XJ_1 + mYJ_2 + \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} Y^2 J_{22} - \frac{\Sigma^\top \Sigma}{2} \frac{J_{12}^2}{J_{11}} \\ & + \min_{\kappa \in \mathbb{R}^M} \left\{ \frac{\kappa^\top \kappa}{2} X^2 J_{11} + \kappa^\top \Theta XY J_{12} \right\} \\ & + \min_{(a,b) \in \mathcal{N}_{a,b}} \left\{ -aXJ_1 + \frac{X^2}{2} \left(b + \mu - (r + a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right)^\top (\sigma \sigma^\top)^{-1} \left(b + \mu - (r + a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right) J_{11} \right\} \end{aligned} \quad (157)$$

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

$$\begin{aligned} \theta J = & \frac{\gamma X^{\frac{\gamma-1}{\gamma}}}{1-\gamma} + XY + (\theta - r)XJ_1 + mYJ_2 + \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} Y^2 \left(J_{22} - \frac{J_{12}^2}{J_{11}} \right) \\ & + \min_{(a,b) \in \mathcal{N}_{a,b}} \left\{ -aXJ_1 + \frac{X^2}{2} \left(b + \mu - (r+a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right)^\top (\sigma \sigma^\top)^{-1} \left(b + \mu - (r+a)\bar{1} - \frac{\sigma \Sigma Y J_{12}}{X J_{11}} \right) J_{11} \right\} \end{aligned} \quad (158)$$

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

$$\min_{(a,b) \in \mathcal{N}_{a,b}} a + \frac{1}{2y} (\eta + y\sigma\Sigma + b - a\bar{1})^\top (\sigma\sigma^\top)^{-1} (\eta + y\sigma\Sigma + b - a\bar{1}). \quad (159)$$

The minimization problem given by Eq. (159) and the maximization problem, given by Eq. (16), 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. (159) is

$$\begin{aligned} a^* = \psi_K = & \frac{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \eta - (1 - (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \sigma \Sigma) y}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K \lambda} \\ b_k^* = & (1 - \lambda_k) a^*, \quad k = 1, \dots, K, \end{aligned} \quad (160)$$

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

$$I_K \frac{z^*}{W} = \frac{(I_K \sigma \sigma^\top I_K^\top)^{-1}}{y} I_K (\eta - y\sigma\Sigma + b^* - a^*\bar{1}). \quad (161)$$

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 $k = K + 1, K + 2, \dots, N$.

Remark. Observe that the right hand side of Eq. (150) 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

$$\widehat{S}_{P,k} = \frac{\mu_k + b_k^* - (r + a^*)}{\sigma_k}. \quad (162)$$

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 $k = 1, \dots, N$ so indeed

$$\left| \widehat{S}_{P,k} \right| < |S_{P,k}|, \quad (163)$$

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 et al. (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

$$\begin{aligned} \max_{(\omega, \omega_M) \in \mathbb{R}^{N+1}} & (\omega, \omega_M)^\top (\eta + y\sigma\Sigma, \eta_M + y\pi^\top\sigma\Sigma) - \frac{y}{2} (\omega, \omega_M)^\top V(\omega, \omega_M) \\ \text{s.t.} & (\omega^+)^\top \lambda^+ + (\omega^-)^\top \lambda^- + \lambda_M^+ \omega_M^+ + \lambda_M^- \omega_M^- \leq 1, \end{aligned} \quad (164)$$

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

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

The Lagrangian of the maximization problem is

$$\begin{aligned} L = & (\omega, \omega_M)^\top (\eta + y\sigma\Sigma, \eta_M + y\pi^\top\sigma\Sigma) - \frac{y}{2} (\omega, \omega_M)^\top V(\omega, \omega_M) \\ & - \psi [(\omega, \omega_M)^\top (\lambda, \lambda_M) - 1] + \left(\frac{\omega}{\lambda}, \frac{\omega_M}{\lambda_M}\right)^\top (\varphi, \varphi_M), \end{aligned} \quad (165)$$

which leads to the following optimal condition

$$(\eta + y\sigma\Sigma, \eta_M + y\pi^\top\sigma\Sigma) - yV(\omega^*, \omega_M^*) - \psi(\lambda, \lambda_M) + \left(\frac{\varphi}{\lambda}, \frac{\varphi_M}{\lambda_M}\right) = 0, \quad (166)$$

where $\psi \geq 0$, with $\psi [(\omega^*, \omega_M^*)^\top (\lambda, \lambda_M) - 1] = 0$ 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 $k = 1, \dots, N$. We note that, by convention, if $(x, y) \in \mathbb{R}^N \times \mathbb{R}^N$, then $z^\top = \left(\frac{x}{y}\right)^\top = \left(\frac{x_1}{y_1}, \frac{x_2}{y_2}, \dots, \frac{x_N}{y_N}\right)$.

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

$$\psi(\lambda_M - \pi^\top \lambda) - \frac{\varphi_M}{\lambda_M} + \pi^\top \frac{\varphi}{\lambda} = 0. \quad (167)$$

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

$$\omega^* + \omega_M^* \pi = \frac{1}{y} (\sigma\sigma^\top)^{-1} (\eta + \frac{\varphi}{\lambda} + y\sigma\Sigma - \psi\lambda). \quad (168)$$

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

$$\omega^* + \omega_M^* \pi = \frac{1}{y} (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}) = \frac{\varpi \pi}{y} \in \mathbb{R}_+^N. \quad (169)$$

Let $(\lambda^+)^\top = (\lambda_1^+, \lambda_2^+, \dots, \lambda_N^+)^\top$ 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: $\lambda_M^+ = \pi^\top \lambda^+$.

From Eq. (169), at $y = y_B^*$ we obtain that

$$1 = \frac{(\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}})}{y_B^*} + (\lambda_M^+ - \pi^\top \lambda^+) \omega_M^*, \quad (170)$$

so that

$$y_B^* = (\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}). \quad (171)$$

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: $\lambda_M^+ > \pi^\top \lambda^+$.

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

$$1 = \frac{(\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1})}{y_B^*} + (\lambda_M^+ - \pi^\top \lambda^+) \omega_M^*. \quad (172)$$

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 $\lambda_M^+ > \pi^\top \lambda^+$, choosing $\omega_M^* = 0$ is optimal and

$$y_B^* = (\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}). \quad (173)$$

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: $\lambda_M^+ < \pi^\top \lambda^+$.

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 $\lambda_M^+ < \pi^\top \lambda^+$, 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

$$y_B^* = \lambda_M^+ \bar{1}^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) = \frac{\lambda_M^+}{\varpi} \quad (174)$$

Observe that

$$y_B^* < (\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) \quad (175)$$

as

$$\frac{\lambda_M^+}{\pi^\top \lambda^+} (\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) = \lambda_M^+ \bar{1}^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) \quad (176)$$

and, by assumption

$$\frac{\lambda_M^+}{\pi^\top \lambda^+} < 1. \quad (177)$$

This strategy is feasible since at $y = y_B^*$, if $\omega_M^* = \frac{1}{\lambda_M^+}$, from Eq. (169) we have

$$\omega^* = \left(\frac{1}{y_B^*} - \frac{\varpi}{\lambda_M^+} \right) (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) = 0 \in \mathbb{R}_+^N. \quad (178)$$

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*

$$V_K = \begin{bmatrix} J_K \sigma \sigma^\top J_K^\top & J_K \sigma \sigma^\top \pi \\ \pi^\top \sigma \sigma^\top J_K^\top & \pi^\top (\sigma \sigma^\top) \pi \end{bmatrix}. \quad (179)$$

Then, we have that

$$V_K^{-1} [(J_K(\mu - r\bar{1}), \mu_M - r)] = (0, 0, \dots, 0, \varpi^{-1}). \quad (180)$$

Proof of Lemma J.1.

First, notice that

$$V_K = \begin{bmatrix} J_K \sigma \sigma^\top J_K^\top & \varpi J_K(\mu - r\bar{1}) \\ \varpi(\mu - r\bar{1})^\top J_K^\top & \varpi^2(\mu - r\bar{1})^\top (\sigma \sigma^\top)^{-1}(\mu - r\bar{1}) \end{bmatrix}. \quad (181)$$

Set

$$d = \varpi^2 \left[(\mu - r\bar{1})^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) - (\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) \right] > 0. \quad (182)$$

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

$$\begin{bmatrix} (J_K \sigma \sigma^\top J_K^\top)^{-1} \left(1 + \frac{\varpi^2}{d} J_K (\mu - r\bar{1}) (\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} \right) & -\frac{\varpi}{d} (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) \\ -\frac{\varpi}{d} (\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} & \frac{1}{d} \end{bmatrix}. \quad (183)$$

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

$$\begin{aligned} & I_{K,K+1} V_K^{-1} [(J_K (\mu - r\bar{1}), \mu_M - r)] \\ &= \frac{\varpi^2}{d} \left[\frac{d}{\varpi^2} + (\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) - (\mu - r\bar{1})^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) \right] \\ & \quad \times (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) \\ &= 0 \text{ (by definition of } d \text{ in Eq. 182)}. \end{aligned} \quad (184)$$

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

$$\begin{aligned} e_{K+1}^\top V_K^{-1} [(J_K (\mu - r\bar{1}), \mu_M - r)] &= -\frac{\varpi}{d} (\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) + \frac{1}{d} \pi^\top (\mu - r\bar{1}) \\ &= \frac{\varpi}{d} \left[-(\mu - r\bar{1})^\top J_K^\top (J_K \sigma \sigma^\top J_K^\top)^{-1} J_K (\mu - r\bar{1}) \right. \\ & \quad \left. + (\mu - r\bar{1})^\top (\sigma \sigma^\top)^{-1} (\mu - r\bar{1}) \right] \\ &= \varpi^{-1}. \end{aligned} \quad (185)$$

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: $\lambda_M^+ = \pi^\top \lambda^+$.

