

Solvency Constraint, Underdiversification, and Idiosyncratic Risks *

Hong Liu

Washington University in St. Louis and CAFR

This version: September 4, 2010

*I thank Nick Barberis, Jonathan Berk, John Campbell, Jon Ingersoll, Ravi Jagannathan, Mark Loewenstein, Todd Milbourn, Stijn Van Nieuwerburgh, Valery Polkovnichenko, Paolo Sodini, Matthew Spiegel, Anjan Thakor, Harald Uhlig, Raman Uppal, Dimitri Vayanos, Laura Veldkamp, Yajun Wang, Guofu Zhou, and seminar participants at Ajou University, KAIST, Rutgers University, Singapore Management University, UC San Diego, Washington University in St. Louis, Yale University, the 2007 WFA, the 2008 ASAP, and the 2009 AFA conferences for very helpful suggestions. I am indebted to Kerry Back, Phil Dybvig, and Pietro Veronesi for their constructive comments that have significantly improved the paper. Send correspondence to Hong Liu at Olin Business School, Washington University, Campus Box 1133, One Brookings Drive, St. Louis, MO 63130. Email: liuh@wustl.edu.

Solvency Constraint, Underdiversification, and Idiosyncratic Risks

Abstract

Contrary to the prediction of standard portfolio diversification theory, many investors place a large fraction of their stock investment in a small number of stocks. We show that underdiversification may be a result of the solvency requirement. The key assumption is that investors must remain solvent after certain committed consumption at which the marginal utility is finite. We show that investors underdiversify when wealth is low and that even wealthy investors may underdiversify. In addition, variance and higher moments do not affect stock selection and a less diversified stock portfolio has a higher expected return, a higher volatility, a higher skewness, and maybe a lower Sharpe ratio. Underdiversification in equilibrium implies that idiosyncratic risks are priced. All of these theoretical predictions are consistent with empirical findings.

Journal of Economic Literature Classification Numbers: D11, D91, G11, C61.

Keywords: Diversification, Underdiversification, Solvency Constraint, Idiosyncratic Risks.

In contrast to the theory of the celebrated Capital Market Pricing Model (CAPM), extensive empirical literature has shown that many investors underdiversify and idiosyncratic risks are priced. For example, among the households that hold individual stocks directly, the median number of directly held stocks was two until 2001, when it increased to three.¹ The main empirical findings on underdiversification include (1) the number of stocks directly held by poor investors is small (e.g., Campbell 2006) and increases as investors' wealth increases (e.g., Goetzmann and Kumar 2005); (2) even wealthy investors may hold a small number of stocks in the directly held portfolio (e.g., Polkovnichenko 2005); (3) many households simultaneously invest in well-diversified funds and in extremely underdiversified stock portfolios (e.g., Polkovnichenko 2005); (4) less diversified investors hold stocks with higher expected returns (e.g., Goetzmann and Kumar 2005); and (5) underdiversified portfolios have higher expected returns (e.g., Calvet, Campbell, and Sodini 2007, Table 5, Kroll, Levy, and Rapoport, 1988a, 1988b), higher volatility (e.g., Goetzmann and Kumar 2005), lower Sharpe ratios (e.g., Goetzmann and Kumar 2005), and higher skewness (e.g., Mitton and Vorkink 2007). There is also a large literature on whether or not idiosyncratic risks are priced. For example, Bessembinder (1992) finds strong evidence that idiosyncratic risk was priced in the foreign currency and agricultural futures markets. Ang, Hodrick, Xing, and Zhang (2003) provide empirical evidence suggesting that idiosyncratic volatility affects stock return. Obviously, the relevance of idiosyncratic risks for asset pricing can be a result of underdiversification.

Some possible explanations for underdiversification such as trading costs, differ-

¹See, for example, Kelly (1995), Goetzmann and Kumar (2005), Polkovnichenko (2005), and Campbell (2006). Both the fraction of stockholders who directly hold stocks and the value of their directly held portfolios are significant. For example, using the Survey of Consumer Finance 1983-2001 data, Polkovnichenko (2005) shows that about 30% of all stockholders directly hold individual stocks and the directly held portfolio is worth more than 40% of their financial assets that include both directly and indirectly held stock portfolios (e.g., through retirement account, mutual funds).

ential ambiguity aversion, psychological and behavioral factors, and “special” preferences, have been proposed.² However, it is unlikely that these existing models can explain many of the above main findings. For example, if trading cost is the main concern, it is unclear why investors do not diversify through index funds or S&P 500 SPDR ETF, which have small trading costs. In most of the models based on ambiguity aversion, psychological and behavioral factors, and “special” preferences, the widely documented wealth effect on underdiversification is largely absent. Also, none of these models predict that underdiversified portfolios have higher expected returns and lower Sharpe ratios.

In this article, we propose a new and simple explanation: underdiversification, and thus the importance of idiosyncratic risks for asset pricing, can be a result of the solvency requirement in the presence of committed consumption.³ Specifically, we consider an investor who can trade one risk-free asset and multiple risky stocks to maximize his expected utility from the terminal wealth in a one-period setting. Our main results hold for quite general preferences and stock return distributions. The main assumption is that the investor commits to a nonnegative terminal consumption at which the marginal utility is finite and must remain solvent after the committed consumption.⁴

Different from the existing literature, our model can help explain *all* of the five main empirical findings listed above with unique predictions (4) and (5). In addition, our model can also help explain that (6) young or male investors underdiversify more

²Brennan (1975), Kraus and Litzenberger (1976), Huberman (2001), Merton (1987), Nieuwerburgh and Veldkamp (2005), Barberis and Huang (2005), Uppal and Wang (2003), Mitton and Vorkink (2007), and Boyle, Garlappi, Uppal and Wang (2009) .

³Consumption commitment is well documented in the existing literature (e.g., Chetty and Szeidl 2007, Postlewaite, Silverman, and Samuelson 2008) and can be motivated by habit formation, meeting fixed financial obligations such as mortgage and tuition payments, or precautionary savings against unemployment or health shocks, etc.

⁴For utility functions with finite marginal utility at zero wealth (e.g., CARA preferences), consumption commitment is not required and solvency requirement alone is sufficient for our results.

(e.g., Mitton and Vorkink 2007) and (7) idiosyncratic risks are priced (e.g., Bessembinder 1992). Also different from the existing models, our model implies that only expected return and covariance with other stocks affect stock selection. Other moments such as variance and skewness (and thus Sharpe ratio) are irrelevant for this choice.⁵ In addition, in contrast to models based on psychological and behavioral factors and “special” preferences that use distorted probabilities, our model is fully rational without assuming any distortion of probabilities.

As an illustration of the economic significance of our model, we show that with reasonable parameter values, an investor with \$20,000 and a relative risk aversion of 2 invests only in four stocks, which is consistent with empirical evidence. For example, Mitton and Vorkink (2007) find that in their database with the portfolios of 78,000 households the median portfolio value with only four stocks is \$21,903. We also find that if an investor holds a well-diversified portfolio, then the directly held stock portfolio can be even more underdiversified. For example, suppose an investor holds a \$300,000 well-diversified portfolio in a retirement account and he requires at least \$10,000 outside the retirement account for committed consumption. We show that if the investor’s wealth outside the retirement account is lower than \$61,500, then he holds at most two stocks directly. In fact, for this investor to hold four stocks, he needs to have at least \$299,000 outside the retirement account.

To explain the essential intuitions for our main results in the simplest case, sup-

⁵It is important to note that, while other moments such as variance are irrelevant for stock selection, they do affect how much is invested in a stock *once the stock is selected*. The amount invested in a highly risky asset (e.g., an option) can be very small. Indeed, for any given asset (including an option), an investor in our model never holds more than what is predicted by the standard theory. Thus, small learning costs, such as those considered by Nieuwerburgh and Veldkamp (2005), would prevent him from holding it at all. This may reconcile our model’s predictions with the fact that most investors do not hold options, despite the high expected returns, because options are highly risky and the associated learning costs may be nontrivial. In addition, low covariance with existing assets such as durable goods and retirement portfolios may be another reason for the low holdings of options.

pose an investor has a mean-variance preference and assume all stocks have different expected returns. Because of the solvency constraint and discrete-time trading, only limited borrowing or shortselling is feasible. This implies that when the investor's wealth is low, he can invest only a small amount in stocks, which implies that expected return has a first-order effect on utility, while risk has only a second-order effect, by local risk neutrality. Therefore when his wealth is low, the marginal benefit of diversification (i.e., reducing risk) is smaller than the marginal cost of diversification (i.e., lowering expected return), and thus he only invests in the stock with highest expected return and does not diversify. As wealth increases, he invests more in the stock, his portfolio risk increases, and thus the marginal benefit of diversification increases. At a critical wealth level, the marginal benefit of diversification surpasses the marginal cost of diversification and thus the investor adds the second stock to his portfolio. Among all the return moments of a stock, only expected return affects the marginal cost of diversification and only the covariance with the first stock affects the marginal benefit of diversification. Therefore the selection of a stock as the second stock to be added to the portfolio only depends on the stock's expected return and its covariance with the first stock, but not on other moments such as variance and skewness. This process continues with further increases in wealth. Given high enough wealth, the investor may invest in all stocks and thus fully diversify. However, for some preferences (e.g., CRRA or mean-variance), because of limited borrowing and shortselling, the marginal benefit of full diversification may be always lower than the marginal cost of full diversification. In these cases, even wealthy investors underdiversify.

In equilibrium, while rich investors fully diversify and hold all the stocks, poor investors only hold the stocks with the highest expected returns. Therefore, no one holds the market portfolio in equilibrium and idiosyncratic risks are priced. As poor investors' wealth increases, they sequentially add stocks with lower expected returns.

Thus, a more diversified stock portfolio has a lower expected return. Due to the diversification effect, the return on a more diversified stock portfolio has lower volatility and may have a higher Sharpe ratio. In addition, because adding lower expected return stocks shifts the portfolio return distribution to the left, a more diversified portfolio also has lower skewness. Finally, more risk-averse investors diversify more because they are less risk tolerant. Combined with the empirical finding that younger investors and male investors are less risk averse, our model then predicts that younger or male investors underdiversify more.

The key assumption driving our main results is that the investor commits to a non-negative minimum consumption level at which the marginal utility is finite and must remain solvent after the committed consumption. Without the solvency requirement, a poor investor would always borrow or shortsell to hold a fully diversified portfolio as a rich investor. If the marginal utility at the committed consumption level were infinite, then the marginal benefit of diversification would be high no matter how low an investor's wealth is, and therefore he would always fully diversify.

Our paper is related to the large literature on how habit-formation preferences affect portfolio selection (e.g., Constantinides 1990). The key difference from most of the literature is that, in our model, the marginal utility at the committed consumption level is finite and investors face limited borrowing and shortselling constraints implied by the solvency requirement.⁶ As discussed above, both of these features are important for our main results.

Our paper is also related to the literature on portfolio selection with portfolio constraints (e.g., Ross 1977, Dybvig 1984, and Cuoco and Liu 2006). Most previous works are done in different contexts and for different purposes. In addition, they

⁶The model of Dybvig (1995) also implies that marginal utility at the required living standard is finite. In contrast to our model, both Constantinides (1990) and Dybvig (1995) assume a financial market with a single risky asset.

assume either special preferences or special asset return distributions and offer only partial equilibrium analyses. For example, Ross (1977) and Dybvig (1984) examine the shape of a mean-variance efficient frontier with short-sale constraints. While short-sale constraints can prevent investors from shorting, they do not prohibit investors from borrowing to buy all the stocks with positive expected returns. Cuoco and Liu (2006) consider the impact of the Basel II capital requirements on the riskiness of a financial institution. Basel II capital requirements do not apply to individual investors and are much more stringent than the solvency constraint we assume. In addition, in contrast to our model, they restrict their analysis to CRRA preferences, lognormal stock prices, and portfolio allocation problems, without examining any equilibrium impact of the capital requirements.

The rest of the article is organized as follows. In Section 1, we use some simple examples to illustrate the essential intuitions for our main results. In Section 2, we explicitly solve an equilibrium model to show the main results. In Section 3, we extend our main results to a more general setting with general preferences, general stock return distributions, and the presence of a nontraded asset. Section 4 concludes. We prove the main results in the Appendix.

