**Market Equilibrium and the Cost of Capital with Heterogeneous Investment Horizons**

**Online Appendix**

**Section S1– Extension of Theorem 1 to Intermediate Income and Consumption**

Consider investors who maximize time-separable expected utility defined over both intermediate and terminal consumption of the form: $EU = E(\sum_{t=0}^{T} \rho^{-t} u_t(C_t))$, where $C_t$ is the period $t$ consumption and $\rho$ is a discount rate reflecting time preference. The utility function may generally be age-dependent, hence the subscript $t$ in $u_t$. The only assumption about the utility function is that it is non-decreasing, i.e. $u_t' \geq 0$ for all $t$. The investor may also have a stream of stochastic and periodic time-dependent labor income, $\tilde{Y}_t$. Following Fama and Schwert (1977) and Cocco et. al. (2005), it is assumed that $\tilde{Y}_t$ is independent of stock returns, and also across time (but it could be correlated across investors). As before, expectations about the return parameters are assumed to be homogeneous, the 1-period returns are assumed to be normally distributed, and returns are independent over time. Consider strategy $G$ of investing in portfolio $G_1$ in period 1, in portfolio $G_2$ in period 2, etc., where at least some of these portfolios are mean-variance inefficient. The alternative strategy $F$ implies investment in portfolio $F_1$ vertically above $G_1$, in period 1, in $F_2$ vertically above $G_2$ in period 2, etc., as illustrated by Figure 1. Consider an investor who follows strategy $G$ and dynamically chooses his optimal consumption in each period so as to maximize his lifetime expected utility. We prove that if the investor invests in strategy $F$ instead, he can obtain the exact same stochastic consumption stream, but with a terminal wealth distribution that FSD dominates the terminal wealth distribution obtained by investing following strategy $G$.

Assume, without loss of generality, that the period-$t$ labor income and the portfolio returns are realized, and then the consumption decision for the next period is made. Namely, at time $t$ the labor income and returns for the period (*t-1*, $t$) are realized, and then the investor determines $C_t$, which is the optimal consumption for the time period ($t$, $t$+1). As the labor income and returns are random variables, so is the optimal consumption. We denote the optimal period-$t$ consumption for an investor who follows strategy $G$ by $\tilde{C}_t^G(\tilde{R}_t^G, \tilde{Y}_t)$, where the "~" indicates that the optimal consumption is a random variable that depends on the random return and labor realizations (and possibly on other parameters, such as current wealth, expectations, and preferences). To prove our result, we do not need to explicitly solve for the optimal $\tilde{C}_t^G$.

This is a big advantage, as the optimal consumption depends on the investor's utility function, and our goal here is a general result for all non-decreasing utility functions.

If the investor invests in portfolio $G_1$ in period 1, her wealth at the end of the period is given by:

$$\widetilde{W}_1^G = W_0 \tilde{R}_1^G + \tilde{Y}_1 - \tilde{C}_1^G(\tilde{R}_1^G, \tilde{Y}_1), \tag{S1}$$

where $W_0$ denotes the initial wealth (after the initial consumption, $C_0$), and $\tilde{R}_1^G$ is the total return on portfolio $G_1$ in period 1 (i.e. 1+rate of return). Now consider the following alternative: investing in the FSD dominating portfolio $F_1$. We denote the consumption of the investor holding portfolio $F_1$ by $\tilde{C}_1^F(\tilde{R}_1^F, \tilde{Y}_1)$. Below we prove that $\tilde{C}_1^F$ can be constructed so that it has the following two properties:

i) $\tilde{C}_1^F$ has the exact same univariate distribution as $\tilde{C}_1^G$;

ii) $\tilde{C}_1^F$ ensures that $\widetilde{W}_1^F$ dominates $\widetilde{W}_1^G$ by FSD,

where $\widetilde{W}_1^F$ is the wealth of the investor investing in the portfolio $F_1$ at the end of period 1, which is given by:

$$\widetilde{W}_1^F = W_0 \tilde{R}_1^F + \tilde{Y}_1 - \tilde{C}_1^F(\tilde{R}_1^F, \tilde{Y}_1). \tag{S2}$$

Thus, the investor obtains the same distribution of intermediate consumption, but an FSD dominating distribution of period-1 wealth. Note that in contrast to $\tilde{C}_1^G$, $\tilde{C}_1^F$ is *not* necessarily the optimal consumption for the investor holding portfolio $F_1$. As we prove below dominance with non-optimal consumption, such dominance a *fortiori* exists with optimal consumption - the investor holding portfolio $F_1$ could potentially obtain even higher expected utility by choosing her consumption optimally.