For y slightly below y_B^* , we have

$$\omega^* + \omega_M^* \pi = \frac{(\sigma \sigma^\top)^{-1} [\mu - r\bar{1} - \psi_N(\lambda^+, y) \lambda^+]}{y}, \quad (186)$$

where

$$\psi_N(\lambda^+, y) = \frac{(\lambda^+)^\top(\sigma\sigma^\top)^{-1}(\mu - r\bar{\mathbf{1}}) - y}{(\lambda^+)^\top(\sigma\sigma^\top)^{-1}\lambda^+} > 0 \quad (187)$$

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: $\lambda_M^+ > \pi^\top\lambda^+$.

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: $\lambda_M^+ < \pi^\top\lambda^+$.

If

$$\frac{\mu_M - r}{\lambda_M^+} > \max_{k \in \{1, \dots, N\}} \frac{\mu_k - r}{\lambda_k^+}, \quad (188)$$

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

$$\frac{\mu_M - r}{\lambda_M^+} < \max_{k \in \{1, \dots, N\}} \frac{\mu_k - r}{\lambda_k^+}. \quad (189)$$

All assets cannot be re-integrated into the portfolio at $y = y_B^* - \varepsilon$, $\varepsilon > 0$, otherwise the condition $\psi(\lambda_M^+ - \pi^\top\lambda^+) - \frac{\varphi_M}{\lambda_M^+} + \pi^\top\frac{\varphi}{\lambda^+} = 0$ 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

$$\begin{aligned} (J_K \omega^*, \omega_M^*) &= \frac{V_K^{-1} [(J_K(\mu - r\bar{1}), \mu_M - r) - \psi_{K+1}(y)(J_K \lambda^+, \lambda_M^+)]}{y} \\ &= \frac{1}{y} \left(-\psi_{K+1}(y) I_{K,K+1} V_K^{-1} (J_K \lambda^+, \lambda_M^+), \varpi^{-1} - \psi_{K+1}(y) e_{K+1}^\top V_K^{-1} (J_K \lambda^+, \lambda_M^+) \right), \end{aligned} \quad (190)$$

with

$$\begin{aligned} \psi_{K+1}(y) &= \frac{(J_K \lambda^+, \lambda_M^+)^\top V_K^{-1} (J_K(\mu - r\bar{1}), \mu_M - r) - y}{(J_K \lambda^+, \lambda_M^+)^\top V_K^{-1} (J_K \lambda^+, \lambda_M^+)} \\ &= \frac{y_B^* - y}{(J_K \lambda^+, \lambda_M^+)^\top V_K^{-1} (J_K \lambda^+, \lambda_M^+)}. \end{aligned} \quad (191)$$

We note that since $I_{K,K+1} V_K^{-1} I_K^\top J_K(\mu - r\bar{1}) = 0$, 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 $\varpi^{-1} - \psi_{K+1}(y_{M,D}^*) e_{K+1}^\top V_K^{-1} (J_K \lambda^+, \lambda_M^+) = 0$, which implies that

$$y_{M,D}^* = -\frac{(J_K \lambda^+)^\top I_{K,K+1} V_K^{-1} (J_K \lambda^+, \lambda_M^+)}{\varpi} > 0. \quad (192)$$

- 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 $\lambda_M^+ = \pi^\top \lambda^+$.

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

$$y_B^* = (\lambda^+)^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{1}), \quad (193)$$

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

$$\omega_k^* + \omega_M^* \pi_k = \frac{\mu_k - r}{y \sigma_k^2}, \quad (194)$$

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

$$(\lambda^+)^\top \omega^* + \lambda_M \omega_M^* = (\lambda^+)^\top (\omega^* + \omega_M^* \pi) \quad (195)$$

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

$$\omega_k^* + \omega_M^* \pi_k = \frac{1}{y \sigma_k^2} (\mu_k - r - \psi_N(\lambda^+, y) \lambda_k^+) \quad (196)$$

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

$$\omega_k^* + \omega_M^* \pi_k = \frac{\alpha_k}{y \lambda_k^+ \alpha^\top \bar{1}} (y - \alpha^\top (\xi - \xi_k \bar{1})). \quad (197)$$

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

$$y_B^* = \alpha^\top \xi. \quad (198)$$

Next, without loss of generality, assume that $0 < \xi_N < \xi_{N-1} < \dots < \xi_1$. From Eq. (197), 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

$$y_{N,D}^* = \alpha^\top (\xi - \xi_N \bar{1}). \quad (199)$$

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

$$y_{K,D}^* = (I_K \alpha)^\top [I_K (\xi - \xi_K \bar{\mathbf{1}})], \quad (200)$$

and by convention, set $y_{N+1,D}^* = y_B^*$; observe that $0 = y_{1,D}^* < y_{2,D}^* < \dots < y_{N+1,D}^*$. For $k = 1, \dots, N-1$, we have $K \in \{1, \dots, N\}$ with

$$\omega_k^* = \frac{\alpha_k [y - y_{K,D}^*]^+}{y \lambda_k^+ (I_K \alpha)^\top \bar{\mathbf{1}}}. \quad (201)$$

Thus, there are exactly $N+1$ regions: if $K = 1, \dots, N-1$, 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, $y \in [y_{N-1,D}^*, y_{N,D}^*]$, 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

$$\omega^* = \left(\frac{1}{y} - \varpi \omega_M^* \right) \frac{(\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}})}{y}. \quad (202)$$

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 $e_i^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}})$, 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 $(\lambda_M - \pi^\top \lambda)\omega_M^*$ 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

$$y_B^* = \begin{cases} \lambda^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}), & \text{if } \lambda_M^+ - \pi^\top \lambda > 0, \text{ when } \omega_M^* = 0 \text{ is optimal,} \\ \lambda_M^+ \bar{\mathbf{1}}^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}), & \text{if } \lambda_M^+ - \pi^\top \lambda = 0, \text{ when } \omega_M^* = \left[0, \frac{1}{\lambda_M^+}\right] \text{ is optimal} \\ \lambda_M^+ \bar{\mathbf{1}}^\top (\sigma \sigma^\top)^{-1} (\mu - r \bar{\mathbf{1}}), & \text{if } \lambda_M^+ - \pi^\top \lambda < 0, \text{ when } \omega_M^* = \frac{1}{\lambda_M^+} \text{ is optimal and } \omega^* = 0. \end{cases} \quad (203)$$

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

$$\frac{\mu_M - r}{\lambda_M^+} > \max_{k \in \{1, \dots, N\}} \max_{\lambda_k \in \{-\lambda_k^-, \lambda_k^+\}} \frac{\mu_k - r}{\lambda_k}, \quad (204)$$

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

$$\frac{\mu_M - r}{\lambda_M^+} < \max_{k \in \{1, \dots, N\}} \max_{\lambda_k \in \{-\lambda_k^-, \lambda_k^+\}} \frac{\mu_k - r}{\lambda_k}. \quad (205)$$

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

$$v + B = Af'(v)^{-\frac{1}{\gamma}} + Kf'(v)^{\frac{\beta-1}{\gamma}}. \quad (206)$$

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

$$z^* = z^f - \beta K \frac{(\sigma\sigma^\top)^{-1}\eta}{\gamma} f'(v)^{\frac{\beta-1}{\gamma}} Y. \quad (207)$$

When $e_i^\top(\sigma\sigma^\top)^{-1}\eta > 0 (< 0)$, 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. (206) yields

$$\frac{f'(v)}{f''(v)} = -\frac{A}{\gamma} f'(v)^{-\frac{1}{\gamma}} + \frac{\beta-1}{\gamma} K f'(v)^{\frac{\beta-1}{\gamma}} < 0. \quad (208)$$