1. Some simple examples

In this section, we provide some simple examples to explain our main results that an investor underdiversifies when poor and invests in more stocks as his wealth increases. In a one-period setting, consider the simplest case where there are two independent stocks and a risk-free asset with interest rate r normalized to zero. Stock gross returns are unbounded above and can get arbitrarily close to 0. An investor has a mean-variance preference with a risk-aversion coefficient of A and a committed consumption

$\underline{C} > 0$. Given discrete-time trading and the full support of stock gross returns, the investor cannot borrow or shortsell; otherwise, the committed consumption (and solvency) cannot be guaranteed. For $i = 1, 2$, let μ_i and σ_i be respectively the expected return and the return volatility of Stock i with $\mu_1 > \mu_2 > r = 0$. We now explain how the optimal portfolio composition changes as the initial wealth W_0 increases from \underline{C} . When $W_0 = \underline{C}$, the investor can only invest in the risk-free asset to guarantee the committed consumption. Suppose now wealth is slightly above the minimum level, i.e., $W_0 = \underline{C} + \eta$ for some small $\eta > 0$. Since the investor cannot borrow or shortsell and must invest at least \underline{C} in the risk-free asset, the investor can invest at most the (small) fraction $\alpha \equiv \eta/W_0$ of his wealth in stocks. The investor's problem is then

$$\max_{\{\alpha_1, \alpha_2\}} \alpha_1 \mu_1 + \alpha_2 \mu_2 - \frac{1}{2} A \alpha_1^2 \sigma_1^2 - \frac{1}{2} A \alpha_2^2 \sigma_2^2,$$

subject to the no-borrowing and no-shortselling constraint

$$\alpha_1 + \alpha_2 \leq \alpha, \alpha_1 \geq 0, \alpha_2 \geq 0, \tag{1}$$

where α_i denotes the fraction of wealth W_0 (not η) invested in Stock i for $i = 1, 2$. The no-borrowing constraint (the first inequality in (1)) is binding for small enough α because stocks have higher expected returns than the risk-free asset. Thus the investor's problem becomes

$$\max_{0 \leq \alpha_1 \leq \alpha} V(\alpha_1) \equiv \alpha_1 \mu_1 + (\alpha - \alpha_1) \mu_2 - \frac{1}{2} A \alpha_1^2 \sigma_1^2 - \frac{1}{2} A (\alpha - \alpha_1)^2 \sigma_2^2,$$

which implies that

$$V'(\alpha_1) = (\mu_1 - \mu_2) - A (\alpha_1 \sigma_1^2 - (\alpha - \alpha_1) \sigma_2^2). \tag{2}$$

When α approaches zero, so does α_1 . Therefore, if α is small enough, $V'(\alpha_1)$ is always strictly positive because $\mu_1 > \mu_2$, which implies that $\alpha_1 = \alpha$, i.e., we have a corner solution. Thus, the investor invests all his disposable wealth η in the stock with the highest expected return (Stock 1) and does not diversify.

We can also show this result from comparing the marginal utilities from the two stocks. The marginal utility of investing $\alpha \geq 0$ in Stock i is equal to $\mu_i - A\alpha\sigma_i^2$, which approaches μ_i as α approaches 0. Therefore when α is small, the stock with the highest expected return provides the greatest marginal utility, which makes the investor only invest in this stock. $V'(\alpha_1)$ in (2) is exactly the difference in the marginal utilities from the two stocks (with $\alpha_2 = \alpha - \alpha_1$), which approaches $\mu_1 - \mu_2$ as the investment amount α approaches 0.

Intuitively, a risk-neutral investor only invests in the stock with the highest expected return. A risk-averse investor may hold stocks with lower returns to reduce risk. The right-hand side of equation (2) is also the difference between the marginal cost and the marginal benefit of diversification. More specifically, the marginal cost of diversification is the reduction in the expected return ($\mu_1 - \mu_2$), while the marginal benefit is the reduction in the risk (the second term in (2)). As α approaches 0, so does α_1 , which implies that the marginal benefit of diversification goes to 0 too. Therefore, the marginal benefit of diversification is smaller than the marginal cost of diversification when wealth is low and poor investors do not diversify.

As the investor's wealth further increases, he invests more in Stock 1, his portfolio risk increases, and thus the marginal benefit of diversification increases. When this marginal benefit surpasses the marginal cost of diversification, he adds the second stock. Suppose the critical wealth level beyond which the investor adds Stock 2 is $W_0 = \hat{W}$ and let $\hat{\alpha} \equiv (1 - 1/\hat{W})$ be the fraction of wealth invested in Stock 1. Then

we must have

$$\mu_1 - \mu_2 = A\hat{\alpha}\sigma_1^2, \quad (3)$$

where the right-hand side follows from setting $\alpha_1 = \alpha = \hat{\alpha}$ in the second term of (2), and thus

$$\hat{\alpha} = \frac{\mu_1 - \mu_2}{A\sigma_1^2}, \quad (4)$$

which implies that

$$\hat{W} = \frac{\underline{C}}{1 - \frac{\mu_1 - \mu_2}{A\sigma_1^2}}. \quad (5)$$

Note that among all the moments of the second stock, only the expected return affects the marginal cost of diversification. Therefore, if there were other uncorrelated stocks, the second stock the investor adds would be the stock with the second highest expected return. If a stock is correlated with the first stock, then its covariance with the first stock affects the diversification effectiveness and thus the marginal benefit of diversification. In this case, both expected return and covariance affect stock selection. However, other moments such as variance do not affect this choice.

As wealth increases further, the investor invests more in both stocks. If it is optimal to have no leverage in the unconstrained case, i.e.,

$$w_1^* + w_2^* < 1$$

where

$$w_1^* = \frac{\mu_1 - r}{A\sigma_1^2}, w_2^* = \frac{\mu_2 - r}{A\sigma_2^2},$$

then when the wealth is high enough, the investor holds the tangency portfolio. However, if

$$\frac{\mu_1 - \mu_2}{A\sigma_1^2} > 1, \quad (6)$$

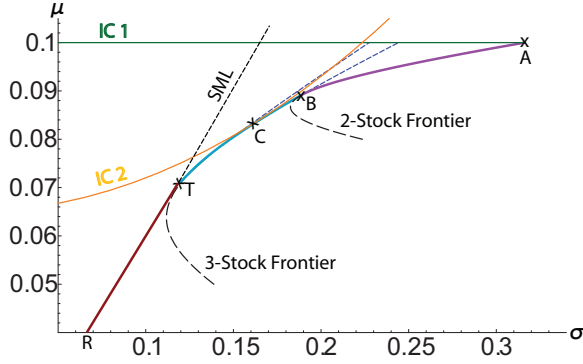


Figure 1: Mean-Variance Efficient Frontier.

then because of the no-borrowing constraint ($\alpha < 1$), the marginal benefit of diversification is always lower than the marginal cost, no matter how high W_0 is. Therefore, in this case, the investor always underdiversifies regardless of his wealth level.

Example 1: Suppose $r = 0$, $A = 2$, $\mu_1 = 0.15$, $\mu_2 = 0.05$, $\sigma_1 = 0.3$, $\sigma_2 = 0.05$, $\underline{C} = \$10,000$. If the investor's wealth W_0 is equal to $\$10,000$, then all $\$10,000$ must be invested in the risk-free asset. If W_0 is above $\$10,000$, then he starts to invest only in Stock 1 because it has the highest expected return, even though Stock 2 has a much greater Sharpe ratio. He will invest only in Stock 1 until his wealth gets above $\hat{W} = \$22,500$ by (5), i.e., he will invest up to 56% of his wealth in one stock that has a higher risk and lower Sharpe ratio. In addition, if $A = 1$, then by (6) the investor will never invest in Stock 2 no matter how wealthy he is.

Figure 1 plots the mean-variance frontier with three uncorrelated stocks. Stocks 1–3 are sorted by expected returns from the highest to the lowest. Because of the no-borrowing constraint, the security market line above the tangency point T is no longer relevant. Because of the short-sales constraints, the dotted segments are no longer achievable so the relevant frontier becomes the curve ABCR. The investor's

utility function is $u = \mu\alpha - \frac{1}{2}\alpha^2\sigma^2$ and thus the indifference curve at utility level u is

$$\mu = u/\alpha + \frac{1}{2}A\alpha\sigma^2. \quad (7)$$

IC 1 in Figure 1 is the indifference curve when the investor's wealth is just slightly above the subsistence level (i.e., α is small), which is almost horizontal by (7), because the slope $\partial\mu/\partial\sigma$ is almost 0. To maximize utility, the investor chooses point A, which represents investment only in Stock 1. As α increases, the indifference curve becomes more curved. When it becomes tangent to ABCTR at point A, the investor starts to add Stock 2. As α continues to increase, the tangency point moves from A to B, at which point the investor adds Stock 3. As α increases further, the investor invests more in each of the three stocks, as shown by point C implied by IC 2. After the tangency portfolio T is reached, the investor also increases the investment in the risk-free asset. If the tangency point between the indifference curve and the frontier ABCTR never moves beyond point C for any wealth level, then even wealthy investors will not hold Stock 3 and thus underdiversify.

Next, we consider the same setup as above except that the investor has a CRRA utility with a relative risk-aversion coefficient of γ . Suppose in addition the gross return \tilde{P}_i of Stock i is lognormally distributed with mean $\mu_i + 1$ and volatility σ_i . The investor's problem is then

$$\max_{\{\theta_1, \theta_2\}} E \left[u(\tilde{W}_1) \right] \quad (8)$$

subject to

$$\tilde{W}_1 = (W_0 - \theta_1 - \theta_2) + \theta_1\tilde{P}_1 + \theta_2\tilde{P}_2, \quad (9)$$

and

$$\tilde{W}_1 \geq \underline{C}, \quad (10)$$

where \tilde{W}_1 is the end-of-period wealth, θ_i denotes the dollar amount invested in Stock i and

$$u(W) = \frac{W^{1-\gamma}}{1-\gamma}.$$

Given the lognormally distributed gross returns, (10) is equivalent to the following no-borrowing and no-shortselling constraint:

$$\theta_1 + \theta_2 \leq W_0 - \underline{C}, \theta_1 \geq 0, \theta_2 \geq 0. \quad (11)$$

First consider the case with $\underline{C} = 0$, i.e., no committed consumption. In this case, the marginal utility at \underline{C} ($u'(0)$) is infinity. The optimal fractions of wealth invested in the stocks ($w_1^* \equiv \theta_1^*/W_0$ and $w_2^* \equiv \theta_2^*/W_0$) are independent of the initial wealth level $W_0 > 0$, which is a well-known result for CRRA preferences.

Now suppose $\underline{C} > 0$. Then at $W_0 = \underline{C}$, constraint (11) implies that $\theta_1 = 0$ and $\theta_2 = 0$, i.e., the investor must invest all the wealth in the risk-free asset. Now suppose the initial wealth is slightly above \underline{C} , i.e., $W_0 = \underline{C} + \eta$ for some small $\eta > 0$. In this case, the investor can invest η into the stocks. In contrast to the case with $\underline{C} = 0$, he will not invest the same w_i^* fraction of η in Stock i . To show this, note that if the investor invests η in Stock i , then

$$\tilde{W}_1 = \underline{C} + \eta \tilde{P}_i. \quad (12)$$

Thus

$$\lim_{\eta \downarrow 0} \frac{\partial E[u(\tilde{W}_1)]}{\partial \eta} = \lim_{\eta \downarrow 0} E \left[u'(\tilde{W}_1) \tilde{P}_i \right] = u'(\underline{C})(\mu_i + 1), i = 1, 2. \quad (13)$$

Since the marginal utility at \underline{C} is finite (i.e., $u'(\underline{C}) < \infty$), this shows that the stock with the highest expected return (Stock 1) provides the highest marginal utility when

the investor's wealth is just above the committed consumption level. Therefore, if the investor is poor enough, he only invests in Stock 1. Note that if $u'(\underline{C}) = \infty$, then all stocks have the same (infinite) marginal utility and thus the investor should fully diversify no matter how poor he is, as the standard theory predicts. This result shows the critical importance of the finiteness of the marginal utility at the committed consumption \underline{C} . It also shows the important difference of our model from the standard habit formation models where the utility function is in the form of $u(W - \underline{C})$, which implies that the marginal utility at \underline{C} is infinite for CRRA preferences and thus investors fully diversify regardless of their wealth level.