$W_0 \tilde{R}_1^F$ FSD dominates $W_0 \tilde{R}_1^G$ (because the return on portfolio $F_1$, $\tilde{R}_1^F$, FSD dominates the return on portfolio $G_1$, $\tilde{R}_1^G$). Note that in general, if a random variable $\tilde{x}$ FSD dominates another random variable $\tilde{y}$, then adding a third random variable $\tilde{z}$ to both $\tilde{x}$ and $\tilde{y}$ preserves the dominance relationship if $\tilde{z}$ is independent of both $\tilde{x}$ and $\tilde{y}$ (Levy and Sarnat 1971). If $\tilde{z}$ is not independent of $\tilde{x}$ and $\tilde{y}$, the dominance may remain, but it is not guaranteed. As the labor income is independent of the stock and bond returns, we have that $W_0 \tilde{R}_1^F + \tilde{Y}_1$ FSD dominates $W_0 \tilde{R}_1^G + \tilde{Y}_1$. In contrast, the consumption generally *does* depend on the labor income and on the portfolio return realization, and therefore the dominance of $W_0 \tilde{R}_1^F + \tilde{Y}_1 - \tilde{C}_1^F(\tilde{R}_1^F, \tilde{Y}_1)$ over $W_0 \tilde{R}_1^G + \tilde{Y}_1 - \tilde{C}_1^G(\tilde{R}_1^G, \tilde{Y}_1)$ is not guaranteed. However, we can define $\tilde{C}_1^F(\tilde{R}_1^F, \tilde{Y}_1)$ to have properties i) and ii) above, ensuring dominance. To do so, consider a particular labor income realization $Y_1$. Given this value, the optimal consumption of the investor holding portfolio $G_1$ is a function of the portfolio $G_1$ return, i.e. $C_1^G(R_1^G, Y_1)$. For the investor holding portfolio $F_1$, we choose the consumption as a function of the stock portfolio return $R_1^F$ in the following way: if the portfolio $F_1$ return realization is $R_1^F$, the investor consumes $C_1^F(R_1^F, Y_1) \equiv C_1^G(R_1^G, Y_1)$, where $R_1^G$ is the portfolio $G_1$ return that satisfies:

$$G(R_1^G) = F(R_1^F), \tag{S3}$$

and $G(\cdot)$ and $F(\cdot)$ denote the cumulative distribution functions of portfolios $G_1$ and $F_1$, respectively. Thus, $R_1^G$ is defined by:

$$R_1^G \equiv Q_G\big(F(R_1^F)\big), \tag{S4}$$

where $Q_G$ is quantile distribution of portfolio $G_1$. This implies that in the case of a portfolio $F_1$ return realization of $R_1^F$, the investor consumes what she would have consumed if holding portfolio $G_1$and realizing a return $R_1^G$ with the same quantile value as $R_1^F$.

By construction, $\tilde{C}_1^F$ and $\tilde{C}_1^G$ have the same univariate distribution. To see that $\widetilde{W}_1^F$ dominates $\widetilde{W}_1^G$ by FSD, consider a specific quantile value $q$, and define the returns $R_1^F$ and $R_1^G$ by $F(R_1^F) = G(R_1^G) = q$. As $\tilde{R}_1^F$ FSD dominates $\tilde{R}_1^G$, we are ensured that $R_1^F \geq R_1^G$ (and this is true for any value of $q$, see Levy and Kroll 1978, Theorem 1). By construction, we have $C_1^F(R_1^F, Y_1) = C_1^G(R_1^G, Y_1)$. Thus:

$$W_1^F = W_0 R_1^F + Y_1 - C_1^F(R_1^F, Y_1) \;\geq\; W_1^B = W_0 R_1^G + Y_1 - C_1^G(R_1^G, Y_1). \tag{S5}$$

As this holds for any quantile value $q$ and any labor income realization $Y_1$, we have that $\widetilde{W}_1^F$ dominates $\widetilde{W}_1^G$ by FSD (Levy and Kroll 1978, Theorem 1).