From Eqs. (208) and (206) it is easy to check that the margin requirement is not binding for $f'(v)^{\frac{1}{\gamma}} \leq Z^*$, for some $0 < Z^* < \widehat{Z}$ where $\widehat{Z} = \beta A / ((\beta - 1)B)$. Then, we have

$$\begin{aligned} \frac{\partial z^*}{\partial Y} &= (\sigma\sigma^\top)^{-1}\eta \left(B - \beta K f'(v)^{\frac{\beta-1}{\gamma}} \left(1 - \frac{\beta-1}{\gamma} \frac{v f''(v)}{f'(v)} \right) \right) \\ &= \frac{(\sigma\sigma^\top)^{-1}\eta}{A - (\beta-1)K f'(v)^{\frac{\beta}{\gamma}}} \left(AB + (\beta-1)^2 B K f'(v)^{\frac{\beta}{\gamma}} - \beta^2 A K f'(v)^{\frac{\beta-1}{\gamma}} \right). \end{aligned} \quad (209)$$

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

$$h(Z) = AB + (\beta-1)^2 B K Z^\beta - \beta^2 A K Z^{\beta-1}. \quad (210)$$

h is a smooth function with $h'(Z) = \beta(\beta-1)^2 B K Z^{\beta-2} (Z - \widehat{Z}) < 0$, 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 $Z \leq Z^*$ we have $(\lambda_B^*)^\top z^* \leq W$ or, equivalently, using the expression of z^*

$$v \left(1 - \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} (\mu - r\bar{1})}{\gamma} \right) \geq \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \eta}{\gamma} (B - \beta K f'(v)^{\frac{\beta-1}{\gamma}}). \quad (211)$$

Using Eq. (206) we obtain that for all Z in $[0, Z^*]$

$$\begin{aligned} K Z^\beta &\geq \vartheta(Z - \bar{Z}) \\ K (Z^*)^\beta &= \vartheta(Z^* - \bar{Z}), \end{aligned} \quad (212)$$

where

$$\begin{aligned} \bar{Z} &= \frac{1 - \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} (\mu - r\bar{1})}{\gamma} A}{1 - (\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \sigma \Sigma B} > 0 \\ \vartheta &= \frac{B (1 - (\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \sigma \Sigma)}{1 - (\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \sigma \Sigma + (\beta-1) \frac{(\lambda_B^*)^\top (\sigma\sigma^\top)^{-1} \eta}{\gamma}} > 0. \end{aligned} \quad (213)$$

Finally, we have

$$h(Z^*) = \frac{B}{Z^*} (\beta \bar{Z} - (\beta-1) Z^*). \quad (214)$$

It remains to show that

$$Z^* \leq \beta \bar{Z} / (\beta-1). \quad (215)$$

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

$$x^\beta \geq \frac{x - x^*}{1 - x^*}. \quad (216)$$

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

$$x^* \geq (\beta - 1)/\beta, \quad (217)$$

or, equivalently,

$$\beta \leq \frac{1}{(1 - x^*)}. \quad (218)$$

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

$$f(x) = x^\beta - \frac{x - x^*}{1 - x^*}. \quad (219)$$

Observe that

$$\begin{aligned} f(0) &= \frac{x^*}{1 - x^*} > 0 \\ f(1) &= 0 \\ f'(x) &= \beta x^{\beta-1} - (1 - x^*)^{-1}. \end{aligned} \quad (220)$$

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. (212) f is non-negative on $[0, 1]$. Thus, we must have $\beta \leq 1/(1 - x^*)$ or, equivalently,

$$Z^* \leq \frac{\beta \bar{Z}}{\beta - 1}. \quad (221)$$

It follows that $h(Z^*) \geq 0$ 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

$$\frac{z^*}{W} = (\sigma \sigma^\top)^{-1} \sigma \Sigma + \frac{(\sigma \sigma^\top)^{-1} \eta}{y}, \quad (222)$$

we deduce that

$$\frac{\partial}{\partial Y} \left(\frac{1}{y} \right) \geq 0, \quad (223)$$

which implies that

$$\frac{\partial y}{\partial Y} \leq 0. \quad (224)$$

Furthermore, since

$$\frac{\partial y}{\partial Y} = -v \frac{\partial y}{\partial v}, \quad (225)$$

we find that

$$\frac{\partial y}{\partial v} \geq 0. \quad (226)$$

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

$$c^* = Y f'(v)^{-\frac{1}{\gamma}}, \quad (227)$$

so

$$\frac{\partial c^*}{\partial W} = -\frac{f''(v) f'(v)^{-\frac{1}{\gamma}-1}}{\gamma} > 0. \quad (228)$$

Then

$$\frac{\partial c^*}{\partial Y} = \frac{f'(v)^{-\frac{1}{\gamma}}}{\gamma} (\gamma - y). \quad (229)$$

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 et al. (1997). In Proposition 1 of their paper, these authors establish that

$$\lim_{v \downarrow 0} f'(v) \quad (230)$$

exists, is positive and finite. They also show that

$$\limsup_{v \downarrow 0} -v f''(v) = 0. \quad (231)$$

Since

$$0 < -v f''(v) \leq \sup_{0 < x \leq v} -x f''(x), \quad (232)$$

it follows that

$$\lim_{v \downarrow 0} -v f''(v) = 0. \quad (233)$$

Hence, we have

$$\lim_{v \downarrow 0} -\frac{v f''(v)}{f'(v)} = 0. \quad (234)$$

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

$$f(v) \underset{0}{\sim} d_0 + v - d_1 v^{\frac{3}{2}} + d_2 v^2 + o(v^2), \quad (235)$$

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

$$\frac{\gamma}{1-\gamma} (f'(v))^{\frac{\gamma-1}{\gamma}} + f'(v) \quad (236)$$

achieves its maximum value for $v = 0$. Using the Hamilton-Jacobi-Bellman Eq. (112) for $K = 1$ and identifying coefficients, we obtain

$$f(0) = d_0 = \frac{1}{(1-\gamma) \left(\theta + (\gamma-1) \left(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} \right) \right)} > 0, \quad (237)$$

and

$$\theta + (\gamma-1) \left(m - \gamma \frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2} \right) = \frac{9}{8\gamma} d_1^2 + (r - m + \gamma(\Sigma^\top \Sigma + \Theta^\top \Theta) + \frac{\eta_1}{\lambda_1^*}). \quad (238)$$

It follows that

$$d_1 = \frac{2\sqrt{2\gamma(\theta + \gamma(m - (\gamma+1)\frac{\Sigma^\top \Sigma + \Theta^\top \Theta}{2})) - (r + \eta_1/\lambda_1^*)}}{3} > 0. \quad (239)$$

This implies that

$$c^* \underset{0}{\sim} Y \quad (240)$$

and,

$$y = -\frac{vf''(v)}{f'(v)} \underset{0}{\sim} \frac{3d_1v^{\frac{1}{2}}}{4}. \quad (241)$$

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

$$dX_t^{a^*, b^*, 0} = X_t^{a^*, b^*, 0}(-(r + a_t^*)dt + (\zeta_t^{a^*, b^*})^\top dw_t), \quad (242)$$

where

$$\zeta_t^{a^*, b^*} = -\sigma^{-1} (b_t^* - a_t^* \bar{1} + \mu - r \bar{1}). \quad (243)$$

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

$$\frac{dc_t^*}{c_t^*} = \left(\frac{r + a_t^* - \theta}{RR(c_t^*)} + \frac{1}{2} \frac{RP(c_t^*)}{(RR(c_t^*))^2} \left\| \zeta_t^{a^*, b^*} \right\|^2 \right) dt + \frac{(\zeta_t^{a^*, b^*})^\top}{RR(c_t^*)} dw_t, \quad (244)$$

where

$$RR(c) = -\frac{cu''(c)}{u'(c)} \quad (245)$$

is the relative risk aversion ratio and

$$RP(c) = -\frac{cu'''(c)}{u''(c)} \quad (246)$$

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 $t \geq 0$, $\left\| \zeta_t^{a^*, b^*} \right\|^2 \leq \|\zeta^{0,0}\|^2$. 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

$$\begin{aligned} b^* &= (\bar{1} - \lambda)a^* \\ a^* &= \frac{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K (\mu - r \bar{1}) - y}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} (I_K \lambda)} > 0, \end{aligned} \quad (247)$$

so that

$$\|\zeta^{a^*, b^*}\|^2 = (I_K(\mu - r\bar{1} - \lambda a^*))^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K(\mu - r\bar{1} - \lambda a^*). \quad (248)$$

Hence for all $\lambda \in \Lambda$

$$\frac{\partial}{\partial y} \|\zeta^{a^*, b^*}\|^2 = -2 \frac{\partial a^*}{\partial y} (I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} I_K(\mu - r\bar{1} - \lambda a^*) = \frac{2}{(I_K \lambda)^\top (I_K \sigma \sigma^\top I_K^\top)^{-1} (I_K \lambda)} > 0. \quad (249)$$