As before, we can also explain the intuition for this result by comparing the marginal cost and marginal benefit of diversification. Since the marginal utility at \underline{C} is finite, the marginal benefit of diversification approaches 0 when the amount of investment approaches 0 by local risk neutrality. By (13), the marginal cost of diversification when the amount of investment approaches 0 approaches $u'(\underline{C})(\mu_1 - \mu_2) > 0$. Therefore, if the investor is poor enough, he only invests in the stock with the highest expected return and does not diversify. If $u'(\underline{C}) = \infty$, however, then the marginal benefit of diversification is infinity even when the amount of investment approaches 0. Thus, the investor should fully diversify no matter how poor he is.

By the same arguments as in the mean-variance case, as wealth increases further, the investor invests more in Stock 1 and may fully diversify when wealth is high enough.

Example 2: Suppose the parameter values are the same as in Example 1 except that the investor has a CRRA preference with $\gamma = 3$. If W_0 is slightly above \$10,000, then he starts to invest only in Stock 1 because it has a higher expected return, even though Stock 2 has a much greater Sharpe ratio. Applying (66) in the Appendix with $m = 1$, we find that the critical wealth level beyond which it is optimal to add Stock

2 is \$18,100. Therefore, as long as his wealth is below \$18,100 he will invest only in Stock 1. In addition, if $\gamma = 1$, then (66) implies that the investor will never invest in Stock 2 no matter how wealthy he gets, because he cannot borrow and the marginal cost of diversification at the investment of 100% of his wealth in Stock 1 is still higher than the marginal benefit of diversification.

2. A simple equilibrium model for underdiversification

In this section, we show that underdiversification can arise in equilibrium and our model can help explain many of the main empirical findings. Specifically, we consider a one-period model where investors maximize the expected utility from the final wealth on date 1. We assume that investors have the same constant absolute risk aversion (CARA) preferences, i.e.,

$$u(W) = -e^{-AW},$$

where $A > 0$ is the CARA coefficient. There is one storable consumption good: the only risk-free asset in the economy. We choose the consumption good as the numeraire and thus the interest rate is normalized to 0. There are $n > 0$ risky stocks with a positive total supply of $\bar{\omega}$ ($n \times 1$) shares. The per-share payoffs \tilde{P} ($n \times 1$) are independently Gamma distributed with $n \times 1$ parameter vectors $\alpha > 1$ and $\beta > 0$.⁷

⁷CARA preferences and Gamma distributions are used only for tractability. Allowing different risk aversions is an straightforward extension. The main qualitative results in this section remain valid with other preferences and payoff distributions, although closed-form solutions would become unlikely. Gamma distributions have similar properties to those of lognormal distributions, including the support set and moment characteristics.

The probability density function for Stock j payoff \tilde{P}_j is then

$$f_j(x) = \frac{x^{\alpha_j-1} e^{-x/\beta_j}}{\beta_j^{\alpha_j} \Gamma(\alpha_j)},$$

with mean $\kappa_j = \alpha_j \beta_j$ and variance $\varphi_j^2 = \alpha_j \beta_j^2$, for $1 \leq j \leq n$.

There are two groups of investors: the rich, with mass 1, and the poor, with mass $\lambda \geq 0$. Both types of investors are subject to the solvency constraint $\widetilde{W}_1 \geq 0$.⁸ The rich are endowed with $W_r \geq 0$ units of the consumption good and $\bar{\omega}$ shares of the stocks. The poor are only endowed with $W_p \geq 0$ units of the consumption good, but no stocks. Since stock payoffs are unbounded above and can get arbitrarily close to 0, to ensure solvency, no one in the economy can borrow or shortsell. Let p denote the $n \times 1$ date 0 equilibrium stock price vector. The rich solve

$$\max_{\omega} E \left[-e^{-A \widetilde{W}_1} \right], \quad (14)$$

s.t.,

$$\widetilde{W}_1 = W_r + (\bar{\omega}^\top - \omega^\top) p + \omega^\top \tilde{P} \geq 0, \quad (15)$$

where the $n \times 1$ vector ω denotes the number of shares held in stocks until date 1.

First, suppose there are no poor investors in the economy, i.e., $\lambda = 0$. In this case, the solvency constraint is not binding for the rich, because $W_r \geq 0$ and the rich's stock endowment is positive. Then Problem (14) becomes

$$\max_{\omega} E \left[-e^{-A(W_r + \bar{\omega}^\top p + \omega^\top (\tilde{P} - p))} \right] = -e^{-A(W_r + \bar{\omega}^\top p)} \prod_{j=1}^n \min_{\omega_j} \frac{e^{A p_j \omega_j}}{(1 + A \omega_j \beta_j)^{\alpha_j}}. \quad (16)$$

⁸Committed consumption is not required because the marginal utility at zero wealth is finite for CARA preferences.

The first-order conditions then imply that

$$\omega_j = \left(\frac{\kappa_j}{p_j} - 1 \right) \frac{\kappa_j}{A\varphi_j^2}, \quad (17)$$

and thus

$$p_j = \frac{\kappa_j^2}{\kappa_j + A\omega_j\varphi_j^2}. \quad (18)$$

The market clearing condition $\omega_j = \bar{\omega}_j$ then yields the equilibrium price

$$p_j = \frac{\kappa_j^2}{\kappa_j + A\bar{\omega}_j\varphi_j^2}, \quad (19)$$

which implies that the equilibrium expected return is

$$\mu_j = \frac{\kappa_j}{p_j} - 1 = A\bar{\omega}_j \frac{\varphi_j^2}{\kappa_j} \quad (20)$$

and the equilibrium return volatility is

$$\sigma_j = \frac{\varphi_j}{p_j} = \frac{\varphi_j}{\kappa_j} (\mu_j + 1). \quad (21)$$

As in the case with lognormally distributed payoffs, both the expected return and the volatility are increasing in the payoff risk and risk aversion, but decreasing in the expected payoff. To simplify notation, we label the risk-free asset as “Stock” $n + 1$ with price $p_{n+1} = 1$, expected payoff $\kappa_{n+1} = 1$, payoff volatility $\varphi_{n+1} = 0$, expected return $\mu_{n+1} = 0$, and return volatility $\sigma_{n+1} = 0$. In addition, we assume the parameters α and β are such that

$$\infty > \mu_1 > \mu_2 > \dots > \mu_n > \mu_{n+1} = 0. \quad (22)$$

Before we proceed, it is useful to clarify the meaning of underdiversification used in this paper.

Definition 1 *An investor is said to underdiversify if he only invests in a small number of stocks.*

Definition 2 *A portfolio is more underdiversified than another if it holds a smaller number of stocks.⁹*

When there are some poor investors in the economy, by the same arguments as illustrated in the previous section, these investors underdiversify when their wealth is low. For example, if their wealth is close to 0, then they only invest in the stock with the highest expected return (i.e., Stock 1). As their wealth increases, poor investors first increase the investment in Stock 1, then add the stock with the second highest expected return (i.e., Stock 2), then increase the investment in both Stock 1 and Stock 2, then add the stock with the third highest expected return (i.e., Stock 3), and so on until they are rich enough to hold the same portfolio as the rich and thus become fully diversified.¹⁰

This result suggests that when the wealth of the poor investors is low, they hold a different portfolio from the rich and therefore no one in economy holds the market

⁹The number of stocks in a portfolio is the most commonly used measure of diversification in the literature (e.g., Blume, Crockett, and Friend (1974), Vissing-Jorgensen (1999), Goetzmann and Kumar (2005)). Two other common measures of diversification are the volatility of a portfolio and the difference between the portfolio weights on stocks and the market portfolio stock weights (e.g., Goetzmann and Kumar (2005)). As shown later, given our model, as wealth increases, the poor invest in a greater number of stocks, the volatility of the portfolio decreases, and the portfolio weights get closer to the market portfolio weights. So all these three measures are highly positively correlated in our model. The advantage of using the number of stocks as a measure for underdiversification is that there is virtually no estimation error.

¹⁰Different from examples in Section 1, we need to take into account the price impact of the poor as they invest more in the stocks when their wealth increases. As we show in the Appendix, as long as the total wealth of the poor is finite, the order of the equilibrium expected returns remains the same as the case without the poor. So the order in which a stock is added is indeed as in (22).

portfolio and idiosyncratic risks are priced in equilibrium, consistent with the findings of Ang et al. (2003).¹¹

We now summarize the main results in the following theorem that is proven in the Appendix.

Theorem 1 *Let μ_j be as defined in (20) such that (22) holds and let \hat{W}_j be as defined in (48) with $\hat{W}_1 \equiv 0$, for $j = 1, 2, \dots, n + 1$.¹² Then*

1. *if $W_p = 0$, then for $j = 1, 2, \dots, n$, the equilibrium price p_j for Stock j is as stated in (19), the equilibrium expected return μ_j is as stated in (20), and the equilibrium return volatility σ_j is as stated in (21);*
2. *if $W_p \in (\hat{W}_{i-1}, \hat{W}_i]$ for some $2 \leq i \leq n + 1$, then a poor investor invests only in the first $i - 1$ stocks. In addition, for $j = 1, 2, \dots, i - 1$, the equilibrium price \bar{p}_j , the equilibrium expected return $\bar{\mu}_j$, the equilibrium return volatility $\bar{\sigma}_j$, and the dollar amount δ_j invested in Stock j are as stated in (42), (43), (44), and (45) respectively, where k is defined as in (47) in the Appendix;*
3. *as W_p approaches \hat{W}_{n+1} , the poor investor's portfolio converges to that of the rich and thus everyone holds the market portfolio;*
4. *if $W_p \in (0, \hat{W}_{n+1})$, then no one holds the market portfolio, CAPM does not hold, and idiosyncratic risks are priced.*

Equation (45) implies that the amount an investor invests in a stock (once selected) decreases with volatility and risk aversion, which in particular implies that less risk-averse investors underdiversify more. Morin and Suarez (1983), Palsson (1996), and Jianakoplos and Bernasek (1998) find that younger investors and male investors are

¹¹Like other heterogeneity models (e.g., Telmer 1993 and Constantinides and Duffie 1996), a representative agent in the economy does not exist.

¹² \equiv should be read as “is defined as.”

less risk averse. Given these findings, our model may help explain the empirical finding that younger or male investors tend to underdiversify more (Mitton and Vorkink 2007).

Theorem 1 implies the following empirically testable predictions:

1. For low initial wealth, poor investors always invest only in a small number of stocks;
2. Less diversified investors tend to choose stocks with high expected returns regardless of risks;
3. Higher moments (e.g., variance and skewness) do not affect stock selection;
4. As the initial wealth of the poor increases, the number of the stocks the poor investors hold also increases;
5. The amount an investor invests in a stock *after* the stock is selected decreases with its volatility;
6. For the same initial wealth, less risk-averse investors invest in a smaller number of stocks;
7. In equilibrium, CAPM does not hold and idiosyncratic risks are priced.

There is an extensive literature (e.g., Calvet, Campbell, and Sodini 2007, Goetzmann and Kumar 2005, and Mitton and Vorkink 2007) showing that a less diversified stock portfolio has a greater expected return, a higher volatility, a greater skewness, and a lower Sharpe ratio. As far as we know, however, there are no existing models that can help explain all of these findings. In contrast, the following theorem shows that our model can.

