*Econometrica Supplementary Material*

# SUPPLEMENT TO "A NEWS-UTILITY THEORY FOR INATTENTION AND DELEGATION IN PORTFOLIO CHOICE"

(*Econometrica*, Vol. 86, No. 2, March 2018, 491–522)

MICHAELA PAGEL
Columbia GSB, NBER, and CEPR

## APPENDIX A: NEWS UTILITY IN STATIC PORTFOLIO THEORY

### A.1. *Proof of Lemma 1*

START WITH $u(c) = \log(c)$ AND CONSUMPTION IS $C = W(R^f + \alpha(R - R^f)) \sim F_C$. Thus, the agent's maximization problem is

$$E\left[\log(C) + \eta \int_{-\infty}^{C} \big(\log(C) - \log(c)\big)\, dF_C(c) + \eta\lambda \int_{C}^{\infty} \big(\log(C) - \log(c)\big)\, dF_C(c)\right],$$

which can be rewritten as

$$E\left[\log(C) + \eta(\lambda - 1) \int_{C}^{\infty} \big(\log(C) - \log(c)\big)\, dF_C(c)\right]$$

because expected good news and bad news partly cancel. I approximate the log portfolio return $\log(R^f + \alpha(R - R^f))$ by $r^f + \alpha(r - r^f) + \alpha(1 - \alpha)\frac{\sigma^2}{2}$ (as in Campbell and Viceira (2002) among many others). Thus, $\log C = \log W + \log(R^f + \alpha(R - R^f)) \approx \log W + r^f + \alpha(r - r^f) + \alpha(1 - \alpha)\frac{\sigma^2}{2} \sim N(\log W + r^f + \alpha(\mu - \frac{\sigma^2}{2} - r^f) + \alpha(1 - \alpha)\frac{\sigma^2}{2}, \alpha^2\sigma^2)$ as $r \sim N(\mu - \frac{\sigma^2}{2}, \sigma^2)$. Let me denote the cumulative distribution function of $r$ by $F_r$. In turn, I can rewrite $E[\eta(\lambda - 1) \int_C^\infty (\log(C) - \log(c))\, dF_C(c)]$ as

$$E\left[\eta(\lambda - 1) \int_{r}^{\infty} (\alpha r - \alpha\tilde{r})\, dF_r(\tilde{r})\right]$$

(as all constant terms will cancel). I then denote a standard normal variable by $s \sim F_s = N(0, 1)$ and thus

$$E\left[\eta(\lambda - 1) \int_{r}^{\infty} (\alpha r - \alpha\tilde{r})\, dF_r(\tilde{r})\right] = \alpha\sigma E\left[\eta(\lambda - 1) \int_{s}^{\infty} (s - \tilde{s})\, dF_s(\tilde{s})\right].$$

In turn, I can rewrite the maximization problem as

$$r^f + \alpha\left(\mu - \frac{\sigma^2}{2} - r^f\right) + \alpha(1 - \alpha)\frac{\sigma^2}{2} + \alpha\sigma E\left[\eta(\lambda - 1) \int_{s}^{\infty} (s - \tilde{s})\, dF_s(\tilde{s})\right]$$

with the first-order condition given by

$$\mu - r^f - \alpha\sigma^2 + \sigma E\left[\eta(\lambda - 1) \int_{s}^{\infty} (s - \tilde{s})\, dF_s(\tilde{s})\right] = 0,$$

Michaela Pagel: mpagel@columbia.edu

 https://doi.org/10.3982/ECTA14417

which results in the optimal portfolio share stated in Equation (2) if $0 \leq \alpha^*$. If the expression (2) would be negative, then $\alpha^* = 0$. This can be easily inferred from the maximization problem and the fact that expected news utility is negative in the presence of uncertainty, that is, whenever $\alpha \neq 0$, and expected consumption utility would be lower if $\alpha < 0$ as $\mu > r^f$. Thus, the agent would always prefer $\alpha = 0$ over $\alpha < 0$. The second-order sufficient condition is given by $-\sigma^2 < 0$, which implies that $\alpha^*$ is a global maximum. The news-utility agent's optimal portfolio share is

$$\alpha = \frac{\mu - r^f + E\left[\eta(\lambda - 1)\int_r^\infty (r - \tilde{r})\, dF_r(\tilde{r})\right]}{\sigma^2}.$$

Now, let me redefine $\mu \triangleq h\mu$, $\sigma \triangleq \sqrt{h}\sigma$, and $r^f \triangleq hr^f$. The optimal portfolio share is given by

$$\alpha = \frac{\mu - r^f + \frac{\sqrt{h}}{h}\sigma E\left[\eta(\lambda - 1)\int_s^\infty (s - \tilde{s})\, dF_s(\tilde{s})\right]}{\sigma^2}.$$

Samuelson's colleague and time diversification: As can be easily seen, as $\lim \frac{\sqrt{h}}{h} \to \infty$ as $h \to 0$, there exists some $\underline{h}$ such that $\mu - r^f < -\frac{\sqrt{h}}{h}\sigma E[\eta(\lambda - 1)\int_s^\infty (s - \tilde{s})\, dF_s(\tilde{s})]$ and thus $\alpha = 0$. As $\alpha > 0$ only if

$$\frac{\mu - r^f}{\sigma} > -E\left[\eta(\lambda - 1)\underbrace{\int_s^\infty (s - \tilde{s})\, dF_s(\tilde{s})}_{<0}\right] > 0.$$

In contrast, $\alpha^s > 0$ whenever $\mu > r^f$. On the other hand, $\alpha > 0$ for some $h$ if $h \to \infty$, then $\lim \frac{\sqrt{h}}{h} \to 0$ and $\alpha \to \alpha^s$. Furthermore, it can be easily seen that $\frac{\partial \alpha}{\partial h} > 0$.