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

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). 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 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, we can reduce the number of state variables after scaling by income Y_t and obtain the following Bellman equation at $t = 0, \dots, T - 1$:

$$\begin{aligned} f_t(v_t) &= \max_{q_t, \omega_t} u(q_t) + \beta E_t [g_t^{1-\gamma} f_{t+1}(v_{t+1})] \\ \text{s.t.} \quad & v_{t+1} = g_t^{-1}(v_t + 1 - q_t) \left(\sum_{i=1}^N \omega_{i,t} R_{i,t}^e + R^f \right) \\ & \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- \leq 1 \\ & f_T = \phi_\tau \frac{(v_T + 1)^{1-\gamma}}{1-\gamma} \end{aligned} \quad (250)$$

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

$$\begin{aligned} \phi_\tau &= \left[\frac{1 - (\beta\alpha)^{1/\gamma}}{1 - (\beta\alpha)^{(\tau+1)/\gamma}} \right]^{-\gamma}, \\ \alpha &= E \left[\left(\sum_{i=1}^N \omega_i^* R_i^e + R^f \right)^{1-\gamma} \right] \end{aligned} \quad (251)$$

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

N.2. Solution Methodology

To solve the problem described in Eq. (250), we extend the method proposed by Brandt et al. (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. (250) by defining a new variable, total investment I_t :

$$I_t = v_t - q_t + 1. \quad (252)$$

At the optimal value of consumption, q_t^* , Eq. (252) 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. (250) into two subproblems:

[Portfolio Optimization]

$$\begin{aligned}
f_t^p(I_t) &= \max_{\omega_t} \beta E_t [g_t^{1-\gamma} f_{t+1}(v_{t+1})], \quad t = 0, \dots, T-1 \\
\text{s.t.} \quad v_{t+1} &= g_t^{-1} I_t \left(\sum_{i=1}^N \omega_{i,t} R_{i,t}^e + R^f \right) \\
\lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- &\leq 1
\end{aligned} \tag{253}$$

[Consumption Optimization]

$$f_t(v_t) = \max_{q_t} u(q_t) + f_t^p(v_t - q_t + 1), \quad t = 0, \dots, T-1, \tag{254}$$

where $f^p(\cdot)$ is the value function of the portfolio optimization problem in Eq. (253). Given the separation of consumption and portfolio optimization, we use the following algorithm to solve the problem in Eq. (250):

Algorithm

Step 1: Set the terminal condition at time T .

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

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 $I_t^i, i = 1, \dots, n_g$:

Step 2.2.1: [Portfolio optimization] given I_t^i , find $\omega_t^*(I_t^i)$ by solving Eq. (253).

Step 2.2.2: [Consumption optimization] given $\{I_t^i, \omega_t^*(I_t^i)\}$, find $q_t^*(I_t^i)$ by solving Eq. (254).

Step 2.2.3: Recover state variable v_t at grid point i by $v_t^i = I_t^i + q_t^*(I_t^i) - 1$.

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 $I_t^i, i = 1, \dots, n_g$, we want to optimize over ω_t by solving Eq. (253). 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:

$$\begin{aligned}
f_t^p(I_t) &= \max_{\omega_t^+, \omega_t^-} \beta E_t [g_t^{1-\gamma} f_{t+1}(v_{t+1})] \\
s.t. \quad v_{t+1} &= g_t^{-1} I_t \left[\sum_{i=1}^N (\omega_{i,t}^+ - \omega_{i,t}^-) R_{i,t}^e + R^f \right] \\
\lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- &\leq 1 \\
\omega_{i,t}^+, \omega_{i,t}^- &\geq 0, i = 1, \dots, N
\end{aligned} \tag{255}$$

Notice that to maintain equivalence between Eqs. (253) and (255) we also need the constraints $\omega_{i,t}^+ \omega_{i,t}^- = 0$ for $i = 1, \dots, N$, in Eq. (255). 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. (255) are given by:

Lagrangian

$$\begin{aligned}
\mathcal{L}^p(\omega_t^+, \omega_t^-, l_t^+, l_t^-, l_t^m) &= \beta E_t [g_t^{1-\gamma} f_{t+1}(v_{t+1})] + \sum_{i=1}^N l_{i,t}^+ \omega_{i,t}^+ + \sum_{i=1}^N l_{i,t}^- \omega_{i,t}^- \\
&\quad + l_t^m \left(1 - \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ - \lambda^- \sum_{i=1}^N \omega_{i,t}^- \right)
\end{aligned} \tag{256}$$

KKT conditions

$$\begin{aligned}
0 &= \beta I_t E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} R_{i,t}^e \right\} + l_{i,t}^+ - l_t^m \lambda^+, i = 1, \dots, N && \text{FOCs} \\
0 &= -\beta I_t E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} R_{i,t}^e \right\} + l_{i,t}^- - l_t^m \lambda^-, i = 1, \dots, N && \text{FOCs} \\
0 &= l_{i,t}^+ \omega_{i,t}^+, i = 1, \dots, N && \text{Complementarity} \\
0 &= l_{i,t}^- \omega_{i,t}^-, i = 1, \dots, N && \text{Complementarity} \\
0 &= l_t^m \left(1 - \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ - \lambda^- \sum_{i=1}^N \omega_{i,t}^- \right) && \text{Complementarity} \\
1 &\geq \lambda^+ \sum_{i=1}^N \omega_{i,t}^+ + \lambda^- \sum_{i=1}^N \omega_{i,t}^- && \text{Feasibility} \\
0 &\leq \omega_{i,t}^+, \omega_{i,t}^-, l_{i,t}^+, l_{i,t}^-, l_t^m, i = 1, \dots, N && \text{Feasibility}
\end{aligned} \tag{257}$$

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. (255) the KKT conditions are both necessary and sufficient since the objective function is concave in (ω_t^+, ω_t^-) and all constraints are linear in (ω_t^+, ω_t^-) .

Solving the KKT conditions requires enumeration of all the possibilities for the complementary conditions. In general, the $2N + 1$ Lagrange multipliers $(l_t^m, l_{i,t}^+, l_{i,t}^-, i = 1, \dots, N)$ 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 $\omega_{i,t}^+ \omega_t^- > 0, i = 1, \dots, N$, 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:

$$E_t \left\{ g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} R_{i,t}^e \middle| I_t, \omega_t^+, \omega_t^- \right\} \approx \sum_{j=1}^{n_b} \alpha_{ij}(I_t) b_j(\omega_t), i = 1, \dots, N, \quad (258)$$

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 $\alpha_{ij}(I_t)$ 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 $\{\omega_t^{(k)}\}_{k=1}^{n_s}$ 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 $\{b_j(\omega_t^{(k)})\}_{j=1}^{n_b}$; 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. (258); 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. (258) 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. (258) 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 $\|\omega^{(i)} - \omega^{(i-1)}\| / \|\omega^{(i-1)}\|$, where $\|x\|$ is the norm of x , defined by $\|x\|^2 = \text{Trace}(x^\top x)$, 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. (255). 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. (254). The first order condition leads to

$$q_t^{-\gamma} = \frac{\partial f_t^p(I_t)}{\partial I_t}. \quad (259)$$

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. (256) and obtain

$$\frac{\partial f_t^p(I_t)}{\partial I_t} = \frac{\partial \mathcal{L}^p}{\partial I_t} \Big|_{\omega_t^*(I_t)} = \beta E_t \left[g_t^{-\gamma} \frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} \left(\sum_{i=1}^n \omega_t^*(I_t) R_{i,t}^e + R^f \right) \right], \quad (260)$$

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

$$\frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} = \frac{\partial \mathcal{L}(q_{t+1}, v_{t+1})}{\partial v_{t+1}} \Big|_{q_{t+1}^*(v_{t+1})} = \frac{\partial f_{t+1}^p(I_{t+1})}{\partial I_{t+1}} \Big|_{q_{t+1}^*(v_{t+1})} = q_{t+1}^{*-\gamma}(v_{t+1}). \quad (261)$$

Thus, due to the form of the Lagrangian, the value function sensitivity of the problem in Eq. (250) is completely specified by the optimal consumption as

$$\frac{\partial f_{t+1}(v_{t+1})}{\partial v_{t+1}} = \begin{cases} q_{t+1}^{*-\gamma}(v_{t+1}) & \text{if } t < T - 1 \\ \phi_\tau(v_T + 1)^{-\gamma} & \text{if } t = T - 1 \end{cases} \quad (262)$$

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

References

- Black, F., 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45, 444–454.
- Black, F., Litterman, R., 1992. Global portfolio optimization. *Financial Analysts Journal* 48, 28–43.
- 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.
- Calvet, L. E., Campbell, J. Y., Sodini, P., 2008. Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115, 707–747.
- Carroll, C. D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters* 91, 312–320.
- Cass, D., Stiglitz, J. E., 1970. The structure of investors preferences and asset returns and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *Journal of Economic Theory* 2, 122–160.
- Cox, J. C., Huang, C., 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory* 49, 33–83.
- Cuoco, D., 1997. Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. *Journal of Economic Theory* 72, 33–73.
- Cuoco, D., Liu, H., 2000. A martingale characterization of consumption choices and hedging costs with margin constraints. *Mathematical Finance* 10, 355–385.
- Cuoco, D., Liu, H., 2004. An analysis of var-based capital requirements. *Journal of Financial Intermediation* 15, 362–394.
- Cvitanić, J., Karatzas, I., 1992. Convex duality in constrained portfolio choices. *Annals of Applied Probability* 3, 767–818.
- Cvitanić, J., Karatzas, I., 1993. Hedging contingent claims with constrained portfolios. *Annals of Applied Probability* 3, 652–681.