Theorem 2 *In equilibrium, as W_p increases, the expected return, volatility, and skewness of a poor investor's stock portfolio all decrease. In addition, if the investor holds more than one stock and λ is small, then as W_p increases, the Sharpe ratio also increases.*¹³

As W_p increases, the poor investor's portfolio becomes less underdiversified. Theorem 2 implies that consistent with the empirical evidence, a less underdiversified stock portfolio has a lower expected return, a lower volatility, and a lower skewness. When the total wealth of the poor investors is small relative to that of the rich, a less underdiversified portfolio also has a higher Sharpe ratio. Intuitively, as wealth increases, the poor investors invest in a greater number of stocks with lower expected returns, and thus a less underdiversified portfolio has a lower expected return and lower skewness, and, because of diversification, also a lower volatility. Whether the Sharpe ratio is lower or higher for a more diversified portfolio depends on the relative impact of diversification on the expected return and volatility. As the poor investors buy more stocks, their price impact drives down both the volatility and the expected return. In addition to the price impact, volatility is also driven down by diversification. When the total wealth of the poor investors is small relative to that of the rich, the price impact of the poor investors is small and thus expected return decreases by less than the volatility, and therefore the Sharpe ratio increases.

2.1. Graphical Illustrations

Next we provide some graphical illustrations of our main analytical results both qualitatively and quantitatively. We assume the relative risk-aversion coefficient of an investor with \$20,000 initial wealth is 2. This translates into an absolute risk aver-

¹³If the poor investor only holds one stock, then the Sharpe ratio decreases as he buys more of the stock because of his price impact and the lack of diversification.

sion coefficient of $A = 10^{-4}$. We arbitrarily set the total number of shares for stocks at $\bar{\omega} = 10^4 \bar{\mathbf{1}}$, where $\bar{\mathbf{1}}$ is a vector of 1's. For these illustrations, we then choose parameters α and β such that the top five stocks (Stocks 1-5) have expected returns evenly distributed from 25% to 5% and return volatilities evenly distributed from 55% to 25%. Figure 2 shows that when the wealth of a poor investor is low, he invests in a small number of stocks. For example, with \$20,000, he invests only in four stocks, which is consistent with empirical evidence.¹⁴ For example, Mitton and Vorkink (2007) find that in their database with the portfolios of 78,000 households the median portfolio value with only four stocks is \$21,903.

Only when his wealth rises above \$30,000 does he add the fifth stock. Also Figure 2 shows that for the same wealth, a more risk-averse investor underdiversifies less. To help explain underdiversification shown in Figure 2, we plot the ratios of the marginal utility of investing in the five stocks to that of investing in Stock 5 in Figure 3. Figure 3 shows that when the wealth W_p is low enough (i.e., to the left of point B), the marginal utility of investing in Stock 1 is the highest. As W_p increases, the poor investor invests more in Stock 1 and thus the marginal utility of investing in Stock 1 decreases. At point B, the marginal utility of investing more in Stock 1 becomes equal to that of investing a small amount in Stock 2 and the investor adds Stock 2. Between B and C, the investor increases investment in Stocks 1 and 2 so that the marginal utilities become lower but stay always the same across these two stocks as required by optimality. Beyond point C, the investor adds Stock 3. Between C and D, the investor increases investment in Stocks 1, 2 and 3 so that the marginal utilities are driven even lower but still always stay the same across the three stocks. Similarly, between D and E, the investor adds Stock 4 and invests more in Stocks 1–4 as his

¹⁴Note that since he chooses stocks in the order of their expected returns, the presence of additional stocks with expected returns lower than 5% will not affect his selection of these four stocks.

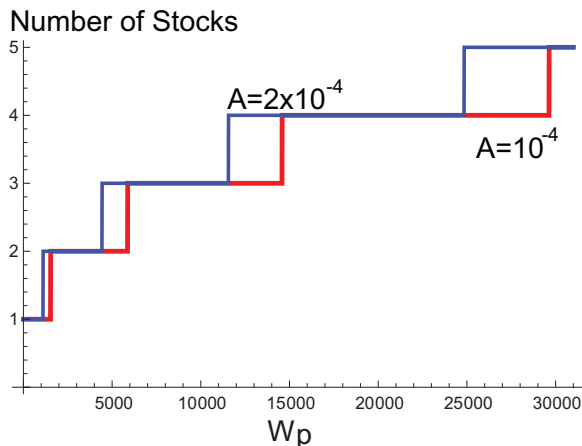


Figure 2: The number of stocks in the poor investor’s stock portfolio against W_p for the following parameters: for $i = 1, 2, \dots, 6$, $\beta_i = 0.25 - (0.25 - 0.01)(i - 1)/5$, $\alpha_i = 20 - (20 - 5)(i - 1)/5$, $\bar{\omega}_i = 10^4$, and $\lambda = 0.1$.

wealth further increases. Beyond point E, the investor also invests in Stock 5.

Figure 4 confirms that a more diversified portfolio has a lower expected return and also a lower volatility. Intuitively, as the wealth of the poor increases, they sequentially add stocks with lower expected returns, which drives down their portfolio expected return. Due to diversification, the portfolio volatility also decreases. When the poor’s wealth increases beyond a critical level, they also fully diversify as the rich, and both the expected return and the volatility of their portfolio are driven down to the lowest. Thus, Figure 4 also suggests that compared to the case without underdiversification, both the expected return and the volatility are higher with underdiversification. Figure 5 shows that the skewness of the portfolio return is also lower for a more diversified portfolio. In addition, it shows that the Sharpe ratio of a more diversified portfolio is higher and therefore diversification improves mean-variance efficiency. All these patterns are consistent with empirical findings such as those in Calvet, Campbell, and Sodini (2007), Goetzmann and Kumar (2005), and

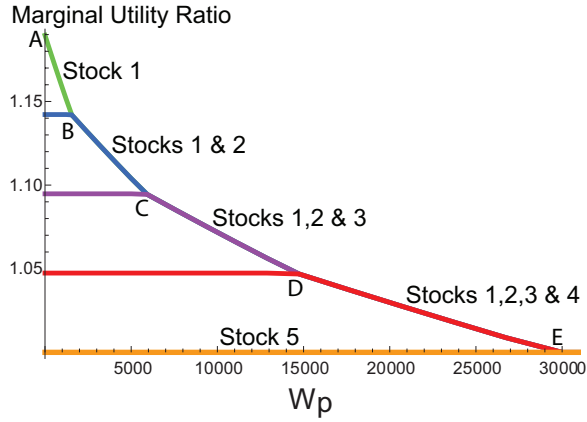


Figure 3: The marginal utility ratios against W_p for the following parameters: for $i = 1, 2, \dots, 6$, $\beta_i = 0.25 - (0.25 - 0.01)(i - 1)/5$, $\alpha_i = 20 - (20 - 5)(i - 1)/5$, $\bar{\omega}_i = 10^4$, and $\lambda = 0.1$.

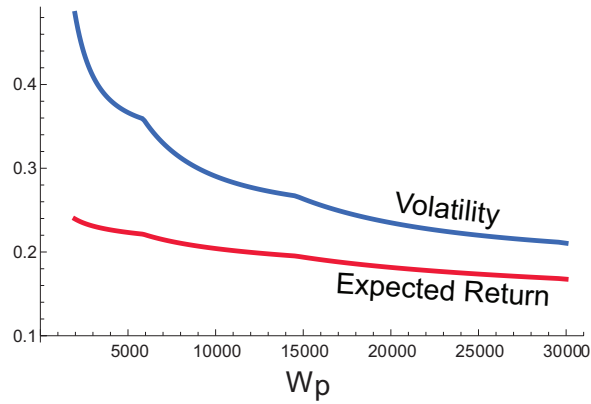


Figure 4: Stock portfolio expected return and volatility against W_p for the following parameters: for $i = 1, 2, \dots, 6$, $\beta_i = 0.25 - (0.25 - 0.01)(i - 1)/5$, $\alpha_i = 20 - (20 - 5)(i - 1)/5$, $\bar{\omega}_i = 10^4$, $\lambda = 0.1$, and $A = 10^{-4}$.

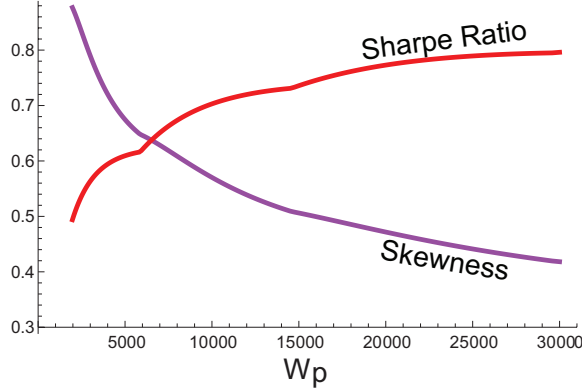


Figure 5: Stock portfolio Sharpe ratio and skewness against wealth W_p for the following parameters: for $i = 1, 2, \dots, 6$, $\beta_i = 0.25 - (0.25 - 0.01)(i - 1)/5$, $\alpha_i = 20 - (20 - 5)(i - 1)/5$, $\bar{\omega}_i = 10^4$, $\lambda = 0.1$, and $A = 10^{-4}$.

Mitton and Vorkink (2007). Although skewness does not impact stock selection in our model, the implied skewness pattern as shown in Figure 5 may appear to indicate that skewness is important for this choice. In our model, the skewness pattern is mainly driven by the fact that as the investor diversifies, he adds stocks with lower expected returns and thus shifts the portfolio return distribution to the left, which results in the reduction in the skewness.

3. Extensions

Next we show that the main results that (1) underdiversification can be a result of solvency requirement, (2) less wealthy investors underdiversify more, and (3) stock selection does not depend on stock return variance or any higher moments hold for general expected utility preferences and general stock payoff distributions. As shown above, equilibrium price impact is not important for any of these results. Accordingly, we restrict our attention to the analysis of the portfolio selection problem. Specifically, we consider a one-period discrete-time portfolio selection model in which an investor

with initial wealth W_p can invest in one risk-free asset and $n \geq 1$ risky stocks and maximizes his expected utility from the end-of-period wealth \widetilde{W}_1 . The utility function $u(W)$ is strictly increasing, strictly concave, and twice continuously differentiable.

According to the consumption commitments literature (e.g., Chetty and Szeidl 2007, Postlewaite, Silverman, and Samuelson 2008), it is costly for households to change certain consumption needs (e.g., durable good consumption). For example, Chetty and Szeidl (2007) show that more than 50% of the average household's budget is committed to ensure a certain critical level of consumption for moderate wealth shocks. On the other hand, for a large negative wealth shock, the consumption level may be lowered to below the critical level. These findings suggest that (1) households tend to commit to a certain critical consumption level and (2) the committed consumption level is generally above the subsistence level and thus the marginal utility at the committed level is finite. Consistent with these findings, we assume that¹⁵

Assumption 1 *The investor commits to a terminal consumption of $\underline{C} \geq 0$ at which the marginal utility is finite and must remain solvent after the committed consumption, i.e., $\widetilde{W}_1 \geq \underline{C} \geq 0$ almost surely and the right derivative $u'(\underline{C}) < \infty$.*

Assumption 1 is also related to the habit formation literature. The key difference from the standard habit formation literature is that the marginal utility at the “habit” level is not infinite, i.e., the “habit” level is above the subsistence level. In other words, even though the investor suffers a huge utility loss when his consumption falls below

¹⁵Alternatively, one can assume a utility function $\hat{u}(W)$ that is defined in \mathbb{R} , where

$$\hat{u}(W) = \begin{cases} u(W), & \text{if } W \geq \underline{C}, \\ u(W) - L, & \text{otherwise,} \end{cases} \quad (23)$$

where $L \gg 0$ represents the large utility loss from not achieving the committed level \underline{C} . It can be shown that in the presence of a solvency constraint, if L is large enough (not necessarily infinite), then an investor starting with a wealth $W_p \geq \underline{C}$ always trades to ensure that the end-of-period wealth \widetilde{W}_1 is never below \underline{C} .

the “habit” level by only a small amount, he can still survive.¹⁶

The risk-free interest rate is normalized to 0. Let \tilde{P} denote the end-of-period gross return vector of the stocks. We assume that the gross return \tilde{P}_i ($i = 1, 2, \dots, n$) is unbounded above and can get arbitrarily close to 0. To ensure solvency, the investor cannot borrow or shortsell. Let $\tilde{z} = \tilde{P} - \bar{1}$ be the return vector, $\mu \equiv (\mu_1, \mu_2, \dots, \mu_n)^\top = E[\tilde{z}]$ be the expected return vector, and $\sigma\sigma^\top \equiv E[(\tilde{z} - \mu)(\tilde{z} - \mu)^\top]$ be the variance-covariance matrix. We assume that stock risk premia are all strictly positive.