## APPENDIX B: DERIVATION OF THE DYNAMIC LIFE-CYCLE PORTFOLIO-CHOICE MODEL

### B.1. *The Monotone-Personal Equilibrium*

I follow a guess and verify solution procedure. The agent adjusts his portfolio share and consumes a fraction $\rho_t$ out of his wealth (if he looks up his portfolio) and a fraction $\rho_t^{\text{in}}$ out of his wealth (if he stays inattentive). Suppose he last looked up his portfolio in $t - i$, that is, he knows $W_{t-i}$. And suppose he will look up his portfolio in period $t + j_1$ but is inattentive in period $t$. Then, his inattentive consumption in periods $t - i + 1$ to $t + j_1 - 1$ is given by (for $k = 1, \dots, i + j_1 - 1$)

$$C_{t-i+k}^{\text{in}} = (W_{t-i} - C_{t-i})\left(R^d\right)^k \rho_{t-i+k}^{\text{in}}$$

and his consumption when he looks up his portfolio in period $t + j_1$ is given by

$$\begin{aligned} C_{t+j_1} &= W_{t+j_1}\rho_{t+j_1} \\ &= \left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i+j_1-1} \frac{C_{t-i+k}^{\text{in}}}{\left(R^d\right)^k}\right)\left(\left(R^f\right)^{i+j_1} + \alpha_{t-i}\left(\prod_{j=1}^{i+j_1} R_{t-i+j} - \left(R^f\right)^{i+j_1}\right)\right)\rho_{t+j_1} \end{aligned}$$

$$= W_{t-i}(1-\rho_{t-i})\left(1-\sum_{k=1}^{i+j_1-1}\rho_{t-i+k}^{\text{in}}\right)\left(\left(R^f\right)^{i+j_1}+\alpha_{t-i}\left(\prod_{j=1}^{i+j_1}R_{t-i+j}-\left(R^f\right)^{i+j_1}\right)\right)\rho_{t+j_1}.$$

Now, suppose the agent looks up his portfolio in period $t$ and then chooses $C_t$ and $\alpha_t$ knowing that he will look up his portfolio in period $t+j_1$ next time. I first explain the optimal choice of $C_t$. First, the agent considers marginal consumption and contemporaneous marginal news utility given by

$$u'(C_t)\big(1+\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big).$$

To understand these terms, note that the agent takes his beliefs as given, his admissible consumption function $C_t$ increases in $r_t+\cdots+r_{t-i+1}$ such that $F_{C_t}^{t-i}(C_t)=F_r(r_t)\cdots F_r(r_{t-i+1})$, and

$$\begin{aligned}&\frac{\partial\left(\eta\int_{-\infty}^{C_t}\big(u(C_t)-u(x)\big)\,dF_{C_t}^{t-i}(x)+\eta\lambda\int_{C_t}^{\infty}\big(u(C_t)-u(x)\big)\,dF_{C_t}^{t-i}(x)\right)}{\partial C_t}\\&\quad=u'(C_t)\big(\eta F_{C_t}^{t-i}(C_t)+\eta\lambda\big(1-F_{C_t}^{t-i}(C_t)\big)\big).\end{aligned}$$

Second, the agent takes into account that he will experience prospective news utility over all consumption in periods $t+1,\dots,T$. Inattentive consumption in periods $t+1$ to $t+j_1-1$ is as above given by (for $k=1,\dots,j_1-1$)

$$C_{t+k}^{\text{in}}=(W_t-C_t)\left(R^d\right)^k\rho_{t+k}^{\text{in}}$$

and thus proportional to $W_t-C_t$. Attentive consumption in period $t+j_1$ is given by

$$\begin{aligned}C_{t+j_1}&=W_{t+j_1}\rho_{t+j_1}\\&=\left(W_t-C_t-\sum_{k=1}^{j_1-1}\frac{C_{t+k}^{\text{in}}}{\left(R^d\right)^k}\right)\left(\left(R^f\right)^{j_1}+\alpha_t\left(\prod_{j=1}^{j_1}R_{t+j}-\left(R^f\right)^{j_1}\right)\right)\rho_{t+j_1}\\&=(W_t-C_t)\left(1-\sum_{k=1}^{j_1-1}\rho_{t+k}^{\text{in}}\right)\left(\left(R^f\right)^{j_1}+\alpha_t\left(\prod_{j=1}^{j_1}R_{t+j}-\left(R^f\right)^{j_1}\right)\right)\rho_{t+j_1}\end{aligned}$$

and thus proportional to $W_t-C_t$. In turn, prospective marginal news utility is

$$\begin{aligned}&\frac{\partial\left(\gamma\sum_{k=1}^{j_1-1}\beta^k\boldsymbol{n}\big(F^{t,t-i}_{(W_t-C_t)(R^d)^k\rho_{t+k}^{\text{in}}}\big)+\gamma\sum_{j=j_1}^{T-t}\beta^j\boldsymbol{n}\big(F^{t,t-i}_{C_{t+j}}\big)\right)}{\partial C_t}\\&\quad=\frac{\partial\log(W_t-C_t)}{\partial C_t}\gamma\sum_{j=1}^{T-t}\beta^j\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big).\end{aligned}$$

To understand this derivation, note that the agent takes his beliefs as given, future consumption increases in today's return realization, and the only terms that realize and thus do not cancel out of the news-utility terms are $W_t - C_t$ such that $F^{t-i}_{W_t-C_t}(W_t - C_t) = F_r(r_t)\cdots F_r(r_{t-i+1})$, whether future consumption is inattentive or attentive. As an example, consider the derivative of prospective news-utility in period $t+1$

$$\frac{\partial \gamma\beta \boldsymbol{n}\big(F^{t,t-i}_{(W_t-C_t)R^d\rho^{\text{in}}_{t+1}}\big)}{\partial C_t} = \frac{\partial \gamma\beta \int_{-\infty}^{\infty} \mu\big(\log\big((W_t - C_t)R^d\rho^{\text{in}}_{t+1}\big) - \log(x)\big)\, dF^{t-i}_{(W_t-C_t)R^d\rho^{\text{in}}_{t+1}}(x)}{\partial C_t}$$
$$= \frac{\partial \log(W_t - C_t)}{\partial C_t}\gamma\beta\big(\eta F^{t-i}_{W_t-C_t}(W_t - C_t) + \eta\lambda\big(1 - F^{t-i}_{W_t-C_t}(W_t - C_t)\big)\big).$$

Third, thanks to log utility, the agent's continuation utility is not affected by expected news utility as $\log(W_t - C_t)$ cancels out of these terms (it is the same in both actual consumption and beliefs). However, expected consumption utility matters. Consumption utility beyond period $t + j_1$ can be iterated back to $t + j_1$ wealth which can be iterated back to $W_t$ as $\log(W_t - C_t - \sum_{k=1}^{j_1-1} \frac{C^{\text{in}}_{t+k}}{(R^d)^k}) = \log((W_t - C_t)(1 - \sum_{k=1}^{j_1-1} \rho^{\text{in}}_{t+k}))$ such that

$$\frac{\partial \sum_{j=1}^{T-t} \beta^j E_t\big[\log(C_{t+j})\big]}{\partial C_t} = \frac{\partial \log(W_t - C_t)}{\partial C_t}\sum_{j=1}^{T-t}\beta^j = -u'(W_t - C_t)\sum_{j=1}^{T-t}\beta^j.$$