- DeMarzo, P. M., Kaniel, R., Kremer, I., 2004. Diversification as a public good: Community effects in portfolio choice. *The Journal of Finance* 59, 1677–1715.
- 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.
- Dybvig, P. H., Liu, H., 2010. Lifetime consumption and investment: Retirement and constrained borrowing. *Journal of Economic Theory* 145, 885–907.
- El Karoui, N., Jeanblanc-Picqué, M., 1998. Optimization of consumption with labor income. *Finance and Stochastics* 2, 409–440.
- Fortune, P., 2000. Margin requirements, margin loans, and margin rates: Practice and principles. *New England Economic Review* 23, 19–44.
- French, K. R., 2008. mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
- Gallmeyer, M., Kaniel, R., Tompaidis, S., 2006. Tax management strategies with multiple risky assets. *Journal of Financial Economics* 80, 243–291.
- Garlappi, L., Skoulakis, G., 2010. Solving consumption and portfolio choice problems: The state variable decomposition method. *Review of Financial Studies* 23, 3346–3400.
- Geczy, C. C., Musto, D. K., Reed, A. V., 2002. Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics* 66, 241–269.
- Goetzmann, W. N., Kumar, A., 2008. Equity portfolio diversification. *Review of Finance* 12, 433–463.
- 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.
- He, H., Pearson, N. D., 1991. Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. *Journal of Economic Theory* 54, 259–304.
- 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.
- Karatzas, I., Lehoczky, J. P., Shreve, S. E., 1987. Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM Journal on Control and Optimization* 25, 1557–1586.
- Karatzas, I., Lehoczky, J. P., Shreve, S. E., Xu, G., 1991. Martingale and duality methods for duality maximization in an incomplete market. *SIAM Journal on Control and Optimization* 29, 702–730.
- Kelly, M., 1995. All their eggs in one basket: Portfolio diversification of us households. *Journal of Economic Behavior and Organization* 27, 87–96.
- Koo, H., 1998. Consumption and portfolio selection with labor income: a continuous time approach. *Mathematical Finance* 8, 49–65.
- Koo, H., 1999. Consumption and portfolio selection with labor income: A discrete-time approach. *Mathematical Methods of Operations Research* 50, 219–243.
- Kubler, F., Willen, P., 2006. Collateralized borrowing and life-cycle portfolio choice, FRB Boston Public Policy Discussion Papers Series, paper no. 06-4.
- Kumar, A., 2009. Who gambles in the stock market? *The Journal of Finance* 64, 1889–1933.
- Levin, J., 2006. Supermodular games. Stanford University. Available at <http://www.stanford.edu/%7Ejdlevin/Econ%20286/Supermodular%20Games.pdf>.
- Liu, H., 2010. Solvency constraint, underdiversification, and idiosyncratic risks, preprint.
- Merton, R., 1971. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3, 373–413.
- Mitton, T., Vorking, K., 2007. Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies* 20, 1255–1288.
- Polkovnichenko, V., 2005. Household portfolio diversification: A case for rank-dependent preferences. *Review of Financial Studies* 18, 1467–1501.

- Shiller, R. J., 2003. Long term stock, bond, interest rate and consumption data, available at www.econ.yale.edu/shiller/data.htm.
- Social Security Administration, 2004. Period life table available at www.ssa.gov/OACT/STATS/table4c6.html.
- Tepla, L., 2000. Optimal portfolio policies with borrowing and shortsale constraints. *Journal of Economic Dynamic and Control* 24, 1623–1639.
- Topkis, D., 1998. *Supermodularity and Complementarity*. Princeton University Press, Princeton, New Jersey.
- Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and underdiversification. *Review of Economic Studies* 77, 779–805.
- Viceira, L. M., 2001. Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance* 56, 433–470.
- Xu, G., Shreve, S. E., 1992. A duality method for optimal consumption and investment under short-selling prohibition. i. general market coefficients. *Annals of Applied Probability* 2, 87–112.
- Yang, C., 2010. *Functional Approximation Methods for Solving Stochastic Control Problems in Finance*. Ph.D. thesis, University of Texas at Austin.

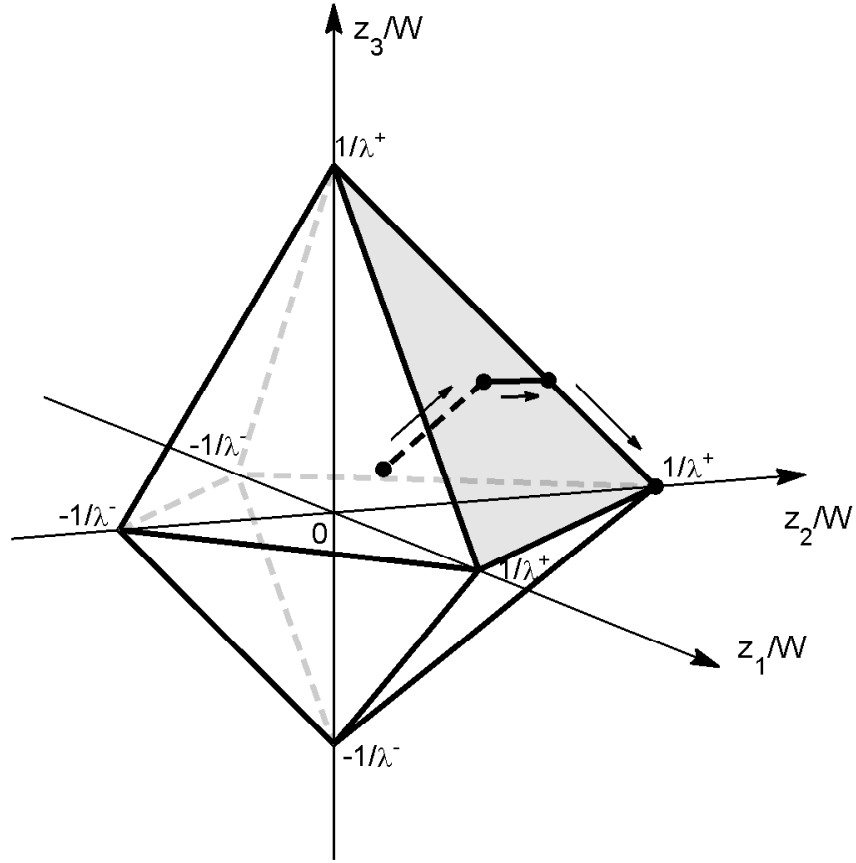


Fig. 1. This figure illustrates the intuition of Proposition 4 for the case of an asset allocation problem with a margin requirement, three risky assets with uncorrelated returns, and income growth uncorrelated with the risky assets returns. The axes correspond to the allocations in each risky asset as a percentage of wealth $z_i/W, i = 1, 2, 3$. The margin coefficients are λ^+ for long positions and λ^- for short positions. The allocations are shown at different values of the wealth to income ratio, with the arrows indicating the direction of change in the allocations as the ratio decreases. The margin requirement binds when the chosen allocation lies on the shaded plane; asset 1 is dropped when the allocation lies on the edge on the $z_2/W - z_3/W$ plane. Asset 3 is dropped when income is much larger than wealth, and the allocation is represented by the vertex on the positive z_2/W axis.