Let θ denote the column vector of the dollar amount invested in the stocks. Given $W_p \geq \underline{C}$, the investor’s problem is then¹⁷

$$\max_{\theta} E \left[u(\tilde{W}_1) \right] \quad (24)$$

subject to

$$\tilde{W}_1 = W_p + \theta^\top \tilde{z} \geq \underline{C} \geq 0. \quad (25)$$

Because \tilde{z} is unbounded above and can be arbitrarily close to -1 , element by element, constraint (25) is equivalent to the no-borrowing and no-shortselling constraint:

$$\theta^\top \bar{1} \leq W_p - \underline{C} \quad \text{and} \quad \theta \geq 0. \quad (26)$$

The intuitions that drive the main results on underdiversification in the previous sections still apply for general preferences and general payoff distributions. Specifically, with the no-borrowing and no-shortselling constraint implied by the solvency

¹⁶For our main results, one can also use the typical habit formation utility function form $u(W - \underline{C})$, as long as $u'(0) < \infty$ and the constraint $W \geq \underline{C}$ is imposed. It is worth noting that for modeling habit-formation investors with preferences such that $u'(0) < \infty$ (e.g., non-CRRA HARA utility functions), the constraint $W \geq \underline{C}$ is also necessary.

¹⁷There always exists a unique solution to the investor’s problem, as the feasible set for θ is compact and the objective function is continuous and strictly concave.

requirement, the amount of investment is limited by the initial wealth W_p . If W_p is small, then the investor only invests in the stocks that provide the highest marginal utility. Since the marginal utility at \underline{C} is finite, investments in different stocks provide different levels of marginal utility in general. Therefore, the investor only invests in a small number of the stocks when he is poor. As his wealth increases, he invests more in these stocks and risks increase, which drives down the marginal utility of investing any additional amount in these stocks. Beyond a critical wealth level, the investor adds a new stock that provides the next highest marginal utility. In addition, for the choice of stocks, since the local risk neutrality still holds in this more general setting, higher moments such as variance and skewness are still irrelevant. However, different from the previous section, if stocks are correlated, then the covariance of the return of a stock with the existing portfolio return affects the magnitude of the diversification benefit and thus the marginal utility that this stock can provide. Therefore in addition to expected returns, covariances with the current portfolio also affect stock selection. With these intuitions in mind, we collect our main analytical results in the following theorem that is proven in the Appendix.

Theorem 3 *We have:*

1. *For low enough initial wealth, the investor always underdiversifies;*
2. *As $W_p - \underline{C}$ increases, the investor invests a greater dollar amount in a greater number of stocks;*
3. *Whether a stock is selected into a portfolio or not depends only on its expected return and its covariance with the portfolio, but not on any other moments (e.g., variance and skewness). In particular, the Sharpe ratio is irrelevant for stock selection;*

4. For some utility functions (e.g., CRRA or mean-variance) and some return distributions, it is optimal for the investor to always underdiversify no matter how wealthy he is,¹⁸

5. If $u'(\underline{C}) = \infty$, then investors hold all the stocks as long as $W_p > \underline{C}$.

Part 4 of Theorem 3 suggests under some conditions on preferences and return distributions, no matter how wealthy an investor is, he always underdiversifies. As explained in Section 1, this is because the investor cannot borrow and for some preferences and return distributions, the marginal cost of diversification is still greater than the marginal benefit of diversification even when he invests 100% of his wealth in a proper subset of available stocks. This result is consistent with the empirical evidence that even the rich may underdiversify.

Part 5 and our main results on underdiversification show that the assumption of infinite marginal utility at \underline{C} is critical for the standard diversification result that investors should diversify regardless of their wealth levels.

In our baseline model, to ensure solvency, an investor cannot borrow or shortsell in a discrete-time setting. However, even if an investor is allowed to borrow and shortsell and to trade continuously, as long as he can only borrow or shortsell a limited multiple of the initial wealth (e.g., with margin requirement), our results still hold. This is because when W_p is small enough, a limited multiple of W_p is also small and the investor still underdiversifies. In addition, the local risk neutrality argument still applies and thus, as before, only expected returns and covariances affect stock selection.

¹⁸The standard mean-variance utility is implied by, for example, CARA preferences and normally distributed stock returns.

3.1. Additional examples: Investor with a nontradable asset

Many investors have illiquid assets such as retirement portfolios, houses, and other durable goods. These illiquid assets are typically too costly to liquidate for daily consumptions. In this subsection, we provide some examples to show that our main results still hold and can even be stronger in the presence of illiquid assets. Thus, our model can also help explain why investors with a diversified retirement portfolio underdiversify in the directly held portfolio (e.g., Goetzmann and Kumar 2005, Polkovnichenko 2005).

We adopt the same setup as before but assume that an investor owns one unit of a nontradable asset, whose end-of-period payoff \tilde{N} is a nonnegative random variable that may be (highly) correlated with stocks. We assume that the nontradable asset is held for future consumption beyond the next period and so cannot be used for the next period's committed consumption $\underline{C} \geq 0$. Therefore, the investor requires the terminal *tradable* wealth be above $\underline{C} \geq 0$.

Given $W_p \geq \underline{C}$, the investor's problem is then

$$\max_{\theta} E \left[u(\tilde{W}_1) \right] \quad (27)$$

subject to

$$\tilde{W}_1 = W_p + \theta^\top \tilde{z} + \tilde{N} \quad (28)$$

and

$$W_p + \theta^\top \tilde{z} \geq \underline{C}, \quad (29)$$

which, as before, is equivalent to

$$\theta^\top \bar{\mathbf{1}} \leq W_p - \underline{C} \quad \text{and} \quad \theta \geq 0. \quad (30)$$

Suppose $W_p = \underline{C} + \eta$ with $\eta > 0$ and the investor invests η in Stock i . Then

$$\widetilde{W}_1 = \underline{C} + \eta(\tilde{z}_i + 1) + \tilde{N}.$$

The marginal utility of investing in Stock i is

$$\frac{\partial E[u(\widetilde{W}_1)]}{\partial \eta} = E[u'(\underline{C} + \eta(\tilde{z}_i + 1) + \tilde{N})(\tilde{z}_i + 1)], \quad (31)$$

which implies that, as η approaches 0, the marginal utility converges to

$$\lim_{\eta \downarrow 0} \frac{\partial E[u(\widetilde{W}_1)]}{\partial \eta} = E[u'(\underline{C} + \tilde{N})(\tilde{z}_i + 1)] = E[u'(\underline{C} + \tilde{N})](\mu_i + 1) + \text{Cov}\left(u'(\underline{C} + \tilde{N}), \tilde{z}_i\right), \quad (32)$$

where the last equality follows from the covariance relation $\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$ for any random variables X and Y . Therefore when wealth is low, the investor chooses to invest in only the stock with the combination of expected return and covariance with the nontradable asset that yields the highest marginal utility. Thus, as before, the investor underdiversifies when tradable wealth is low.

In addition, if a stock is positively correlated with the nontradable asset, then $\text{Cov}\left(u'(\underline{C} + \tilde{N}), \tilde{z}_i\right) < 0$ because the utility function is strictly concave. Therefore the marginal utility from this stock will be lowered by the positive correlation and thus the investor would be less willing to add this stock. As shown in the example below, if the nontradable asset is highly correlated with stocks, then the number of stocks directly held can be quite small, because the diversification benefit of additional stocks is small. This may help explain why investors with a diversified retirement portfolio directly hold a relatively small number of stocks.

Finally, while in the absence of an illiquid asset, our basic model implies that

all investors share the same highest expected return risky assets, investors may hold different stocks in this generalized model with an illiquid asset. This is because covariance with the illiquid asset also matters and investors' illiquid asset holdings may be different.

Next, we graphically illustrate the effect of the ownership of a well-diversified retirement portfolio on underdiversification in the directly held portfolio. For simplicity, we again specialize to the mean-variance preference case.

Suppose there are $n > 0$ risky assets and one risk-free asset that the investor can trade. The $(n + 1)^{st}$ risky asset is nontradable. Let W_p be the initial tradable wealth and W_N be the initial value of the nontradable asset. The $(n + 1) \times 1$ expected return vector is μ and the $(n + 1) \times (n + 1)$ variance-covariance matrix is σ . Let the $n \times 1$ vector w be the initial fraction of total wealth ($W_p + W_N$) invested in the tradable risky assets and $w_N \equiv W_N/(W_p + W_N)$ be the initial fraction of total wealth held in the nontradable asset. Then (30) is equivalent to

$$w^\top \bar{1} \leq 1 - \frac{\underline{C} + W_N}{W_p + W_N}, \quad w \geq 0. \quad (33)$$

The mean-variance investor then solves the following problem:

$$\max_w \left[\left(\begin{array}{c} w \\ w_N \end{array} \right)^\top \mu - \frac{1}{2} A \left(\begin{array}{c} w \\ w_N \end{array} \right)^\top \sigma \sigma^\top \left(\begin{array}{c} w \\ w_N \end{array} \right) \right],$$

subject to (33), where $A > 0$ measures the investor's risk aversion.

Suppose there are 50 stocks, an investor's committed consumption is $\underline{C} > 0$, and he holds an equally weighted (i.e., 2% in each stock) nontradable retirement portfolio worth of $W_N = 30\underline{C}$. Figure 6 plots the optimal number of stocks held in the directly held portfolio against W_p/\underline{C} for two risk-aversion levels. This figure shows

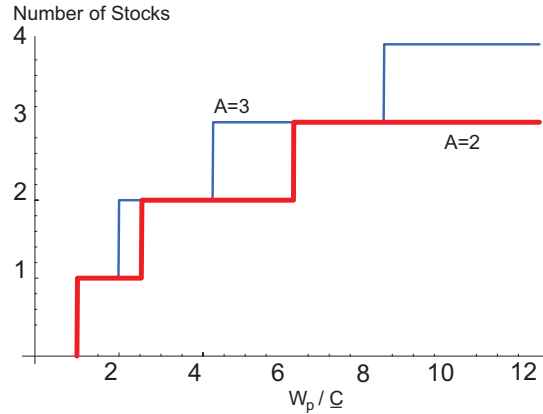


Figure 6: Number of stocks directly held against W_p/\underline{C} given 50 independent stocks. For $i = 1, 2, \dots, 50$, $\mu_i = 0.2 - (0.2 - 0.01)(i - 1)/49$, $\sigma_i = 0.2 - (0.2 - 0.1)(i - 1)/49$, $W_N/\underline{C} = 30$, and the retirement portfolio invests 2% in each stock.

that consistent with the empirical finding, the number of stocks an investor directly holds can be quite small. For example, suppose $\underline{C} = \$10,000$ and thus the investor has $\$300,000$ invested in the nontradable retirement portfolio. Figure 6 shows that if $A = 2$ and the investor's wealth outside the retirement account is smaller than $\$61,500$, then he holds at most two stocks in the directly held portfolio. In fact, for this investor to hold four stocks, he needs to have at least $\$299,000$ wealth outside the retirement account (not shown in the figure). The small number of stocks held in the directly held portfolio is a reflection of the small marginal benefit from additional diversification given an already diversified retirement portfolio.

4. Concluding Remarks

We show that the solvency requirement in the presence of committed consumption can help explain many of the empirical findings on underdiversification and the relevance of idiosyncratic risks for asset pricing. In particular, we demonstrate that investors

always underdiversify when wealth is low and less wealthy investors underdiversify more. In addition, we show that investors choose stocks solely by expected returns and covariances with other stocks, and any other moments (e.g., variance and skewness) are irrelevant for this choice. For investors with a well-diversified illiquid portfolio (e.g., a retirement portfolio), it can be optimal to hold only an even smaller number of stocks directly. In an equilibrium with underdiversification, no one holds the market portfolio and idiosyncratic risks are priced.