Putting the three pieces together, optimal consumption (if the agent looks up his portfolio in period $t$) is determined by the following first-order condition:

$$\begin{aligned} & u'(C_t)\big(1 + \eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big) \\ & \quad - \gamma u'(W_t - C_t)\sum_{j=1}^{T-t}\beta^j\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big) \\ & \quad - u'(W_t - C_t)\sum_{j=1}^{T-t}\beta^j \\ & = 0. \end{aligned}$$

In turn, the solution guess can be verified:

$$\frac{C_t}{W_t} = \rho_t = \frac{1}{1 + \sum_{\tau=1}^{T-t}\beta^\tau \frac{1 + \gamma\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1 + \eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}.$$

The optimal portfolio share depends on prospective news utility in period $t$ for all the consumption levels in periods $t + j_1$ to $T$ that are all proportional to $W_{t+j_1}$ (i.e., determined by $\alpha_t$). Moreover, the agent's consumption and news utility in period $t + j_1$ matters. However, consumption in inattentive periods $t$ to $t + j_1 - 1$ depends only on $W_t - C_t$, which is not affected by $\alpha_t$. Moreover, $\alpha_t$ cancels in all expected news-utility terms after period $t + j_1$. Thus, the relevant terms in the maximization problem for the portfolio share are given by

$$\gamma\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}\boldsymbol{n}\big(F_{C_{t+j_1+\tau}}^{t,t-i}\big)+\beta^{j_1}E_t\Bigg[n\big(C_{t+j_1},F_{C_{t+j_1}}^{t+j_1,t}\big)+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^{\tau}\boldsymbol{n}\big(F_{C_{t+j_1+\tau}}^{t+j_1,t}\big)\Bigg]$$
$$+\beta^{j_1}E_t\Bigg[\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}\log(C_{t+j_1+\tau})\Bigg].$$

The derivative of the first term is

$$\frac{\partial\gamma\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}\boldsymbol{n}\big(F_{C_{t+j_1+\tau}}^{t,t-i}\big)}{\partial\alpha_t}$$
$$=j_1\big(\mu-r^f-\alpha_t\sigma^2\big)\gamma\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big).$$

To illustrate where the derivative comes from,

$$\log(C_{t+j_1})=\log(W_{t+j_1}\rho_{t+j_1})$$
$$=\log\left((W_t-C_t)\left(1-\sum_{j=1}^{j_1-1}\rho_{t+j}^{\text{in}}\right)\left(\big(R^f\big)^{j_1}+\alpha_t\left(\prod_{j=1}^{j_1}R_{t+j}-\big(R^f\big)^{j_1}\right)\right)\rho_{t+j_1}\right),$$

thus the only term determined by $\alpha_t$ is $\log((R^f)^{j_1}+\alpha_t(\prod_{j=1}^{j_1}R_{t+j}-(R^f)^{j_1}))\approx j_1r^f+\alpha_t(\sum_{j=1}^{j_1}r_{t+j}-j_1r^f)+j_1\alpha_t(1-\alpha_t)\frac{\sigma^2}{2}$ and the agent takes his beliefs as given with future consumption being increasing in this period's return as in the derivation of the consumption share above).[31] Moreover,

$$C_{t+j_1+1}^{\text{in}}=(W_t-C_t)\left(1-\sum_{j=1}^{j_1-1}\rho_{t+j}^{\text{in}}\right)\left(\big(R^f\big)^{j_1}+\alpha_t\left(\prod_{j=1}^{j_1}R_{t+j}-\big(R^f\big)^{j_1}\right)\right)(1-\rho_{t+j_1})R^d\rho_{t+j_1+1}^{\text{in}};$$

thus, the only term determined by $\alpha_t$ are $\log((R^f)^{j_1}+\alpha_t(\prod_{j=1}^{j_1}R_{t+j}-(R^f)^{j_1}))$ and beliefs are taken as given. Analogously to above, marginal prospective news utility is composed of the derivative $\frac{\partial\log(C_{t+j_1+\tau})}{\partial\alpha_t}$ and the news-utility terms (prior beliefs are taken as given and thus drop out) determined by the weighted sum of $\eta$ and $\eta\lambda$. Thus, in the derivative, the term $j_1(\mu-r^f-\alpha_t\sigma^2)$ is left and the integrals are determined by $F_{C_{t+j_1+1}^{\text{in}}}^{t,t-1}=F_r(r_t)\cdots F_r(r_{t-i+1})$. Thus, only the sum $\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}$ remains in the derivative.

[31]Note that, if $R_{t+j}$ log-normally distributed, then $\prod_{j=1}^{j_1}R_{t+j}$ is distributed log-normally with $\mu_{\prod}=\sum_{j=1}^{j1}\mu_{R_{t+j}}=j_1(\mu-\frac{\sigma^2}{2})$ and $\sigma_{\prod}^2=\sum_{j=1}^{j_1}\sigma_{R_{t+j}}^2=j_1\sigma^2$.

In turn, the derivative of the continuation value is

$$\frac{\beta^{j_1}E_t\left[n\left(C_{t+j_1},F^{t+j_1,t}_{C_{t+j_1}}\right)+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^{\tau}\boldsymbol{n}\left(F^{t+j_1,t}_{C_{t+j_1+\tau}}\right)\right]+\beta^{j_1}E_t\left[\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}\log(C_{t+j_1+\tau})\right]}{\partial\alpha_t}$$

$$=\beta^{j_1}\left(1+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^{\tau}\right)\sqrt{j_1}\sigma E\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]$$

$$+\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}j_1\left(\mu-r^f-\alpha_t\sigma^2\right),$$

as all terms in expected news utility cancel except $\alpha_t\sum_{j=1}^{j_1}r_{t+j}=\alpha_t\sqrt{j_1}\sigma s$ and all terms in expected consumption utility drop out in the derivative except $j_1(\mu-r^f-\alpha_t\sigma^2)$. Altogether, the optimal portfolio share is given by

$$\alpha_t=\frac{\mu-r^f+\dfrac{\dfrac{\beta^{j_1}\left(1+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^{\tau}\right)}{\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^{\tau}}\dfrac{\sqrt{j_1}}{j_1}\sigma E\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]}{1+\gamma\left(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\left(1-F_r(r_t)\cdots F_r(r_{t-i+1})\right)\right)}}{\sigma^2}.$$

This is an optimum only if $\alpha_t\geq 0$; if $\alpha_t<0$, the agent would choose $\alpha_t=0$ instead. This can be easily inferred from the maximization problem and the fact that expected news utility is negative in the presence of uncertainty, that is, whenever $\alpha_t\neq 0$, and expected consumption utility would be lower if $\alpha_t<0$ as $\mu>r^f$. Thus, the agent would always prefer $\alpha_t=0$ over $\alpha_t<0$.