To illustrate the idea behind eq.(S5), it is convenient to consider the case of a discrete return distribution with $N$ equally-likely outcomes. Let us arrange the $N$ possible period-1 portfolio $G_1$ returns by ascending order: $R_1^{G1} \leq R_1^{G2} \leq \cdots \leq R_1^{GN}$, and similarly for the returns of portfolio $F_1$: $R_1^{F1} \leq R_1^{F2} \leq \cdots \leq R_1^{FN}$. The FSD dominance of portfolio $F_1$ over portfolio $G_1$ implies that:

$$\begin{aligned} &R_1^{G1} \leq R_1^{F1} \\ &R_1^{G2} \leq R_1^{F2} \\ &\quad\vdots \end{aligned} \tag{S6}$$

$$R_1^{GN} \leq R_1^{FN},$$

Eqs.(S3) and (S4) imply that if the worst portfolio $F_1$ return ($R_1^{F1}$) is realized, the investor consumes what he would have optimally consumed if he had held portfolio $G_1$, and the worst portfolio $G_1$ return ($R_1^{G1}$) would have been realized. Hence $C_1^F(R_1^{F1}, Y_1) \equiv C_1^G(R_1^{G1}, Y_1)$. Generally, eq.(S3) implies $C_1^F(R_1^{Fi}, Y_1) \equiv C_1^G(R_1^{Gi}, Y_1)$ for all *i*=1…*N*. We can now write the wealth after labor income and consumption according to this same order of ascending returns:

$$\begin{aligned}
&W_0 R_1^{G1} + Y_1 - C_1^G(R_1^{G1}, Y_1) \leq W_0 R_1^{F1} + Y_1 - C_1^G(R_1^{G1}, Y_1) \\
&W_0 R_1^{G2} + Y_1 - C_1^G(R_1^{G2}, Y_1) \leq W_0 R_1^{F2} + Y_1 - C_1^G(R_1^{G2}, Y_1) \\
&\qquad\qquad\qquad\qquad\vdots \\
&W_0 R_1^{GN} + Y_1 - C_1^G(R_1^{GN}, Y_1) \leq W_0 R_1^{FN} + Y_1 - C_1^G(R_1^{G1}, Y_1),
\end{aligned} \tag{S7}$$

where we have replaced $C_1^F(R_1^{Fi}, Y_1)$ with $C_1^G\left(R_1^{Gi}, Y_1\right)$ on the right hand sides, as these are equal by construction (i.e. by eq.(S3)). Inequalities (S7) hold because of the FSD inequalities in (S6), and because the consumption is the same on both sides of each inequality. Equations (S7) imply that given a labor income realization of $Y_1$, $W_0 R_1^F + Y_1 - C_1^F(R_1^F, Y_1)$ FSD dominates $W_0 R_1^G + Y_1 - C_1^G(R_1^G, Y_1)$. As such dominance holds for any realization of $\tilde{Y}_1$, we have an FSD dominance of $W_0 R_1^F + \tilde{Y}_1 - C_1^F\left(\tilde{R}_1^F, \tilde{Y}_1\right)$ over $W_0 R_1^G + \tilde{Y}_1 - C_1^G\left(\tilde{R}_1^G, \tilde{Y}_1\right)$.[1]

Moving to the second period, we have:

$$\begin{aligned}
\widetilde{W}_2^G &= \widetilde{W}_1^G \tilde{R}_2^G + \tilde{Y}_2 - \tilde{C}_2^G(\tilde{R}_2^G, \tilde{Y}_2) \\
\widetilde{W}_2^F &= \widetilde{W}_1^F \tilde{R}_2^F + \tilde{Y}_2 - \tilde{C}_2^F(\tilde{R}_2^F, \tilde{Y}_2).
\end{aligned} \tag{S8}$$