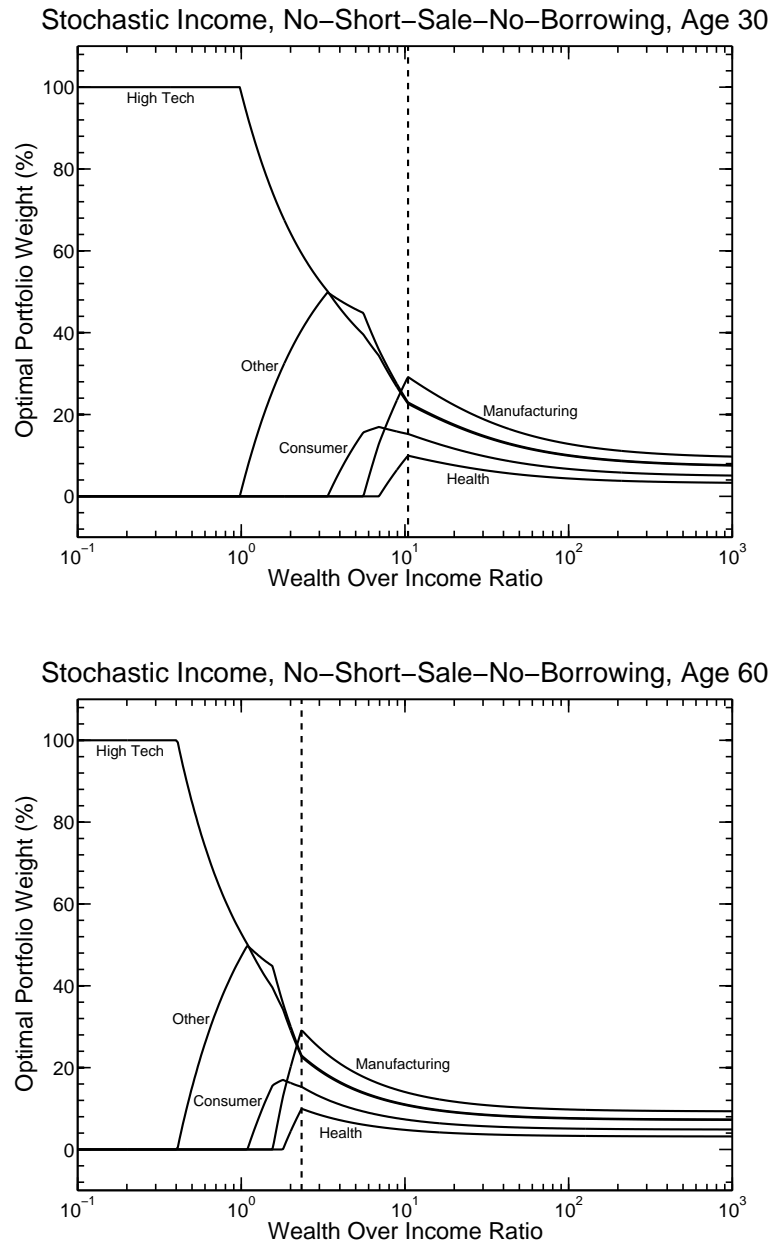


Fig. 2. This figure presents the asset allocations for different levels of the financial wealth to annual income ratio. The investor receives income with a stochastic growth rate. The investor's opportunity set consists of a riskless asset and five risky assets calibrated to the returns of stock industry indices for the industries High Tech, Consumer, Manufacturing, Health, and Other. The parameter values for the processes followed by the risky and riskless assets and the labor income growth are given in Table 1. The investor is not allowed to borrow or short an asset and is required to pay 100% of an asset's value. The top panel corresponds to a 30-year-old investor and the bottom panel to a 60-year-old investor.

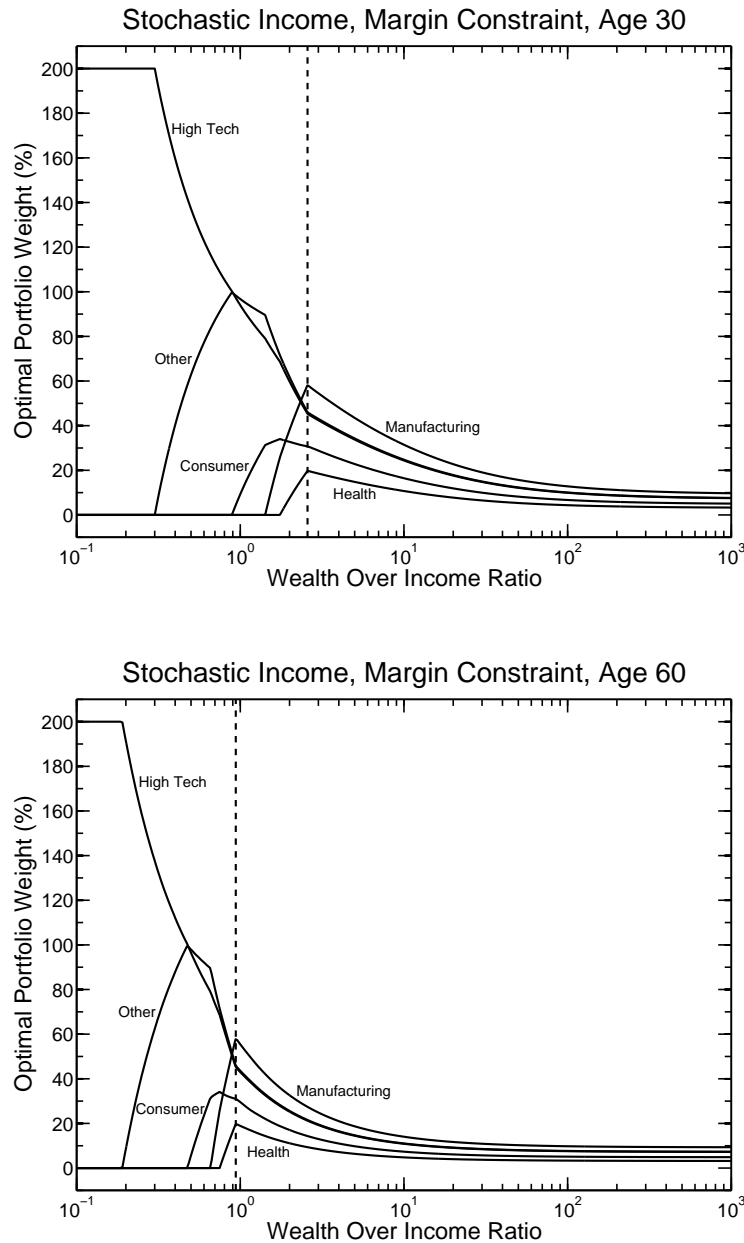
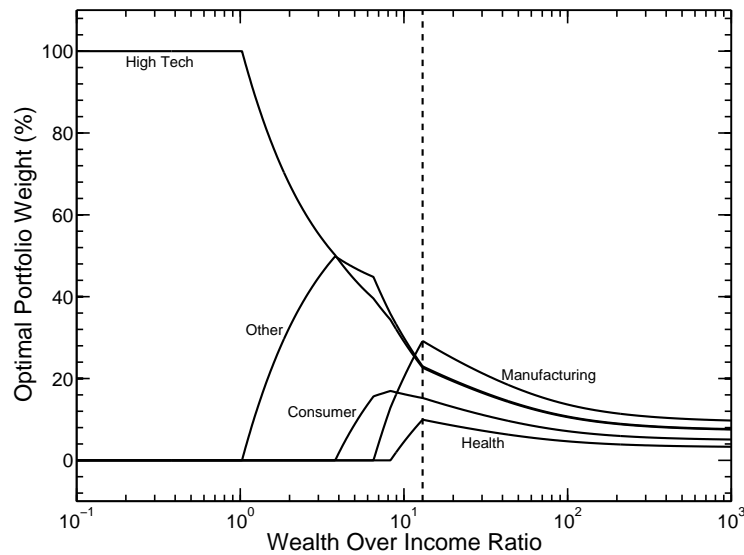


Fig. 3. This figure presents the asset allocations for different levels of the financial wealth to annual income ratio. The investor receives income with a stochastic growth rate. The investor's opportunity set consists of a riskless asset and five risky assets calibrated to the returns of stock industry indices for the industries High Tech, Consumer, Manufacturing, Health, and Other. The parameter values for the processes followed by the risky and riskless assets and the labor income growth are given in Table 1. The investor is allowed to purchase a risky asset on 50% margin and to short it at 150% margin. The top panel corresponds to a 30-year-old investor and the bottom panel to a 60-year-old investor.

Deterministic Income, No-Short-Sale-No-Borrowing, Age 30



Deterministic Income, No-Short-Sale-No-Borrowing, Age 60

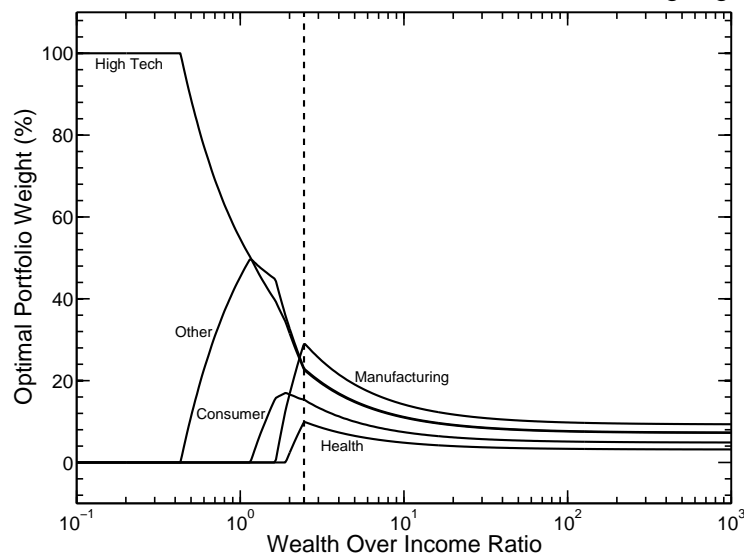


Fig. 4. This figure presents the asset allocations for different levels of the financial wealth to annual income ratio. The investor receives deterministic income. The investor's opportunity set consists of a riskless asset and five risky assets calibrated to the returns of stock industry indices for the industries High Tech, Consumer, Manufacturing, Health, and Other. The parameter values for the processes followed by the risky and riskless assets are given in Table 1. The investor is not allowed to borrow or short an asset and is required to pay 100% of an asset's value. The top panel corresponds to a 30-year-old investor and the bottom panel to a 60-year-old investor.