If the agent is inattentive in period $t$, his consumption is determined by the following first-order condition:

$$u'\left(C_t^{\text{in}}\right)-\sum_{\tau=0}^{T-t-j_1}\beta^{j_1+\tau}u'\left(W_{t-i}-C_{t-i}-\sum_{k=1}^{i+j_1-1}\frac{C^{\text{in}}_{t-i+k}}{\left(R^d\right)^k}\right)\frac{1}{\left(R^d\right)^i}=0.$$

The term concerning consumption in period $t$ is self-explanatory. The terms concerning consumption from periods $t-i$ to $t+j_1-1$ drop out as they are determined by the solution guess $C_t^{\text{in}}=(W_{t-i}-C_{t-i})(R^d)^i\rho_t^{\text{in}}$. The terms concerning consumption from period $t+j_1$ on are all proportional to $\log(W_{t+j_1})$, which equals $\log(W_{t-i}-C_{t-i}-\sum_{k=1}^{i+j_1-1}\frac{C^{\text{in}}_{t-i+k}}{(R^d)^k})$ plus the log returns from period $t-i+1$ to period $t+j_1$, which, however, drop out by taking the derivative with respect to $C_t^{\text{in}}$ thanks to log utility. Accordingly,

$$\frac{1}{C_t^{\text{in}}}-\sum_{\tau=0}^{T-t-j_1}\beta^{j_1+\tau}\frac{1}{W_{t-i}-C_{t-i}-\sum_{k=1}^{i+j_1-1}\dfrac{C^{\text{in}}_{t-i+k}}{\left(R^d\right)^k}}\frac{1}{\left(R^d\right)^i}=0$$

$$\Rightarrow \quad \frac{1}{\rho_t^{\text{in}}} = \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau} \frac{1}{1 - \sum_{k=1}^{i+j_1-1} \rho_{t-i+k}^{\text{in}}},$$

$$\rho_t^{\text{in}} = \frac{1}{1 + \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau}} \left(1 - \sum_{k=1}^{i-1} \rho_{t-i+k}^{\text{in}} - \sum_{k=1}^{j_1-1} \rho_{t+k}^{\text{in}}\right).$$

The set of final $\rho_t^{\text{in}}$ that are determined by this recursion will only depend on the value of $\beta$. For $\beta \approx 1$, for instance, the optimal inattentive consumption share $\rho_t^{\text{in}}$ is the same for each period as can be seen easily:

$$\rho_t^{\text{in}} = \frac{1}{1 + \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau}} \left(1 - \sum_{k=1}^{i-1} \rho_{t-i+k}^{\text{in}} - \sum_{k=1}^{j_1-1} \rho_{t+k}^{\text{in}}\right)$$

$$= \frac{1}{1 + T - t - j_1 + 1} \left(1 - \sum_{k=1}^{i-1} \rho_{t-i+k}^{\text{in}} - \sum_{k=1}^{j_1-1} \rho_{t+k}^{\text{in}}\right) = \rho_{t-1}^{\text{in}}$$

so that

$$\rho_t^{\text{in}} = \frac{1}{1 + T - t - j_1 + 1}\left(1 - (i-1)\rho_t^{\text{in}} - (j_1 - 1)\rho_t^{\text{in}}\right) \quad \Rightarrow \quad \rho_t^{\text{in}} = \frac{1}{T - t + i}.$$

And the solution guess ($C_t^{\text{in}} = (W_{t-i} - C_{t-i})(R^d)^i \rho_t^{\text{in}}$) can be verified. The agent does not deviate and overconsume in inattentive periods $t-i+1$ to $t-1$ so long as the bad news from deviating outweighs the good news from doing so, that is, $u'((R^d)^k \rho_{t-i+k}^{\text{in}})(1+\eta) < \beta^j u'((R^d)^{k+j} \rho_{t-i+k+j}^{\text{in}})(1+\gamma\eta\lambda) \Rightarrow u'(\rho_{t-i+k}^{\text{in}})(1+\eta) < \beta^j (R^d)^{-j} u'(\rho_{t-i+k+j}^{\text{in}})(1+\gamma\eta\lambda)$ for $k = 1, \ldots, i-2$ and $j = 1, \ldots, i-k-1$. In the derivation, I assume that this condition holds so that, in inattentive periods, the agent does not deviate from his consumption path and does not overconsume time inconsistently. For $\beta R^d \approx 1$ and given that$\rho_{t-i+k}^{\text{in}} \approx \rho_{t-i+k+j}^{\text{in}}$, this condition is roughly equivalent to $\lambda > \frac{1}{\gamma}$.

### B.2. *The Monotone-Precommitted Equilibrium*

Suppose the agent has the ability to pick an optimal history-dependent consumption path for each possible future contingency in period zero when he does not experience any news utility. Thus, in period zero, the agent chooses optimal consumption in period $t$ in each possible contingency jointly with his beliefs, which of course coincide with the agent's optimal state-contingent plan. Additional details about the derivation can be found in

Pagel (2017). The optimal precommitted portfolio share is

$$\alpha_t^c = \frac{\mu - r^f + \dfrac{\dfrac{\beta^{j_1}\left(1+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^\tau\right)}{\beta^{j_1}\sum_{\tau=0}^{T-t-j_1}\beta^\tau}\dfrac{\sqrt{j_1}}{j_1}\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]}{1+\gamma\eta(\lambda-1)\big(1-2F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}{\sigma^2}.$$

Moreover, the optimal precommitted attentive and inattentive consumption shares are

$$\rho_t^c = \frac{1}{1+\sum_{\tau=1}^{T-t}\beta^\tau \dfrac{1+\gamma\eta(\lambda-1)\big(1-2F_r(r_t)\cdots F_r(r_{t-i+1})\big)}{1+\eta(\lambda-1)\big(1-2F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}$$

and

$$\rho_t^{\text{cin}} = \frac{1}{1+\sum_{\tau=0}^{T-t-j_1}\beta^{j_1+\tau}}\left(1-\sum_{k=1}^{i-1}\rho_{t-i+k}^{\text{cin}}-\sum_{k=1}^{j_1-1}\rho_{t-i+k}^{\text{cin}}\right).$$