As we have established above that $\widetilde{W}_1^F$ dominates $\widetilde{W}_1^G$ by FSD, and $\tilde{R}_2^F$ dominates $\tilde{R}_2^G$ by FSD and as the period-2 returns are independent of the period-1 wealth, we have that $\widetilde{W}_1^F \tilde{R}_2^F$ FSD dominates $\widetilde{W}_1^G \tilde{R}_2^G$.[2] For any labor income realization $Y_2$ we define the portfolio $F$ consumption again by (S3), and obtain the inequality in (S5) with the time index 2 replacing 1. Thus, it follows that $\widetilde{W}_2^F$ FSD dominates $\widetilde{W}_2^G$. Iterating over all time periods, we finally obtain that $\widetilde{W}_T^F$ FSD dominates $\widetilde{W}_T^G$. This proves that by investing in strategy $F$ the investor can achieve in every period the exact same stochastic consumption distribution as obtained when investing in

[1] Note that this does not imply that strategy F yields the same consumption as strategy G in every state of nature, only that the distribution of consumption is the same under both strategies. A similar approach is employed by Levy and Levy (2021) in the context of the question of stocks versus bonds in the long-run.

[2] If $\tilde{x}_1$ dominates $\tilde{y}_1$ by FSD, $\tilde{x}_2$ dominates $\tilde{y}_2$ by FSD, $\tilde{x}_1$ and $\tilde{x}_2$ are independent, and $\tilde{y}_1$ and $\tilde{y}_2$ are independent, then $\tilde{x}_1 \cdot \tilde{x}_2$ dominates $\tilde{y}_1 \cdot \tilde{y}_2$ by FSD (Levy 1973).

strategy $G$ and consuming optimally, but with a dominating terminal wealth distribution. It follows that all investors with increasing utility functions select strategy $F$, i.e. their optimal equity portfolio will be the 1-period mean-variance tangency portfolio.