While the main results are shown for (quite general) risk-averse expected utility preferences, they also hold for many alternative preferences. For example, the main result that investors underdiversify more when poor than when rich holds as long as all assets do not yield exactly the same marginal utility at zero investment. Therefore this result holds for many expected utility preferences with or without global risk aversion, as well as many non-expected-utility preferences. For expected utility preferences, for example, utility functions can be convex for a certain range of wealth, like the classic Friedman and Savage preferences (Friedman and Savage 1958), as long as concavification exists. For non-expected-utility preferences, this result holds, for example, for disappointment aversion preferences (Gul 2000), recursive preferences (Epstein and Zin 1989), loss aversion preferences (Kahneman and Tversky 1979), and Machina preferences (Machina 1982), as long as one can rank assets by the marginal utility they can yield at zero investment.

Therefore our main results hold for quite general preferences and asset return distributions. Moreover, if the marginal utility is finite at zero wealth (e.g., non-CRRA HARA preferences), then the requirement of solvency itself is sufficient. The generality of these results seems to suggest that underdiversification and thus the relevance of idiosyncratic risks for asset pricing should be the norm, not an exception.

Appendix

In this appendix, we provide the proofs for Theorems 1-3.

PROOF OF THEOREM 1: The case where $W_p = 0$ is the same as the case without poor, which is shown in the text. We now consider the case where $W_p = \eta > 0$. Suppose first η is small. Given the solvency constraint, the poor cannot borrow or shortsell and thus the maximum amount that they can invest in any stock is η . Therefore, as long as investing in different stocks provides different marginal utilities, the investor will choose sequentially the stocks that provide the next highest marginal utility until his budget η is exhausted. We first examine the portfolio allocation problem of the poor investors without taking into account the equilibrium price impact of their trades. The utility from investing a dollar amount $\eta > 0$ in Stock j is

$$U_j(\eta) = E \left[-e^{-A \frac{\eta}{p_j} \tilde{P}_j} \right] = - \left(1 + A \frac{\eta}{p_j} \beta_j \right)^{-\alpha_j}.$$

Accordingly, the marginal utility from investing η in Stock j is

$$U'_j(\eta) = \frac{\frac{A\alpha_j\beta_j}{p_j}}{\left(1 + A \frac{\eta}{p_j} \beta_j \right)^{\alpha_j+1}}, \quad (34)$$

which, as $\eta \downarrow 0$, converges to

$$\lim_{\eta \downarrow 0} U'_j(\eta) = \frac{A\alpha_j\beta_j}{p_j} = A \frac{\kappa_j}{p_j} = A(\mu_j + 1), \quad (35)$$

where the last equality follows from (20). Therefore, irrespective of other moments (e.g., variance and skewness), the stock with the highest expected return yields the greatest marginal utility when the amount of investment η is small. By (22) and the continuity of $U'_j(\eta)$, a poor investor invests the entire amount η in Stock 1, when η is

small enough.

We now derive the new equilibrium price for Stock 1, taking into account the equilibrium price impact of the poor investor's purchase of Stock 1. Let \hat{p}_1 be the new equilibrium price of Stock 1. By (17), the market clearing condition for Stock 1 becomes

$$\lambda \times \frac{\eta}{\hat{p}_1} + 1 \times \left(\frac{\kappa_1}{\hat{p}_1} - 1 \right) \frac{\kappa_1}{A\varphi_1^2} = \bar{\omega}_1,$$

which implies that the new equilibrium price for Stock 1 is

$$\hat{p}_1 = \frac{\kappa_1^2 + A\lambda\eta\varphi_1^2}{\kappa_1 + A\bar{\omega}_1\varphi_1^2} = p_1 \left(1 + A\lambda\eta\frac{\varphi_1^2}{\kappa_1^2} \right), \quad (36)$$

and the new equilibrium expected return becomes

$$\hat{\mu}_1 = \frac{\kappa_1}{\hat{p}_1} - 1 = \mu_1 - \frac{A\lambda\eta\frac{\varphi_1^2}{\kappa_1^2}}{1 + A\lambda\eta\frac{\varphi_1^2}{\kappa_1^2}}(\mu_1 + 1).$$

Therefore, by (22), there exists a small enough $\eta > 0$ such that (22) still holds with μ_1 replaced by $\hat{\mu}_1$.¹⁹ This shows that indeed when their wealth is low enough, poor investors buy only Stock 1 in equilibrium and thus underdiversify.

We next show that as their wealth increases, poor investors first increase the investment in Stock 1, then add the stock with the second highest expected return (i.e., Stock 2), then increase the investment in both Stock 1 and Stock 2, then add the stock with the third highest expected return (i.e., Stock 3), and so on until they are rich enough to hold the same portfolio as the rich and thus become fully diversified.

We show this by induction. Suppose at a higher wealth W_p a poor investor invests only in the first $i - 1$ stocks for $i \geq 2$. Let $\delta_j \geq 0$ denote the dollar amount invested in

¹⁹Since the poor will only buy Stock 1 and by (17), the rich's demand for a stock is independent of any other stocks, the equilibrium prices and expected returns of other stocks remain the same.

Stock j ($j = 1, 2, \dots, n + 1$) with $\delta_j > 0$ only for $j \leq i - 1$ (recall that “Stock” $n + 1$ is the risk free asset). Let \bar{p}_j be the new equilibrium price for Stock j for $1 \leq j \leq n + 1$. For $m \geq i$, since $\delta_m = 0$ and the rich’s demand for a stock is independent of other stocks (as shown in (17)), the equilibrium price for Stock m remains the same as in the case with $\lambda = 0$, i.e., $\bar{p}_m = p_m$. Let

$$\begin{aligned} V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1}) &= E \left[-\exp \left(-A \sum_{j=1}^n \frac{\delta_j}{\bar{p}_j} \tilde{P}_j - A\delta_{n+1} \right) \right] \\ &= -\prod_{j=1}^n \left(1 + A \frac{\delta_j}{\bar{p}_j} \beta_j \right)^{-\alpha_j} e^{-A\delta_{n+1}} \end{aligned} \quad (37)$$

be the value function of the poor. (37) implies that the marginal utility from investing δ_j in Stock j is

$$\frac{\partial V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1})}{\partial \delta_j} = A |V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1})| \frac{\kappa_j}{\bar{p}_j + A\delta_j \varphi_j^2 / \kappa_j}, \quad (38)$$

which shows that for $m \geq i$, the marginal utility at $\delta_m = 0$ is

$$\frac{\partial V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1})}{\partial \delta_m} = A |V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1})| \frac{\kappa_m}{p_m} = A |V(\delta_1, \delta_2, \dots, \delta_n, \delta_{n+1})| (\mu_m + 1). \quad (39)$$

Similar to (35), (39) shows that the marginal utility from investing δ_m in Stock m at $\delta_m = 0$ only depends on its expected return μ_m , but not on any of its higher moments. (39) and (22) then imply that the next stock the poor investor is going to add when his wealth increases enough beyond W_p will be Stock i , the stock with the next highest expected return and thus the highest marginal utility among the remaining stocks.

For investing δ_j in Stock j ($j = 1, 2, \dots, n + 1$) at W_p to be optimal, the marginal utility from each of the first $i - 1$ stocks must be the same and must also be greater

than the marginal utility from investing any positive amount in the rest of the stocks. By (38) and (39), these optimality conditions then imply that

$$\frac{\kappa_j}{\bar{p}_j + A\delta_j\varphi_j^2/\kappa_j} = k + 1, j = 1, 2, \dots, i - 1, \quad (40)$$

for some $k \in [\mu_i, \mu_{i-1})$ that is to be determined later.²⁰

Given the investment of δ_j in Stock j , a similar argument to that for (36) implies that the new equilibrium price of Stock j is

$$\bar{p}_j = \frac{\kappa_j^2 + A\lambda\delta_j\varphi_j^2}{\kappa_j + A\bar{\omega}_j\varphi_j^2}, \quad j = 1, 2, \dots, i - 1. \quad (41)$$

Solving (41) and (40) for \bar{p}_j and δ_j and simplifying, we have that the new equilibrium prices are

$$\bar{p}_j = p_j \frac{\lambda/(k + 1) + 1}{\lambda/(\mu_j + 1) + 1}, \quad j = 1, 2, \dots, i - 1, \quad (42)$$

which yields respectively the new expected returns and volatilities:

$$\bar{\mu}_j = \nu_k k + (1 - \nu_k)\mu_j, \quad j = 1, 2, \dots, i - 1, \quad (43)$$

and

$$\bar{\sigma}_j = \frac{\varphi_j}{\kappa_j}(\bar{\mu}_j + 1), \quad j = 1, 2, \dots, i - 1, \quad (44)$$

where

$$\nu_k = \frac{\lambda}{\lambda + k + 1} \in [0, 1].$$

Thus, the new equilibrium expected returns are weighted averages of the original expected returns (μ_j 's) and k . Since the weight ν_k is the same across all the first $i - 1$

²⁰ $k < \mu_{i-1}$ is because (40) shows that at $\delta_{i-1} = 0$, $k = \mu_{i-1}$, and the left-hand side of (40) is decreasing in δ_j as implied by (40) and (41).

stocks, (43), inequality (22), and $k \in [\mu_i, \mu_{i-1})$ imply that $\bar{\mu}_1 > \bar{\mu}_2 > \dots > \bar{\mu}_{i-1} > k \geq \mu_i > \dots > \mu_n > \mu_{n+1}$. So the poor will indeed invest only in the first $i - 1$ stocks at W_p in equilibrium.

Plugging (42) back into (40) and simplifying, we have that the equilibrium dollar amount invested in Stock j is

$$\delta_j = \frac{\alpha_j(\mu_j - k)}{A(k+1)(\lambda + \mu_j + 1)} = \frac{\bar{\mu}_j + 1}{k+1} \frac{\bar{\mu}_j - k}{A\bar{\sigma}_j^2}, \quad j = 1, 2, \dots, i-1. \quad (45)$$

Without borrowing or shortselling, we must have the following budget constraint

$$W_p = \sum_{j=1}^{i-1} \delta_j = \sum_{j=1}^{i-1} \frac{\alpha_j(\mu_j - k)}{A(k+1)(\lambda + \mu_j + 1)}, \quad (46)$$

which yields that

$$k = \frac{\sum_{j=1}^{i-1} \frac{\alpha_j \mu_j}{\lambda + \mu_j + 1} - A W_p}{\sum_{j=1}^{i-1} \frac{\alpha_j}{\lambda + \mu_j + 1} + A W_p}. \quad (47)$$

(47), (45), and (38) show that as the wealth increases, k decreases and for all $j \leq i - 1$, the dollar amount δ_j invested in Stock j increases, and the marginal utility of investing in Stock j decreases. When the wealth reaches a threshold level at which the marginal utility of investing more in each of the first $i - 1$ stocks is equal to the marginal utility of investing a small amount in Stock i , the investor adds Stock i to his portfolio. By (38), (39), and (40), this threshold wealth level \hat{W}_i above which the poor investor holds Stock i must be such that $k = \mu_i$, which combined with (46) implies that

$$\hat{W}_i = \sum_{j=1}^{i-1} \frac{\alpha_j(\mu_j - \mu_i)}{A(\mu_i + 1)(\lambda + \mu_j + 1)}. \quad (48)$$

By (47), we have $k \in [\mu_i, \mu_{i-1})$ if and only if $W_p \in (\hat{W}_{i-1}, \hat{W}_i]$. As i increases, μ_i decreases, so the threshold wealth level \hat{W}_i increases. Because the above derivation

applies to any $i = 2, 3, \dots, n$, we have shown that for $2 \leq i \leq n$, the poor holds only the stocks with the highest $i - 1$ expected returns if and only if the initial wealth $W_p \in (\hat{W}_{i-1}, \hat{W}_i]$, equivalently if and only if $k \in [\mu_i, \mu_{i-1})$.