## APPENDIX C: PROOFS OF THE DYNAMIC LIFE-CYCLE PORTFOLIO-CHOICE MODEL

### C.1. *Proof of Proposition 1*

If the consumption function derived in Appendix B is admissible, then the equilibrium exists and is unique as the equilibrium solution is obtained by maximizing the agent's objective function, which is globally concave. Moreover, there is a finite period $T$ that uniquely determines the equilibrium. The derivative of the agent's maximization problem for attentive consumption (see Appendix B.1) in any period $t$ is given by

$$\begin{aligned}
&u'(C_t)\big(1+\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big) \\
&\quad -\gamma u'(W_t-C_t)\sum_{j=1}^{T-t}\beta^j\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big) \\
&\quad -u'(W_t-C_t)\sum_{j=1}^{T-t}\beta^j=0.
\end{aligned}$$

Note that $\gamma\sum_{j=1}^{T-t}\beta^j(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda(1-F_r(r_t)\cdots F_r(r_{t-i+1})))$ and $\sum_{j=1}^{T-t}\beta^j$ are constant in $C_t$ and positive. In turn, the sufficient condition for an optimum is

$$u''(C_t)\text{ positive const.}+u''(W_t-C_t)\text{ positive const.}<0,$$

$$-\frac{1}{C_t^2}\text{ positive const.}-\frac{1}{(W_t-C_t)^2}\text{ positive const.}<0,$$

which is always true. Equivalently, the derivative of the agent's maximization problem for inattentive consumption as derived in Appendix B.1 in any period $t$ is given by

$$u'\big(C_t^{\text{in}}\big) - \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau} u'\left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i+j_1-1} \frac{C_{t-i+k}^{\text{in}}}{\big(R^d\big)^k}\right) \frac{1}{\big(R^d\big)^i} = 0,$$

$$-\frac{1}{\big(C_t^{\text{in}}\big)^2} - \frac{1}{\big(R^d\big)^i} \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau} \frac{1}{\left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i-1} \frac{C_{t-i+k}^{\text{in}}}{\big(R^d\big)^k} - \frac{C_t^{\text{in}}}{\big(R^d\big)^i} - \sum_{k=1}^{j_1-1} \frac{C_{t+k}^{\text{in}}}{\big(R^d\big)^{i+k}}\right)^2} \frac{1}{\big(R^d\big)^i} < 0,$$

which is always true. In turn, this derivative has a unique point at which it equals zero, which is the global maximum as the maximization problem is concave. In turn, it has to be ensured that the consumption function is admissible by the condition that $\sigma \geq \sigma_t^*$. For attentive periods, $\sigma_t^*$ is implicitly defined by the log monotone consumption function restriction $\frac{\partial \log(C_t)}{\partial(r_t+\cdots+r_{t-i})} > 0$ as

$$\begin{aligned}
\log(C_t) &= \log(W_t) + \log(\rho_t) \\
&= \log\left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i-1} \frac{C_{t-i+k}^{\text{in}}}{\big(R^d\big)^k}\right) + \log\big(R_t^p\big) \\
&\quad + \log\left(\frac{1}{1 + \sum_{\tau=1}^{T-t} \beta^\tau \frac{1 + \gamma\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1 + \eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}\right)
\end{aligned}$$

as

$$\begin{aligned}
\frac{\partial \log\big(R_t^p\big)}{\partial(r_t+\cdots+r_{t-i})} &= \frac{\partial \log\big(\big(R^f\big)^i + \alpha_{t-i}\big(R_t\cdots R_{t-i+1} - \big(R^f\big)^i\big)\big)}{\partial(R_t\cdots R_{t-i+1})} \frac{\partial(R_t\cdots R_{t-i+1})}{\partial(r_t+\cdots+r_{t-i+1})} \\
&= \frac{1}{\big(R^f\big)^i + \alpha_{t-i}\big(R_t\cdots R_{t-i+1} - \big(R^f\big)^i\big)} \alpha_{t-i} \frac{\partial R_t\cdots R_{t-i+1}}{\partial(r_t+\cdots+r_{t-i+1})} \\
&= \frac{1}{\big(R^f\big)^i + \alpha_{t-i}\big(R_t\cdots R_{t-i+1} - \big(R^f\big)^i\big)} \alpha_{t-i} \frac{1}{\frac{\partial \log(R_t\cdots R_{t-i+1})}{\partial R_t\cdots R_{t-i+1}}} \\
&= \alpha_{t-i} \frac{R_t\cdots R_{t-i+1}}{\big(R^f\big)^i + \alpha_{t-i}\big(R_t\cdots R_{t-i+1} - \big(R^f\big)^i\big)}
\end{aligned}$$

the restrictions are equivalent to $\frac{\partial \log(\rho_t)}{\partial(r_t+\cdots+r_{t-i+1})} > -\alpha_{t-i}\frac{R_t\cdots R_{t-i+1}}{(R^f)^i+\alpha_{t-i}(R_t\cdots R_{t-i+1}-(R^f)^i)}$, then $\sigma_t^*$ is implicitly defined by the restriction

$$\frac{\partial \log(\rho_t)}{\partial\big(F_r(r_t)\cdots F_r(r_{t-i+1})\big)} \frac{\partial\big(F_r(r_t)\cdots F_r(r_{t-i+1})\big)}{\partial(r_t+\cdots+r_{t-i+1})} > -\alpha_{t-i} \frac{R_t\cdots R_{t-i+1}}{\big(R^f\big)^i + \alpha_{t-i}\big(R_t\cdots R_{t-i+1} - \big(R^f\big)^i\big)},$$

noting that $\frac{\partial(F_r(r_t)\cdots F_r(r_{t-i+1}))}{\partial(r_t+\cdots+r_{t-i+1})} > 0$ but

$$\frac{\partial \log(\rho_t)}{\partial\big(F_r(r_t)\cdots F_r(r_{t-i+1})\big)} = -\frac{\dfrac{(1-\gamma)\eta(\lambda-1)\sum_{\tau=1}^{T-t}\beta^\tau}{\big(1+\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)^2}}{\left(1+\sum_{\tau=1}^{T-t}\beta^\tau \dfrac{1+\gamma\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1+\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda\big(1-F_r(r_t)\cdots F_r(r_{t-i+1})\big)}\right)^2} < 0.$$

Increasing $\sigma$ unambiguously decreases

$$\frac{\partial\big(F_r(r_t)\cdots F_r(r_{t-i+1})\big)}{\partial(r_t+\cdots+r_{t-i+1})} = \frac{\partial F_{\log(R_t\cdots R_{t-i+1})}\big(\log(R_t\cdots R_{t-i+1})\big)}{\partial \log(R_t\cdots R_{t-i+1})} = f_{\log(R_t\cdots R_{t-i+1})}\big(\log(R_t\cdots R_{t-i+1})\big).$$