### Section S2– Results with Randomly Selected Stocks

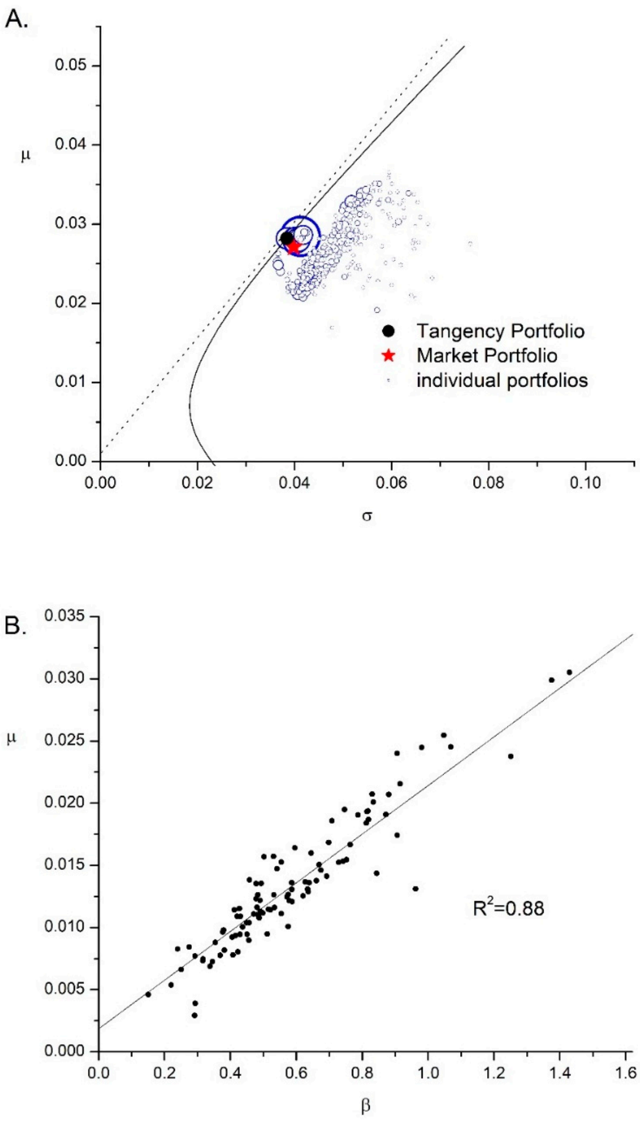


**Figure S1:** Equilibrium in a market with heterogeneous preferences and investment horizons, and the empirical monthly return distributions. Firms are selected randomly from the set of all firms in our sample (rather than taking the largest firms). Monthly returns are assumed to be independent across time, as opposed to the analysis in the main text, which employs the empirical autocorrelations. Panel A: Investor's equity portfolios, shown in the 1-month mean-variance plane. The solid circle represents the mean-variance tangency portfolio. The hollow circles represent investors' portfolios. The area of each hollow circle represents the number of investors holding the portfolio. The star is the market portfolio, which is the aggregate portfolio of all individuals. Panel B: The relationship between beta and expected returns.

**Figure S3– The Effect of Serial Correlations**

To analyze the effect of serial correlations on the results, we repeat the analysis when serial independence is assumed - each month is drawn randomly from the empirical sample (rather than drawing 12 consecutive months, as in the main text). The results, which are very similar to those with serial correlations, are given in Figure S1. We conclude that serial correlations do not have a large effect on the results, and most of the deviations from the theoretical predictions of Theorem 1 are due to the non-normality of the 1-month return distributions. This is not surprising given that the empirical serial correlations are close to zero.

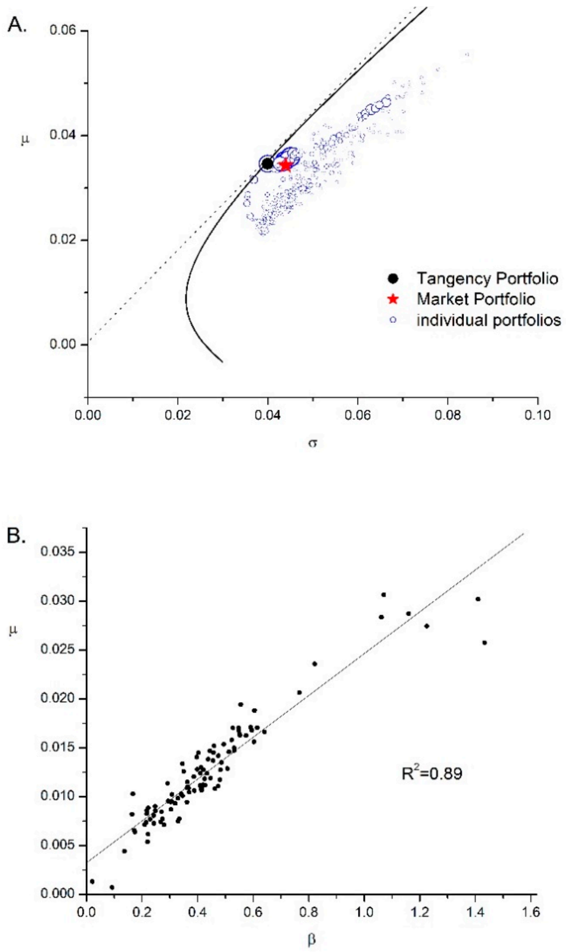


**Figure S2:** Equilibrium in a market with heterogeneous preferences and investment horizons, and the empirical monthly return distributions. Monthly returns are assumed to be independent across time, as opposed to the analysis in the main text, which employs the empirical autocorrelations. Panel A: Investor's equity portfolios, shown in the 1-month mean-variance plane. The solid circle represents the mean-variance tangency portfolio. The hollow circles represent investors' portfolios. The area of each circle represents the number of investors holding the portfolio. The star is the market portfolio, which is the aggregate portfolio of all individuals. Panel B: The relationship between beta and expected returns.