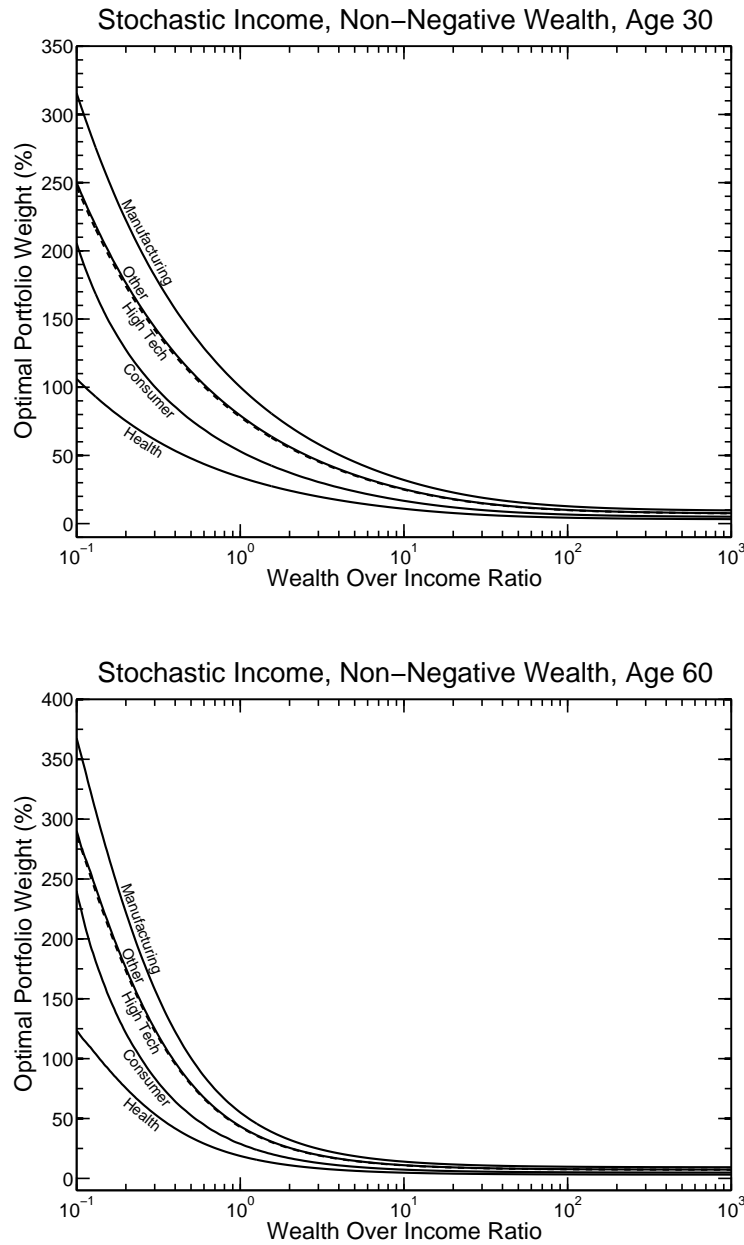


Fig. 5. This figure presents the asset allocations for different levels of the financial wealth to annual income ratio for an investor that faces a non-negative wealth constraint but no borrowing or margin requirements. The investor receives income with stochastic growth rate and has access to a riskless asset and five risky assets calibrated to the returns of stock industry indices for the industries High Tech, Consumer, Manufacturing, Health, and Other. The parameter values for the processes followed by income and the risky and riskless assets are given in Table 1. The top panel corresponds to a 30-year-old investor and the bottom panel to a 60-year-old investor.

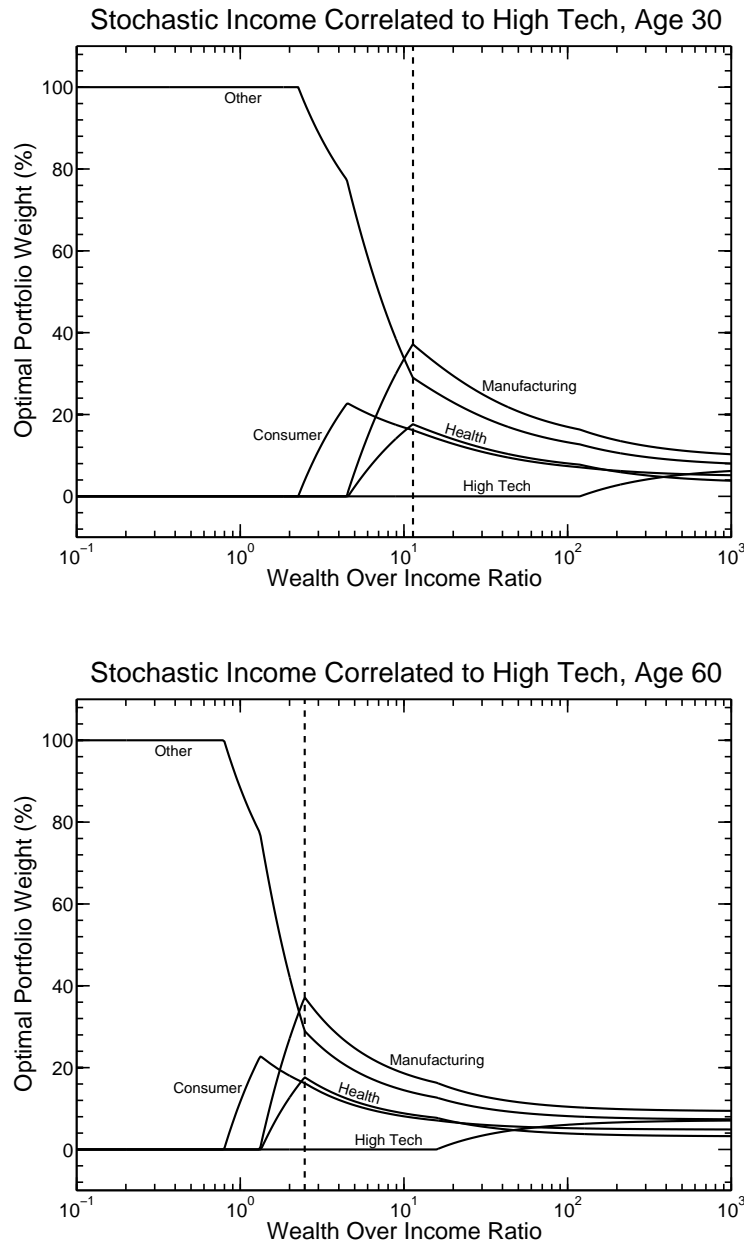


Fig. 6. This figure presents the asset allocations for different levels of the financial wealth to annual income ratio for an investor who receives labor income with stochastic growth that is correlated with the High Tech asset with correlation 20%. The investor's opportunity set consists of a riskless asset and five risky assets calibrated to the returns of stock industry indices for the industries High Tech, Consumer, Manufacturing, Health, and Other. The remaining parameter values for the processes followed by the risky and riskless assets and the labor income process are given in Table 1. The investor is not allowed to borrow or short an asset and is required to pay 100% of an asset's value. The top panel corresponds to a 30-year-old investor and the bottom panel to a 60-year-old investor.

Table 1: Parameter Values for the Base Case

This table includes the base case values of the parameters for the investor characteristics, and the characteristics of the returns of the risky assets. The five risky assets correspond to five industry indexes: Consumer, Manufacturing, High Tech, Health, and Other.

Number of periods	45 years (age 20 to age 65)					
Risk aversion	3					
Long margin	1					
Short margin	∞					
Time discount factor (annually)	0.98					
Interest rate (annually)	1.4%					
Income growth rate (annually)	3%					
Income volatility (annually)	10%					
	Cnsmr	Manuf	HiTec	Hlth	Other	Income
Drift (annually)	8.51%	7.83%	9.51%	6.97%	8.87%	3%
Volatility (annually)	28.9%	25.8%	33.3%	26.6%	29.7%	10%
Beta	1.01	0.91	1.15	0.79	1.06	
Sharpe ratio	25%	25%	24%	21%	25%	
Systematic risk	27%	25%	31%	21%	29%	
Idiosyncratic risk	9%	7%	12%	16%	7%	
Correlations	1.000	0.898	0.832	0.732	0.932	0.000
		1.000	0.848	0.698	0.930	0.000
			1.000	0.772	0.856	0.000
				1.000	0.727	0.000
					1.000	0.000
						1.000

Table 2: Asset Allocation and Diversification Measures of the Base Case

This table presents the optimal asset allocations and diversification measures for the base case: stochastic income growth with no-short-sale-no-borrowing constraint. W/Y is the current wealth to income ratio. Cnsmr, Manuf, HiTec, Hlth, and Other are the portfolio weights (as a percentage of current wealth) of the five industry indices: Consumer, Manufacturing, High Tech, Health, and Other. Margin is the total usage of the margin account in percentage. μ_h and σ_h are the expected value and standard deviation of the excess return of the risky part of the portfolio. $\sigma_{i,h}$ is the idiosyncratic standard deviation. IVarS is the idiosyncratic variance share. S_h is the Sharpe ratio of the risky part of the portfolio. RSRL $_h$ is the relative Sharpe ratio loss. RL $_h$ is the return loss of the total portfolio. UL $_h$ is the utility loss. β_h is the β of the constrained portfolio with respect to the unconstrained portfolio. LRRRA is the lifetime relative risk aversion.