Moreover, increasing $\sigma$ decreases $F_r(r_t)\cdots F_r(r_{t-i+1})$ which increases $\eta F_r(r_t)\cdots \times F_r(r_{t-i+1}) + \eta\lambda(1-F_r(r_t)\cdots F_r(r_{t-i+1})))$ and decreases $-\frac{\partial \log(\rho_t)}{\partial(F_r(r_t)\cdots F_r(r_{t-i+1}))}$. Thus, there exists a condition $\sigma \geq \sigma_t^*$ for all $t$ which ensures that an admissible consumption function exists.[32] Moreover, because the consumption share in inattentive periods is constant, inattentive consumption $C_t^{\text{in}}$ is necessarily increasing in $W_{t-i} - C_{t-i}$ and thus $r_{t-i}+\cdots+r_{t-i-j_0+1}$ (as $\frac{\partial \rho_{t-i}}{\partial(r_{t-i}+\cdots+r_{t-i-j_0+1})} < 0$ and thus $\frac{\partial(1-\rho_{t-i})}{\partial(r_{t-i}+\cdots+r_{t-i-j_0+1})} > 0$). Finally, the agent would not choose a suboptimal portfolio share to affect the attentiveness of future selves because the agent is not subject to a time inconsistency with respect to his plan of when to look up his portfolio so long as the agent is not subject to time-inconsistent overconsumption in inattentive periods. The agent does not deviate and overconsume in inattentive periods if the parameter restriction $\lambda > \frac{1}{\gamma} + \Delta$ holds with $\Delta$ determined by the following inequalities: $u'(C_{t-i+k}^{\text{in}})(1+\eta) < \beta^j u'(C_{t-i+k+j}^{\text{in}})(1+\gamma\eta\lambda)$ such that $u'(\rho_{t-i+k}^{\text{in}})(1+\eta) < (\beta/R^d)^j u'(\rho_{t-i+k+j}^{\text{in}})(1+\gamma\eta\lambda)$ for $k = 1,\ldots,i-2$ and $j = 1,\ldots,i-k-1$. Given that $\rho_{t-i+k}^{\text{in}}(\beta/R^d)^j \approx \rho_{t-i+k+j}^{\text{in}}$, $\Delta$ is small. This condition ensures that the agent is not going to deviate from his inattentive consumption path in periods $t-i+1$ to $t-1$, because the bad news from deviating outweighs the good news from doing so in all inattentive periods. In turn, in inattentive periods, the agent does not experience news utility in equilibrium because no uncertainty resolves and he cannot fool himself.

### C.2. *Proof of Corollary 1*

The proof of Corollary 1 in the dynamic model is very similar to the proof of Proposition 1 in the static model. Please refer to Appendix B for the derivation of the dynamic portfolio share; redefining $\mu \triangleq h\mu$, $\sigma \triangleq \sqrt{h}\sigma$, $r^f \triangleq hr^f$, and $\beta \triangleq \beta^h$, the optimal portfolio

[32] If $\sigma < \sigma_t^*$ for some $t$, the agent would optimally choose a flat section that spans the part his consumption function decreases. In that situation, the admissible consumption function requirement is weakly satisfied and the model's equilibrium is not affected qualitatively or quantitatively.

share (in any period $t$ as the agent has to look up his portfolio every period) if $0 \leq \alpha_t \leq 1$ can be rewritten as

$$\alpha_t = \frac{\mu - r^f + \frac{\sqrt{h}}{h}\sigma \frac{\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} E_t\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]}{1+\gamma\big(\eta F_r(r_t)+\eta\lambda\big(1-F_r(r_t)\big)\big)}}{\sigma^2}.$$

Samuelson's colleague and time diversification: As can be easily seen, $\lim \frac{\sqrt{h}}{h} \to \infty$ as $h \to 0$ whereas $\lim \frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} \to \frac{1+\gamma(T-t-1)}{T-t}$ as $h \to 0$, there exists some $\underline{h}$ such that $\mu - r^f > -\frac{\sqrt{h}}{h}\sigma \frac{\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} E_t[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})]}{1+\gamma(\eta F(s_t)+\eta\lambda(1-F(s_t)))}$ and thus $\alpha_t = 0$ for any $t$. In contrast, $\alpha^s > 0$ whenever $\mu > r^f$. On the other hand, $\alpha_t > 0$ for some $h$ as if $h \to \infty$, then $\lim \frac{\sqrt{h}}{h} \to 0$, $\lim \frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} \to 1$, and $\alpha_t \to \alpha^s$. Furthermore, it can be seen that $\frac{\partial \alpha_t}{\partial h} > 0$ since

$$\begin{aligned}
&\frac{\partial}{\partial h}\frac{\sqrt{h}}{h}\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} \\
&= \frac{\frac{1}{2\sqrt{h}}h-\sqrt{h}}{h^2}\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} \\
&\quad + \frac{\sqrt{h}}{h}\frac{\gamma\sum_{k=1}^{T-t-1}\log(\beta)k\beta^{kh}\sum_{\tau=0}^{T-t-1}\beta^{\tau h} - \sum_{\tau=0}^{T-t-1}\log(\beta)\tau\beta^{\tau h}\left(1+\gamma\sum_{k=1}^{T-t-1}\beta^{kh}\right)}{\left(\sum_{\tau=0}^{T-t-1}\beta^{\tau h}\right)^2} \\
&< 0
\end{aligned}$$

if $\beta \approx 1$. The first term is necessarily negative as $\frac{\frac{1}{2\sqrt{h}}h-\sqrt{h}}{h^2} < 0 \Rightarrow \frac{1}{2\sqrt{h}}h - \sqrt{h} < 0 \Rightarrow \frac{1}{2}h < h$ which is multiplied by $\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}} > 0$. Moreover, if $\beta \approx 1$, the first negative term will necessarily dominate the second term as $\log(\beta) \approx 0$. Moreover, $\frac{1+\gamma\sum_{\tau=1}^{T-t-1}\beta^{\tau h}}{\sum_{\tau=0}^{T-t-1}\beta^{\tau h}}$ decreases in $T-t$ if $\gamma < 1$ such that $\alpha_t$ decreases in $T-t$ while $\alpha^s$ is constant.