## Section S4– Return Parameters Adjusted to Make Tangency Portfolio Weights Consistent with Empirical Market Capitalizations

The tangency portfolio derived based on the sample parameters usually involves many short positions. Levy and Roll (LR, 2010) show that small adjustments to the sample parameters, well within their estimation error bounds, are sufficient to make the tangency portfolio weights equal to the weights of the market proxy based on firms' empirical market capitalizations. Here we repeat the analysis reported in the main text, but we employ the LR adjusted parameters instead of the sample parameters. Following LR we define the distance between the sample parameters and the adjusted parameters as:

$$D\left((\mu,\sigma),(\mu,\sigma)^{sam}\right) \equiv \sqrt{\alpha \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mu_i - \mu_i^{sam}}{\sigma_i^{sam}}\right)^2 + (1-\alpha)\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i - \sigma_i^{sam}}{\sigma_i^{sam}}\right)^2}\ , \qquad \text{(S9)}$$

where $\mu_i^{sam}$ and $\sigma_i^{sam}$ are the sample mean and standard deviation of stock $i$, $\mu_i$ and $\sigma_i$ are their adjusted counterparts, and $\alpha$ is a constant determining the relative importance of deviations from the means and deviations from the standard deviations (the correlations are taken as their sample values, i.e. they are not adjusted). The optimal adjusted parameters, $\mu_i^*$ and $\sigma_i^*$, are the parameters that ensure that the tangency portfolio weights coincide with the market proxy weights, while minimizing the distance *D*.

We take the market proxy as value-weighted portfolio of the 100 stocks in our sample, with market values at March 31, 2022. As in LR, we find that small adjustments to the parameters suffice to make the tangency portfolio coincide with the market proxy. Table D1 provides the sample and adjusted parameters of the 20 largest stocks (the results for the other 80 stocks are very similar). Our goal is to examine the robustness of Theorem 1 to non-normal return distributions. Thus, we apply the LR parameter adjustments, but we maintain the discrete form of the empirical return distribution. Namely, we first multiply all sample returns by $\frac{\sigma_i^*}{\sigma_i^{sam}}$ to obtain a discrete return distribution with the desired standard deviation of $\sigma_i^*$. Then, a constant is added to all returns to ensure that the mean of the distribution is the desired $\mu_i^*$.

Figure S2 shows the 1-month mean-variance frontier, tangency portfolio, individual portfolios and the aggregate portfolio, when the adjusted monthly parameters are employed. As in the case that the sample parameters are used, the market portfolio obtained by aggregating the holdings of all investors in the market is close to the tangency portfolio, and the relationship between betas and expected returns is very close to linear ($R^2$=0.97). The fit is even better than in the case of the unadjusted parameters. The reason could be that the adjusted parameters tend to be "shrunk" towards average values (see Table S1, and the discussion in Levy and Roll 2010). Thus, the weights in individual's portfolios tend to be less extreme, and to deviate less from the tangency portfolio (deviations that are mainly due to the non-normality of the return distributions). Hence, the cancelation of these deviations is more effective, and aggregation across investors leads to a market portfolio that is very close to the 1-month tangency portfolio.

## Table S1

### The Sample Parameters and the Adjusted Parameters Ensuring that the Given Market Proxy Coincides with Tangency Portfolio

For the sake of brevity, only the parameters of the first 20 stocks are reported. The sample parameters are given in the second and fourth columns. The adjusted expected returns and standard deviations which ensure that the tangency portfolio coincides with the market proxy and minimize the distance *D* in eq.(S9) are given in columns (3) and (5). The t-values for the expected returns are given in column (6). The difference between the sample mean is significant only for stock 9 (it is also significant for 3 other stocks of the 80 stocks not reported here). Column (7) reports the ratio between the adjusted variances and the sample variances. The 95% confidence interval for this ratio is [0.844-1.209].[3] All of the ratios in the table, as well as the ratios for all other 80 stocks not shown here, fall within this interval.