Panel A: Age 30

W/Y	Cnsmr (%)	Manuf (%)	HiTec (%)	Hlth (%)	Other (%)	Margin (%)	μ_h (%)	σ_h (%)	$\sigma_{i,h}$ (%)	IVarS (%)	S_h (%)	RSRL $_h$ (%)	RL $_h$ (%)	UL $_h$ (%)	β_h	LRRRA
∞	5	9	7	3	7	32	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	3.00
1000.0	5	10	8	3	8	33	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.87
100.0	7	13	10	4	10	44	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.18
20.0	12	22	17	8	18	76	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.26
15.0	13	25	19	9	20	86	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.12
12.0	14	28	21	9	22	94	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.01
10.0	15	28	24	9	24	100	7.47	27.3	0.2	0.0	27.3	0.0	0.00	0.00	1.01	0.94
9.0	16	24	27	7	27	100	7.58	27.7	0.8	0.1	27.3	0.0	0.00	0.00	1.02	0.91
8.0	16	19	30	4	31	100	7.72	28.2	1.5	0.3	27.3	0.0	0.00	0.00	1.04	0.86
7.0	17	13	34	0	35	100	7.88	28.9	2.3	0.7	27.3	0.1	0.01	0.00	1.06	0.82
6.0	16	5	38	0	42	100	8.01	29.4	3.1	1.1	27.2	0.3	0.02	0.01	1.08	0.76
5.0	13	0	41	0	46	100	8.11	29.8	3.7	1.5	27.2	0.5	0.04	0.01	1.09	0.71
4.0	6	0	46	0	48	100	8.16	30.1	4.2	2.0	27.1	0.7	0.06	0.02	1.10	0.62
3.0	0	0	53	0	47	100	8.24	30.5	5.0	2.7	27.0	1.0	0.09	0.03	1.10	0.53
2.5	0	0	57	0	43	100	8.27	30.7	5.5	3.2	27.0	1.3	0.11	0.03	1.11	0.49
2.0	0	0	64	0	36	100	8.32	31.0	6.4	4.2	26.8	1.7	0.15	0.04	1.12	0.42
1.6	0	0	73	0	27	100	8.38	31.4	7.5	5.7	26.6	2.5	0.21	0.06	1.12	0.37
1.2	0	0	87	0	13	100	8.48	32.4	9.6	8.9	26.2	4.1	0.36	0.10	1.14	0.31
1.0	0	0	99	0	1	100	8.56	33.2	11.4	11.8	25.8	5.6	0.51	0.14	1.15	0.27
0.6	0	0	100	0	0	100	8.57	33.3	11.6	12.2	25.7	5.8	0.53	0.14	1.15	0.19
0.3	0	0	100	0	0	100	8.57	33.3	11.6	12.2	25.7	5.8	0.53	0.14	1.15	0.10
0.0	0	0	100	0	0	100	8.57	33.3	11.6	12.2	25.7	5.8	0.53	0.14	1.15	0.00

Panel B: Age 60

W/Y	Cnsmr (%)	Manuf (%)	HiTec (%)	Hlth (%)	Other (%)	Margin (%)	μ_h (%)	σ_h (%)	$\sigma_{i,h}$ (%)	IVarS (%)	S_h (%)	RSRL _h (%)	RL _h (%)	UL _h (%)	β_h	LRRA
∞	5	9	7	3	7	32	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	3.00
1000.0	5	9	7	3	7	32	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.98
100.0	5	10	8	3	8	33	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.85
20.0	6	12	9	4	9	40	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.38
15.0	7	12	10	4	10	43	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.23
12.0	7	13	10	5	10	45	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	2.10
10.0	7	14	11	5	11	48	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.99
9.0	8	15	11	5	11	50	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.94
8.0	8	15	12	5	12	52	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.83
7.0	8	16	12	5	13	55	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.74
6.0	9	17	13	6	13	59	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.63
5.0	10	19	14	6	15	64	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.50
4.0	11	21	16	7	17	72	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.34
3.0	13	25	19	8	20	85	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.11
2.5	15	28	22	10	22	96	7.43	27.2	0.0	0.0	27.3	0.0	0.00	0.00	1.00	1.01
2.0	16	20	29	4	30	100	7.69	28.1	1.4	0.2	27.3	0.0	0.00	0.00	1.03	0.88
1.6	16	3	38	0	43	100	8.04	29.5	3.3	1.2	27.2	0.4	0.03	0.01	1.08	0.75
1.2	5	0	47	0	48	100	8.18	30.1	4.3	2.1	27.1	0.7	0.06	0.02	1.10	0.60
1.0	0	0	53	0	47	100	8.24	30.5	5.0	2.7	27.0	1.0	0.09	0.03	1.10	0.54
0.6	0	0	74	0	26	100	8.39	31.5	7.7	5.9	26.6	2.6	0.22	0.06	1.12	0.34
0.3	0	0	100	0	0	100	8.57	33.3	11.6	12.2	25.7	5.8	0.53	0.14	1.15	0.19
0.0	0	0	100	0	0	100	8.57	33.3	11.6	12.2	25.7	5.8	0.53	0.14	1.15	0.00

Table 3: Base Case Simulation Results

This table presents summary statistics of the simulated wealth as well as the portfolio and consumption choices for an individual investor starting from a given initial wealth to annual income ratio and following the optimal investment and consumption strategy. The results are based on 10,000 simulation paths. W/Y and C/Y are the realized wealth to annual income ratio and consumption to income ratio. Cnsmr, Manuf, HiTec, Hlth, and Other are the portfolio weights, as a percentage of current wealth, of the five industry indices: Consumer, Manufacturing, High Tech, Health, and Other. Margin is the total usage of the margin account in percentage. Q_{25} , Q_{50} , and Q_{75} are the 25% percentile, the 50% percentile (median), and the 75% percentile. SD is the standard deviation.

Panel A: Initial wealth equal to two years of income									
		W/Y	C/Y	Cnsmr (%)	Manuf (%)	HiTec (%)	Hlth (%)	Other (%)	Margin (%)
Age 30	Q_{25}	2.1	0.9	0	0	54	0	10	100
	Q_{50}	2.7	1.0	0	0	69	0	30	100
	Q_{75}	3.7	1.0	0	0	90	0	43	100
	Mean	3.3	1.0	2	1	71	0	26	100
	SD	2.0	0.1	5	4	21	1	17	1
Age 45	Q_{25}	3.4	0.8	0	0	26	0	23	100
	Q_{50}	4.9	0.9	9	0	43	0	37	100
	Q_{75}	8.0	1.1	15	19	55	6	46	100
	Mean	6.8	1.0	8	8	43	2	34	96
	SD	5.7	0.2	7	11	19	4	12	10
Age 60	Q_{25}	7.8	0.6	7	13	10	4	10	43
	Q_{50}	11.0	0.8	7	14	11	5	11	48
	Q_{75}	15.7	1.1	8	16	13	5	13	55
	Mean	13.0	0.9	8	15	11	5	12	50
	SD	8.2	0.4	2	3	3	1	3	11
Panel B: Initial wealth equal to ten years of income									
		W/Y	C/Y	Cnsmr (%)	Manuf (%)	HiTec (%)	Hlth (%)	Other (%)	Margin (%)
Age 30	Q_{25}	5.0	1.1	7	0	20	0	20	88
	Q_{50}	8.7	1.3	12	17	31	3	28	100
	Q_{75}	15.3	1.6	15	24	46	8	44	100
	Mean	11.5	1.4	10	13	35	4	30	93
	SD	9.0	0.4	6	11	18	4	12	12
Age 45	Q_{25}	5.0	0.9	7	0	15	0	15	67
	Q_{50}	9.6	1.1	10	16	22	5	21	96
	Q_{75}	18.5	1.5	14	22	42	7	40	100
	Mean	14.2	1.3	10	14	30	4	27	84
	SD	13.7	0.6	5	10	18	4	13	19
Age 60	Q_{25}	9.5	0.7	6	12	9	4	9	40
	Q_{50}	14.0	1.0	7	13	10	4	10	44
	Q_{75}	22.0	1.4	8	15	11	5	12	51
	Mean	18.1	1.2	7	14	11	5	11	47
	SD	14.8	0.8	2	3	2	1	2	10