### C.3. *Proof of Proposition 2*

The basic intuition of the proof is as follows: the benefits to be gained from consumption smoothing are proportional to the length of a period $h$ and are second order because the agent deviates from an initially optimal consumption path. The costs of experiencing news utility are proportional to $\sqrt{h}$ and are first order. Thus as $h$ becomes smaller, the benefits of consumption smoothing decrease relative to the costs of news utility. Moreover, inattention has the additional benefit that the agent overconsumes less when inattentive. More formally, I pick $t$ such that $T-t$ is large and I can simplify the exposition by replacing $\sum_{k=1}^{T-t}\beta^k$ with $\frac{\beta}{1-\beta}$. If the agent is attentive every period, his value function is given by

$$\beta E_{t-1}\big[V_t(W_t)\big] = E_{t-1}\left[\frac{\beta}{1-\beta}\log(W_t)\right] + \psi_t^{t-1}(\alpha_{t-1}),$$

$$\begin{aligned}\psi_t^{t-1}(\alpha_{t-1}) = \beta E_{t-1}\Bigg[&\log(\rho_t) + \alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]\\ &+\frac{\beta}{1-\beta}\log(1-\rho_t) + \frac{\beta}{1-\beta}\left(r^f + \alpha_t\big(r_{t+1}-r^f\big) + \alpha_t(1-\alpha_t)\frac{\sigma^2}{2}\right)\\ &+\psi_{t+1}^{t}(\alpha_t)\Bigg].\end{aligned}$$

Now, I show when the expected utility from being inattentive for one period is larger than the expected utility from being attentive for all periods:

$$\begin{aligned}&E_{t-1}\left[\log\big(C_t^{\text{in}}\big) + \frac{\beta}{1-\beta}\log(W_{t+1}) + \psi_{t+1}^{t-1}(\alpha_{t-1})\right]\\ &\quad= E_{t-1}\Bigg[\log(W_{t-1}-C_{t-1}) + \log\big(R^d\big) + \log\big(\rho_t^{\text{in}}\big)\\ &\qquad+\frac{\beta}{1-\beta}\log\left(W_{t-1}-C_{t-1}-\frac{C_t^{\text{in}}}{R^d}\right)\\ &\qquad+\frac{\beta}{1-\beta}\big(2r^f + \alpha_{t-1}\big(r_t + r_{t+1} - 2r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\sigma^2\big) + \psi_{t+1}^{t-1}(\alpha_{t-1})\Bigg]\\ &\quad> E_{t-1}\Bigg[\log(C_t) + \alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]\\ &\qquad+\frac{\beta}{1-\beta}\log(W_{t+1}) + \psi_{t+1}^{t}(\alpha_t)\Bigg]\\ &\quad= E_{t-1}\Bigg[\log(W_{t-1}-C_{t-1}) + r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} + \log(\rho_t)\\ &\qquad+\alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^{\infty}(s-\tilde{s})\,dF(\tilde{s})\right]\\ &\qquad+\frac{\beta}{1-\beta}\log(W_{t+1}) + \psi_{t+1}^{t}(\alpha_t)\Bigg]\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad & E_{t-1}\Big[r^d + \log\big(\rho_t^{\text{in}}\big) + \frac{\beta}{1-\beta}\log(W_{t-1}-C_{t-1}) + \frac{\beta}{1-\beta}\log\big(1-\rho_t^{\text{in}}\big) \\
& + \frac{\beta}{1-\beta}\big(2r^f + \alpha_{t-1}\big(r_t + r_{t+1} - 2r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\sigma^2\big) + \psi_{t+1}^{t-1}(\alpha_{t-1})\Big] \\
& > E_{t-1}\Big[r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} + \log(\rho_t) \\
& + \alpha_{t-1}\Big(1 + \gamma\frac{\beta}{1-\beta}\Big)\sigma E\Big[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\Big] \\
& + \frac{\beta}{1-\beta}\log(W_{t-1}-C_{t-1}) + \frac{\beta}{1-\beta}\Big(r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2}\Big) \\
& + \frac{\beta}{1-\beta}\log(1-\rho_t) + \frac{\beta}{1-\beta}\Big(r^f + \alpha_t\big(r_{t+1} - r^f\big) + \alpha_t(1-\alpha_t)\frac{\sigma^2}{2}\Big) + \psi_{t+1}^{t}(\alpha_t)\Big] \\
\Rightarrow \quad & E_{t-1}\Big[r^d + \log\big(\rho_t^{\text{in}}\big) + \frac{\beta}{1-\beta}\log\big(1-\rho_t^{\text{in}}\big) \\
& + \frac{\beta}{1-\beta}\big(2r^f + \alpha_{t-1}\big(r_t + r_{t+1} - 2r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\sigma^2\big) + \psi_{t+1}^{t-1}(\alpha_{t-1})\Big] \\
& > E_{t-1}\Big[r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} + \log(\rho_t) \\
& + \alpha_{t-1}\Big(1 + \gamma\frac{\beta}{1-\beta}\Big)\sigma E\Big[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\Big] \\
& + \frac{\beta}{1-\beta}\Big(r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2}\Big) + \frac{\beta}{1-\beta}\log(1-\rho_t) \\
& + \frac{\beta}{1-\beta}\Big(r^f + \alpha_t\big(r_{t+1} - r^f\big) + \alpha_t(1-\alpha_t)\frac{\sigma^2}{2}\Big) + \psi_{t+1}^{t}(\alpha_t)\Big].
\end{aligned}$$

As $T-t$ is large, for an average period $t-1$ it holds that $E_{t-1}[\alpha_t] \approx \alpha_{t-1}$ (as can be easily seen by looking at the equation for the portfolio share that converges) such that

$$\begin{aligned}
& E_{t-1}\Big[r^d + \log\big(\rho_t^{\text{in}}\big) + \frac{\beta}{1-\beta}\log\big(1-\rho_t^{\text{in}}\big) + \psi_{t+1}^{t-1}(\alpha_{t-1})\Big] \\
& > E_{t-1}\Big[r^f + \alpha_{t-1}\big(r_t - r^f\big) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} \\
& + \log(\rho_t) + \alpha_{t-1}\Big(1 + \gamma\frac{\beta}{1-\beta}\Big)\sigma E\Big[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\Big] \\
& + \frac{\beta}{1-\beta}\log(1-\rho_t) + \psi_{t+1}^{t}(\alpha_t)\Big].
\end{aligned}$$

The agent's continuation utilities are given by

$$\begin{aligned}\psi_{t+1}^{t-1}(\alpha_{t-1}) = \beta E_{t-1}\biggl[&\log(\rho_{t+1}) + \sqrt{2}\alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]\\&+\frac{\beta}{1-\beta}\log(1-\rho_{t+1}) + \frac{\beta}{1-\beta}\left(r^f + \alpha_{t+1}\left(r_{t+2}-r^f\right) + \alpha_{t+1}(1-\alpha_{t+1})\frac{\sigma^2}{2}\right)\\&+\psi_{t+2}^{t+1}(\alpha_{t+1})\biggr],\\ \psi_{t+1}^{t}(\alpha_t) = \beta E_t\biggl[&\log(\rho_{t+1}) + \alpha_t\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]\\&+\frac{\beta}{1-\beta}\log(1-\rho_{t+1}) + \frac{\beta}{1-\beta}\left(r^f + \alpha_{t+1}\left(r_{t+2}-r^f\right) + \alpha_{t+1}(1-\alpha_{t+1})\frac{\sigma^2}{2}\right)\\&+\psi_{t+2}^{t+1}(\alpha_{t+1})\biggr].\end{aligned}$$