| (1) Stock # | (2) $\mu_i^{sam}$ | (3) $\mu_i^*$ | (4) $\sigma_i^{sam}$ | (5) $\sigma_i^*$ | (6) t-value for $\mu_i^*$ | (7) $(\sigma_i^*)^2/(\sigma_i^{sam})^2$ |
|---|---|---|---|---|---|---|
| 1 | 0.011 | 0.015 | 0.074 | 0.072 | 0.532 | 0.960 |
| 2 | 0.014 | 0.015 | 0.066 | 0.065 | 0.044 | 0.980 |
| 3 | 0.011 | 0.016 | 0.066 | 0.064 | 0.794 | 0.934 |
| 4 | 0.011 | 0.013 | 0.063 | 0.062 | 0.221 | 0.985 |
| 5 | 0.007 | 0.008 | 0.047 | 0.047 | 0.108 | 0.998 |
| 6 | 0.007 | 0.010 | 0.062 | 0.062 | 0.473 | 0.970 |
| 7 | 0.011 | 0.010 | 0.052 | 0.053 | -0.456 | 1.029 |
| 8 | 0.011 | 0.013 | 0.066 | 0.065 | 0.384 | 0.974 |
| 9 | 0.001 | 0.014 | 0.086 | 0.081 | 2.073 | 0.881 |
| 10 | 0.005 | 0.013 | 0.066 | 0.062 | 1.557 | 0.894 |
| 11 | 0.008 | 0.009 | 0.044 | 0.045 | -0.154 | 1.012 |
| 12 | 0.013 | 0.009 | 0.066 | 0.067 | -1.088 | 1.048 |
| 13 | 0.012 | 0.014 | 0.088 | 0.088 | 0.097 | 0.991 |
| 14 | 0.009 | 0.010 | 0.070 | 0.070 | -0.165 | 1.009 |
| 15 | 0.011 | 0.012 | 0.065 | 0.065 | 0.076 | 0.992 |
| 16 | 0.031 | 0.023 | 0.096 | 0.104 | -1.330 | 1.167 |
| 17 | 0.012 | 0.020 | 0.097 | 0.093 | 0.992 | 0.923 |
| 18 | 0.012 | 0.017 | 0.080 | 0.077 | 0.750 | 0.940 |
| 19 | 0.011 | 0.011 | 0.066 | 0.066 | -0.127 | 1.007 |
| 20 | 0.015 | 0.014 | 0.072 | 0.073 | -0.333 | 1.023 |

[3] The ratio $\frac{(n-1)s^2}{\sigma^2}$ is distributed according to the $\chi^2_{n-1}$distribution, where $\sigma^2$is the population variance, $s^2$is the sample variance (or $(\sigma^{sam})^2$in the notation used in this paper), and *n* is the number of observations. We have 240 monthly return observations, hence *n*=240. As we are looking for the 95% confidence interval for $s^2/\sigma^2$, we need to find the critical values $c_1$ and $c_2$ for which $P(\chi^2_{239} > c_1) = 0.025$, and $P(\chi^2_{239} < c_2) = 0.025$. For large $n$, $\sqrt{2\chi^2_n} - \sqrt{2n-1}$ can be approximated by the standard normal distribution. Thus, the critical values $c_1$ and $c_2$ satisfy $\sqrt{2c_1} - \sqrt{2 \cdot 239 - 1} = 1.96$ and $\sqrt{2c_2} - \sqrt{2 \cdot 239 - 1} = -1.96$, which yield: $c_1 = 283.2$ and $c_2 = 197.6$. Thus, the 95% confidence interval for $s^2/\sigma^2$ is given by $197.6 < 239 \cdot s^2/\sigma^2 < 283.2$ or: $0.827 < s^2/\sigma^2 < 1.185$. Alternatively, this range can be also stated as $0.844 < \sigma^2/s^2 < 1.209$.