Therefore, as wealth increases, the poor investors sequentially add stocks with the next highest expected returns. Eventually, the poor will start to invest in the risk-free asset (which has the lowest expected return) when the poor's wealth increases to a critical level \hat{W}_{n+1} . Beyond this critical level, because investing more in the risk-free asset does not increase risk, (38) and (39) imply that the marginal utility of investing more in the risk-free asset decreases less than that of investing more in any of the risky stocks. Therefore, the investor optimally invests any amount above \hat{W}_{n+1} in the risk-free asset and no additional amount in any of the risky stocks (a standard result for CARA preferences). This implies that \hat{W}_{n+1} is also the critical wealth level at which the poor hold the same unconstrained optimal stock portfolio as the rich. Therefore, to get \hat{W}_{n+1} , we can simply set $k = \mu_{n+1} = 0$ and $i = n + 1$ in (48). This shows that as W_p approaches \hat{W}_{n+1} , the poor's portfolio converges to that of the rich.

Finally, as shown above, as long as $W_p \in (0, \hat{W}_{n+1})$, poor investors only hold a proper subset of the available stocks, while the rich hold all of them. Therefore, the poor and the rich hold different portfolios and thus no one in the economy holds the market portfolio in equilibrium. The market portfolio's weight on Stock j is

$$w_j^M = \frac{\bar{\omega}_j \bar{p}_j}{\sum_{i=1}^n \bar{\omega}_i \bar{p}_i}, j = 1, 2, \dots, n.$$

Direct computation using (42), (43), and (44) shows that CAPM does not hold, i.e.,

$$\bar{\mu}_j - r \neq \beta_{jM}(\bar{\mu}_M - r),$$

where $r = 0$,

$$\bar{\mu}_M = \sum_{i=1}^n w_i^M \bar{\mu}_i,$$

$$\beta_{jM} = \frac{w_j^M \bar{\sigma}_j^2}{\sum_{i=1}^n (w_i^M)^2 \bar{\sigma}_i^2}.$$

Thus idiosyncratic risks are priced. What holds is a modified CAPM equation with a nonzero alpha term. Specifically, we have for $j = 1, 2, \dots, n$,

$$\tilde{r}_j - r = \alpha_j + \beta_j(\tilde{r}_M - r) + \tilde{\varepsilon}_j, \quad (49)$$

where \tilde{r}_j is the Stock j 's return, \tilde{r}_M is the market portfolio return, $\tilde{\varepsilon}_j$ is the mean-zero error term, and

$$\alpha_j \equiv (\bar{\mu}_j - r) - \beta_{jM}(\bar{\mu}_M - r). \quad (50)$$

□

The following lemma will be used repeatedly in the proof of Theorem 2.

Lemma 1 *Given positive integer i , if $x_j > 0$, $b_j > 0$, $\hat{b}_j > 0$, and both x_j and b_j/\hat{b}_j decrease with j for $j = 1, 2, \dots, i$, then*

$$\frac{\sum_{j=1}^i b_j x_j}{\sum_{j=1}^i b_j} \geq \frac{\sum_{j=1}^i \hat{b}_j x_j}{\sum_{j=1}^i \hat{b}_j}. \quad (51)$$

PROOF OF LEMMA 1: Let $\pi_j = \frac{b_j}{\sum_{l=1}^i b_l}$ and $\hat{\pi}_j = \frac{\hat{b}_j}{\sum_{l=1}^i \hat{b}_l}$. Then the left- (right-) hand side of (51) can be viewed as the mean of a random variable \tilde{x} with support $\{x_1, x_2, \dots, x_i\}$ and a probability of π_j ($\hat{\pi}_j$, respectively) for x_j ($j = 1, 2, \dots, i$). Next we show that the assumption that both x_j and b_j/\hat{b}_j decrease with j implies that the probability distribution for the left-hand side of (51) first stochastically dominates that for the right-hand side and thus the mean on the left-hand side is greater than that

on the right-hand side, and accordingly, (51) holds. Since $x_j > 0$ decreases with j , for the first stochastic dominance we only need to show that for any $1 \leq m \leq i - 1$,

$$\sum_{j=1}^m \pi_j \geq \sum_{j=1}^m \hat{\pi}_j, \quad (52)$$

which is equivalent to

$$\frac{\sum_{j=1}^m b_j}{\sum_{j=1}^m \hat{b}_j} \geq \frac{\sum_{j=m+1}^i b_j}{\sum_{j=m+1}^i \hat{b}_j}. \quad (53)$$

Since $b_j/\hat{b}_j > 0$ decreases with j , we have

$$\frac{\sum_{j=1}^m b_j}{\sum_{j=1}^m \hat{b}_j} = \frac{\sum_{j=1}^m \hat{b}_j \frac{b_j}{\hat{b}_j}}{\sum_{j=1}^m \hat{b}_j} \geq \frac{\sum_{j=1}^m \hat{b}_j \frac{b_m}{\hat{b}_m}}{\sum_{j=1}^m \hat{b}_j} = \frac{b_m}{\hat{b}_m} = \frac{\sum_{j=m+1}^i \hat{b}_j \frac{b_m}{\hat{b}_m}}{\sum_{j=m+1}^i \hat{b}_j} \geq \frac{\sum_{j=m+1}^i \hat{b}_j \frac{b_j}{\hat{b}_j}}{\sum_{j=m+1}^i \hat{b}_j} = \frac{\sum_{j=m+1}^i b_j}{\sum_{j=m+1}^i \hat{b}_j}. \quad (54)$$

Therefore (53) indeed holds and thus (51) also holds. \square

Intuitively, Lemma 1 holds because b_j assigns higher weights to larger values of \tilde{x} than \hat{b}_j . We are now ready to prove Theorem 2.

PROOF OF THEOREM 2: Since as shown in (47), k decreases as W_p increases, we can equivalently show how the properties of the poor's portfolio change with k . The m th central moment of the poor's stock portfolio gross return is

$$\xi_m(k) = \sum_{j=1}^i w_j^m \frac{M_{mj}}{\bar{p}_j^m}, \quad (55)$$

where the portfolio weight

$$w_j = \frac{\bar{\delta}_j}{\sum_{j=1}^i \bar{\delta}_j}, \quad j = 1, 2, \dots, i,$$

and M_{mj} is the m th central moment of the payoff of the j th stock. Using (42) and

(45), we have

$$\xi_m(k) = \frac{\sum_{j=1}^i C_{mj}(\mu_j - k)^m}{((\lambda/(k+1) + 1) \sum_{j=1}^i a_j(\mu_j - k))^m}, \quad (56)$$

where $C_{mj} \equiv M_{mj}/\beta_j^m$ and $a_j \equiv \alpha_j/(\lambda + \mu_j + 1)$. Computing $\xi'_m(k)$ and rearranging yield that $\xi_m(k)$ is strictly increasing in k for all $\lambda > 0$ if and only if

$$\frac{\sum_{j=1}^i b_{mj}(\mu_j + 1)}{\sum_{j=1}^i b_{mj}} \geq \frac{\sum_{j=1}^i a_j(\mu_j + 1)}{\sum_{j=1}^i a_j}, \quad (57)$$

where

$$b_{mj} = C_{mj}(\mu_j - k)^{m-1}.$$

By Lemma 1, for (57) to hold, we only need to show b_{mj}/a_j decreases with j . First, consider the expected return. Since $M_{1j} = \alpha_j\beta_j$, we have $b_{1j} = \alpha_j$. So

$$\frac{b_{1j}}{a_j} = \lambda + \mu_j + 1,$$

which indeed decreases with j by (22) and thus the expected return decreases as k decreases. Next, consider the return volatility. Since $M_{2j} = \alpha_j\beta_j^2$, we have $b_{2j} = \alpha_j(\mu_j - k)$ and

$$\frac{b_{2j}}{a_j} = (\mu_j - k)(\lambda + \mu_j + 1),$$

which also decreases with j by (22) and thus the volatility also decreases as k decreases.

The skewness of the stock portfolio is equal to

$$s(k) \equiv \frac{\xi_3(k)}{\xi_2(k)^{3/2}} = \frac{\sum_{j=1}^i C_{3j}(\mu_j - k)^3}{\left(\sum_{j=1}^i C_{2j}(\mu_j - k)^2\right)^{3/2}}.$$

It is easy to verify that $C_{3j} = 2\alpha_j$, $C_{2j} = \alpha_j$, and the skewness is increasing in k if

and only if

$$\frac{\sum_{j=1}^i b_j(\mu_j - k)}{\sum_{j=1}^i b_j} \geq \frac{\sum_{j=1}^i \hat{b}_j(\mu_j - k)}{\sum_{j=1}^i \hat{b}_j}, \quad (58)$$

where

$$b_j = \alpha_j(\mu_j - k)^2, \quad \hat{b}_j = \alpha_j(\mu_j - k).$$

Since $b_j/\hat{b}_j = \mu_j - k$, which decreases with j , by Lemma 1, we have (58) holds and therefore the skewness of the stock portfolio also decreases as k decreases.

For $W_p > \hat{W}_2$, the poor investor holds at least two stocks, i.e., $i \geq 2$. The Sharpe ratio of the stock portfolio at $\lambda = 0$ is

$$SR(k) \equiv \frac{\xi_1(k) - 1}{\sqrt{\xi_2(k)}} = \frac{\sum_{j=1}^i (C_{1j} - a_j)(\mu_j - k)}{\sqrt{\sum_{j=1}^i C_{2j}(\mu_j - k)^2}}.$$

Computing $SR'(k)$ shows that the Sharpe ratio strictly *decreases* in k if and only if

$$\frac{\sum_{j=1}^i d_j(\mu_j - k)}{\sum_{j=1}^i d_j} > \frac{\sum_{j=1}^i \hat{d}_j(\mu_j - k)}{\sum_{j=1}^i \hat{d}_j}, \quad (59)$$

where

$$d_j = \alpha_j(\mu_j - k), \quad \hat{d}_j = \alpha_j \frac{\mu_j}{\mu_j + 1}.$$

Since $d_j/\hat{d}_j = \mu_j - k + 1 - k/\mu_j$, which strictly decreases with j , by Lemma 1, we have that (59) holds and therefore by continuity the Sharpe ratio of the stock portfolio increases as k decreases when λ is small. \square

PROOF OF THEOREM 3: Define the adjusted initial wealth $W_p^A \equiv W_p - \underline{C}$. When $W_p^A = 0$, obviously, the investor can only invest in the risk-free asset. Now suppose his wealth increases to $W_p^A = \eta$, where $\eta > 0$ is small. Since he cannot borrow or shortsell, the most he can invest in stocks is η . Given this restriction, he will first invest in the stock that provides the highest marginal utility. To identify the stock

that provides the highest marginal utility, we first suppose he invests \underline{C} in the risk-free asset and η in Stock j , then the end-of-period wealth

$$\widetilde{W}_1 = \underline{C} + \eta(\tilde{z}_j + 1).$$

As η approaches 0, the investor's marginal utility from investing in Stock j converges to

$$\lim_{\eta \downarrow 0} \frac{\partial E[u(\widetilde{W}_1)]}{\partial \eta} = \lim_{\eta \downarrow 0} E[h(\eta, \tilde{z}_j)] = E[h(0, \tilde{z}_j)] = E[u'(\underline{C})(\tilde{z}_j + 1)] = u'(\underline{C})(\mu_j + 1), \quad (60)$$

where

$$h(\eta, \tilde{z}_j) \equiv u'(\underline{C} + \eta(\tilde{z}_j + 1))(\tilde{z}_j + 1),$$

the first equality in (60) follows from Leibnitz's rule, and the second equality follows from the Monotone Convergence Theorem (e.g., Williams 1991, p. 59) because $h(\eta, \tilde{z}_j)$ is strictly decreasing in η since $\partial h(\eta, \tilde{z}_j)/\partial \eta = u''(\underline{C} + \eta(\tilde{z}_j + 1))(\tilde{z}_j + 1)^2 < 0$ by the strict concavity of $u(\cdot)$. Therefore when η is small enough, since $\infty > u'(\underline{C}) > 0$, the marginal utility from investing in the stock with the highest expected return is strictly the highest. Thus, if wealth W_p is slightly above the committed level \underline{C} , the investor invests \underline{C} in the risk-free asset and the rest in the stock with the highest expected return, say, Stock 1. Higher moments, such as variance, skewness, and kurtosis, do not affect this choice.