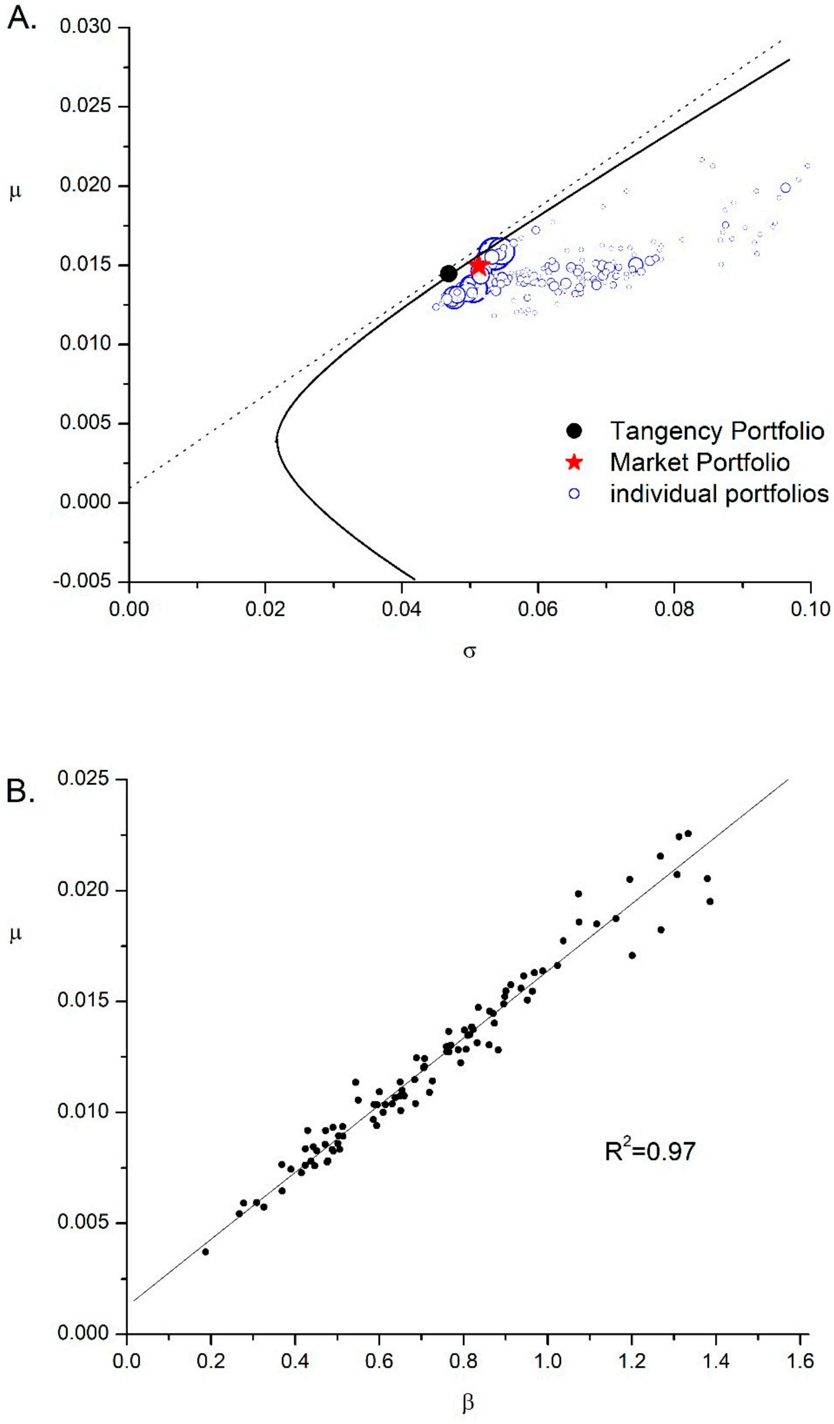


**Figure S3:** Equilibrium in a market with heterogeneous preferences and investment horizons, when the empirical monthly return distributions are adjusted according to Levy and Roll (2010). Panel A: Investor's equity portfolios, shown in the 1-month mean-variance plane. The solid circle represents the mean-variance tangency portfolio. The hollow circles represent investors' portfolios. The area of each hollow circle represents the number of investors holding the portfolio. The star is the market portfolio, which is the aggregate portfolio of all individuals. Panel B: The relationship between beta and expected returns.