We now prove by induction. Suppose that, given the initial adjusted wealth W_p^A , the investor optimally invests \underline{C} in the risk-free asset, $\theta_i > 0$ in Stock i for $1 \leq i \leq m < n$, and 0 in the rest of the stocks. Since the investor is investing in only $m < n$ of n stocks, the constraint $\sum_{i=1}^m \theta_i \leq W_p - \underline{C} = W_p^A$ must be still binding. This is because the marginal utility of investing a bit in the rest of the stocks is

strictly greater than that of investing more in the risk-free asset, since the expected stock returns are strictly greater than the risk-free rate. The end-of-period wealth is $\widetilde{W}_1 = \underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1)$. The Lagrangian is

$$V(W_p^A) = \max_{\theta \geq 0} \left\{ E \left[u \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) \right] + \nu \left(W_p^A - \sum_{i=1}^m \theta_i \right) \right\},$$

where ν is the Lagrangian multiplier. From the first-order conditions, we have that the marginal utility of investing in Stock i for $1 \leq i \leq m$ is

$$\frac{\partial E \left[u \left(\widetilde{W}_1 \right) \right]}{\partial \theta_i} = E[u'(\widetilde{W}_1)(\tilde{z}_i + 1)] = \nu. \quad (61)$$

By the Envelop theorem, we have

$$\nu = \frac{\partial V(W_p^A)}{\partial W_p^A} > 0,$$

since utility is strictly increasing in wealth. Therefore, at the initial level W_p^A , the marginal utilities of investing in these m stocks are equal, strictly positive, and strictly greater than the marginal utilities of investing more than \underline{C} in the risk-free asset and the rest of the stocks (because it is optimal for the investor not to hold the rest of the stocks). Since

$$\frac{\partial E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) (\tilde{z}_i + 1) \right]}{\partial \theta_i} = E[u''(\widetilde{W}_1)(\tilde{z}_i + 1)^2] < 0,$$

the marginal utility of investing in Stock i decreases as the investment in the stock increases. Thus, as the adjusted initial wealth increases from W_p^A , the investor will

increase the investment in all these m stocks, i.e.,²¹

$$\frac{\partial \theta_i}{\partial W_p^A} > 0, \quad i = 1, 2, \dots, m. \quad (62)$$

Similar to (60), the marginal utility of investing a small amount in a new Stock j ($j > m$) at W_p^A is

$$\begin{aligned} E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) (\tilde{z}_j + 1) \right] &= E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) \right] (\mu_j + 1) \\ &+ \text{Cov} \left(u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right), \tilde{z}_j \right). \end{aligned} \quad (63)$$

Therefore, when it becomes optimal to add another stock to the portfolio, which stock to add only depends on its expected return and its covariance with the stocks in the current portfolio. Other moments of Stock j , such as variance and skewness, are irrelevant for this choice. If stocks are uncorrelated, then this choice only depends on the expected return of Stock j .

We next derive the critical adjusted wealth level \hat{W}_{m+1}^A above which it is optimal to add another stock, say, Stock $m+1$. By (61) and (63) with $j = m+1$, we must have that at this critical level, the marginal utilities of investing in these $m+1$ stocks are exactly the same, i.e., for $i = 1, 2, \dots, m$, we have

$$E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) (\tilde{z}_i + 1) \right] = E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) (\tilde{z}_{m+1} + 1) \right], \quad (64)$$

²¹At W_p^A , the marginal utilities across all the m stocks are the same. When the wealth increases above W_p^A , if the investor increases investment only in some of the m stocks, then the marginal utilities from these stocks would be lowered and the marginal utilities from the rest of the m stocks would be strictly higher, which is a contradiction to optimality.

which can be simplified to

$$E \left[u' \left(\underline{C} + \sum_{i=1}^m \theta_i (\tilde{z}_i + 1) \right) (\tilde{z}_i - \tilde{z}_{m+1}) \right] = 0, \quad i = 1, 2, \dots, m. \quad (65)$$

Therefore, we have $\hat{W}_{m+1}^A = \sum_{i=1}^m \theta_i$ where θ_i 's are the solution to (65). By (62), we have that as long as the constraint is binding, the critical wealth level \hat{W}_i^A at which it is optimal to add Stock i strictly increases with i . Combining these results shows that both the number and the amount of stocks optimally held increase as the adjusted wealth W_p^A increases. This completes the proof of Parts 1-3.

Now we show Part 4. If the investor has a CRRA utility, i.e.,

$$u(W) = \frac{W^{1-\gamma}}{1-\gamma},$$

then equation (65) is equivalent to

$$E \left[u' \left(1 + \sum_{i=1}^m w_i \tilde{z}_i \right) (\tilde{z}_i - \tilde{z}_{m+1}) \right] = 0, \quad i = 1, 2, \dots, m, \quad (66)$$

where $w_i \equiv \theta_i/W_p$ represent the fraction of the initial wealth W_p invested in stock i . And the constraints (26) become

$$\sum_{i=1}^n w_i \leq 1 - \frac{C}{W_p} \quad \text{and} \quad w_i \geq 0, \quad i = 1, 2, \dots, n.$$

Due to the no-borrowing constraint, if the w_i 's that solve (66) are such that $\sum_{i=1}^m w_i > 1$ for some $m < n$, then the investor never holds more than m stocks, no matter how wealthy he is. To show that this can happen, suppose we have

$$E [u' (1 + w_1 \tilde{z}_1) (\tilde{z}_1 - \tilde{z}_j)] > 0 \quad (67)$$

for all $w_1 \in [0, 1]$ and $j = 2, 3, \dots, n$, which can hold if $\mu_1 - \mu_j$ is large enough, Stock 1's volatility is small enough, and Stock 1 is independent of other stocks. Then it is never optimal for the investor to hold more than one stock. This is due to the fact that the marginal utility of investing in Stock 1 is strictly greater than that of investing in any other stocks for all feasible w_1 (i.e., for $w_1 \in [0, 1]$). A similar proof applies to mean-variance preferences.

For Part 5, if $u'(\underline{C})$ is infinite, then (60) implies that all stocks have the same (infinite) marginal utility at \underline{C} irrespective of their expected returns and risks. It is therefore optimal to hold all the stocks even when the wealth of the poor investor is just slightly above the committed level \underline{C} . As the wealth of the poor increases, he will increase the investment in every stock such that the marginal utility from the investment in each stock stays the same. Therefore, if $u'(\underline{C})$ is infinite, an investor always invests in all the stocks as long as his wealth $W_p > \underline{C}$, as predicted in the standard portfolio selection theory. \square

References

- Barberis, N. C., and M. Huang. 2007. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, Working Paper, Yale University.
- Blume, M. E., J. Crockett, and I. Friend. 1974. Stockholders in the United States: Characteristics and Trends, *Survey of Current Business* 54:16–40.
- Blume, M., and I. Friend. 1975. The Asset Structure of Individual Portfolios and Some Implications for Utility Functions, *Journal of Finance* 30:585–603.
- Boehmer, E., C. M. Jones, and X. Zhang. 2005. Which Shorts Are Informed? Working Paper, Columbia University.
- Boyle, P., L. Garlappi, R. Uppal, and T. Wang. 2009. Keynes Meets Markowitz: The Trade-off between Familiarity and Diversification, Working Paper, London Business School.
- Brennan, M. J. 1975. The Optimal Number of Securities in a Risky Asset Portfolio when There Are Fixed Costs of Transacting: Theory and Some Empirical Results, *Journal of Financial and Quantitative Analysis* 10:483–496.
- Bucks, B. K., A. B. Kennickell, and K. B. Moore. 2006. Recent Changes in U.S. Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances, *Federal Reserve Bulletin*, 1–38.
- Calvet, L. E., Campbell, J. Y., and P. Sodini. 2007. Down or Out: Assessing the Welfare Costs of Household Investment Mistakes, *Journal of Political Economy* 115:707–747.
- Campbell, J. Y. 2006. Household Finance, *Journal of Finance* 61:239–244.

- Campbell, J. Y., L. Martin, B. G. Malkiel, and Y. Xu. 2001. Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, *Journal of Finance* 56:1–43.
- Chetty R., and Szeidl, A. 2007. Consumption Commitments and Risk Preferences, *Quarterly Journal of Economics* 122:831–877.
- Constantinides, G. M. 1990. Habit Formation: A Resolution of the Equity Premium Puzzle, *Journal of Political Economy* 98:519–543.
- Constantinides, G., and D. Duffie. 1996. Asset Pricing with Heterogeneous Consumers, *Journal of Political Economy* 104:219–240.
- Cuoco, D., and H. Liu. 2000. A Martingale Characterization of Consumption Choices and Hedging Costs with Margin Requirements, *Mathematical Finance* 10:355–385.
- Cuoco, D., and H. Liu. 2006. An Analysis of VaR-based Capital Requirements, *Journal of Financial Intermediation* 15:362–394.
- Dybvig, P. H. 1984. Short Sales Restrictions and Kinks on the Mean Variance Frontier, *Journal of Finance* 36:239–244.
- Dybvig, P. H. 1995. Duesenberry’s Ratcheting of Consumption: Optimal Dynamic Allocation and Investment Given Intolerance for any Decline in Standard of Living, *Review of Economic Studies* 62:287–313.
- Epstein, L. G., and S. E. Zin. 1989. Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns. *Econometrica* 57:937–969.
- Friedman, M., and L. J. Savage. 1948. The Utility Analysis of Choices Involving Risk, *Journal of Political Economy* 56:279–304.

- Goetzmann, W. N., and A. Kumar. 2005. Why Do Individual Investors Hold Underdiversified Portfolios? Working Paper, Yale University.
- Gul, F. 1991. A Theory of Disappointment Aversion, *Econometrica* 59:667–686.
- Huberman, G. 2001. Familiarity Breeds Investment, *Review of Financial Studies* 14:659–680.
- Kraus, A., and R. Litzenberger. 1976. Skewness Preference and the Valuation of Risk Assets, *Journal of Finance* 31:1085–1100.
- Kroll, Y., H. Levy, and A. Rapoport. 1988a. Experimental Tests of the Mean-variance Model for Portfolio Selection, *Organizational Behavior and Human Decision Processes* 42:388–410.
- Kroll, Y., H. Levy, and A. Rapoport. 1988b. Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model, *American Economic Review* 78:500–519.
- Machina, M. J. 1982. “Expected Utility” Analysis without the Independence Axiom, *Econometrica* 50:277–323.
- Merton, R. C. 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42:483–510.
- Mitton, T., and K. Vorkink. 2007. Equilibrium Underdiversification and the Preference for Skewness, *Review of Financial Studies* 20:1255–1288.
- Morin, R. A., and F. Suarez. 1983. Risk Aversion Revisited, *Journal of Finance* 38:1201–1216.

- Nieuwerburgh, S. V., and L. Veldkamp. 2005. Information Acquisition and Portfolio Underdiversification, Working Paper, New York University.
- Palsson, J. M. 1996. Does the Degree of Relative Risk Aversion Vary with Household Characteristics? *Journal of Economic Psychology* 17:771–787.
- Polkovnichenko, V. 2005. Household Portfolio Diversification: A Case for Rank Dependent Preferences, *Review of Financial Studies* 18:1467–1502.
- Postlewaite, A., D. Silverman, and L. Samuelson. 2008. Consumption Commitments and Employment Contracts, *Review of Economic Studies* 75:559–578.
- Ross, S. A. 1977. The Capital Asset Pricing Model (CAPM), Short-sale Restrictions and Related Issues, *Journal of Finance* 32:177–183.
- Telmer, C. 1993. Asset-Pricing Puzzles and Incomplete Markets, *Journal of Finance* 48:1803–1832.
- Uppal, R., and T. Wang. 2003. Model Misspecification and Underdiversification, *Journal of Finance* 58:2465–2486.
- Vissing-Jørgensen, A. 1999. Comments on “Stock Prices and Fundamentals” by John Heaton and Deborah Lucas, *NBER Macroeconomics Annual*.
- Vissing-Jørgensen, A. 2002. Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structure, Working Paper, University of Chicago.
- Williams, D. 1991. Probability with Martingales, Cambridge, Cambridge University Press.