The agent's behavior from period $t+2$ on is not going to be affected by his period $t$ (in)attentiveness (as his period-$t+1$ self can be forced to look up the portfolio by his period-$t$ self). Moreover, as $T-t$ is large, for an average period $t-1$ it holds that $E_{t-1}[\alpha_t] \approx \alpha_{t-1}$ such that

$$\begin{aligned}&E_{t-1}\left[\psi_{t+1}^{t-1}(\alpha_{t-1}) - \psi_{t+1}^{t}(\alpha_t)\right]\\&\quad = E_{t-1}\left[(\sqrt{2}-1)\alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]\right]\\&\quad\Rightarrow\quad E_{t-1}\biggl[r^d + \log\left(\rho_t^{\text{in}}\right) + \frac{\beta}{1-\beta}\log\left(1-\rho_t^{\text{in}}\right)\\&\qquad\qquad + (\sqrt{2}-2)\alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]\biggr]\\&\qquad\quad > E_{t-1}\left[r^f + \alpha_{t-1}\left(r_t - r^f\right) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} + \log(\rho_t) + \frac{\beta}{1-\beta}\log(1-\rho_t)\right],\end{aligned}$$

which finally results in the following comparison:

$$\begin{aligned}&E_{t-1}\left[\log\left(\rho_t^{\text{in}}\right) + \frac{\beta}{1-\beta}\log\left(1-\rho_t^{\text{in}}\right)\right]\\&\quad > E_{t-1}\biggl[\underbrace{r^f + \alpha_{t-1}\left(r_t - r^f\right) + \alpha_{t-1}(1-\alpha_{t-1})\frac{\sigma^2}{2} - r^d}_{>0 \text{ in expectation and increasing with } h}\\&\qquad + \log(\rho_t) + \frac{\beta}{1-\beta}\log(1-\rho_t)\\&\qquad + \underbrace{(2-\sqrt{2})\alpha_{t-1}\left(1+\gamma\frac{\beta}{1-\beta}\right)\sigma E\left[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\right]}_{<0 \text{ and increasing with } \sqrt{h}}\biggr].\end{aligned}$$

In turn, I can prove that the agent will be inattentive for at least one period if $h < \underline{h}$ (redefining $\mu \triangleq h\mu$, $\sigma \triangleq \sqrt{h}\sigma$, $r^f \triangleq hr^f$, and $\beta \triangleq \beta^h$). If I decrease $h$, I decrease the positive return part, which becomes quantitatively less important (as it is proportional to $h$) relative to the negative news-utility part (as it is proportional to $\sqrt{h}$). Moreover, the difference in consumption utilities speaks towards not looking up the portfolio, too, that is, $E_{t-1}[\log(\rho_t^{\text{in}}) + \frac{\beta}{1-\beta}\log(1-\rho_t^{\text{in}})] - E_{t-1}[\log(\rho_t) + \frac{\beta}{1-\beta}\log(1-\rho_t)] > 0$ for any $h$. The difference in consumption utilities is positive because $\log(\rho_t^{\text{in}}) + \frac{\beta}{1-\beta}\log(1-\rho_t^{\text{in}})$ is maximized for $\rho_t^{\text{in}} = \frac{1}{1+\frac{\beta}{1-\beta}}$, which corresponds to the standard agent's portfolio share as the agent is inattentive for just one period. The intuition for this additional reason to ignore the portfolio is that inattentive consumption is not subject to a self-control problem while attentive consumption is. Furthermore, $h$ affects the prospective news utility term via $1 + \gamma\frac{\beta}{1-\beta}$. However, as $\frac{\beta}{1-\beta}$ increases if $h$ decreases, this will only make the agent more likely to be inattentive. Thus, I conclude that the agent will find it optimal to be inattentive for at least one period if $h < \underline{h}$. It cannot be argued that the agent would behave differently than what is assumed from period $t+1$ on unless he finds it optimal to do so from the perspective of period $t$ because the agent can restrict the funds in the checking account and determine whether or not his period-$t+1$ self is attentive.

In turn, if I increase $h$, I increase the positive return part, which becomes quantitatively relatively more important than the negative news-utility part. Increasing $h$ implies that the difference between the positive return part and the negative news-utility part will at some point exceed the difference in consumption utilities $E_{t-1}[\log(\rho_t^{\text{in}}) + \frac{\beta}{1-\beta}\log(1-\rho_t^{\text{in}})] - E_{t-1}[\log(\rho_t) + \frac{\beta}{1-\beta}\log(1-\rho_t)]$, which is positive as shown above. Moreover, any increase in the difference in consumption utilities due to the increase in $h$ will be less than the rate at which the difference between the return and news-utility part increases if $h$ becomes large. The reason is that an increase in $h$ will result in a decrease in $\beta$ and thereby a decrease in $\frac{\beta}{1-\beta}$ as $\frac{\partial \frac{\beta^h}{1-\beta^h}}{\partial h} = \frac{\beta^h \log(\beta)}{(1-\beta^h)^2}$, which goes to zero as $h \to \infty$ (and $\frac{\beta}{1-\beta}$ fully determines $\rho_t^{\text{in}}$ and $\rho_t$). Thus, I conclude that there exists some $h > \bar{h}$ such that the agent will find it optimal to be attentive in every period.

## C.4. *Proof of Proposition 3*

Please refer to Appendix B for the derivation of the dynamic portfolio share. The expected benefit of inattention (as defined in the text) is given by

$$-\Bigg(\sqrt{i}E[\alpha_{t-i}]\Bigg(1+\gamma\sum_{\tau=1}^{T-t}\beta^\tau\Bigg) + \beta^{j_1}\big(\sqrt{j_1}E[\alpha_t] - E[\alpha_{t-i}]\sqrt{j_1+i}\big)\Bigg(1+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^\tau\Bigg)\Bigg)\sigma E\bigg[\eta(\lambda-1)\int_s^\infty (s-\tilde{s})\,dF(\tilde{s})\bigg],$$

and is always positive if $E[\alpha_{t-i}] \approx E[\alpha_t]$, which is necessarily the case if $T-t$ is large. As can be easily inferred, $\frac{1+\gamma\sum_{\tau=1}^{T-t-j_1}\beta^\tau}{\sum_{\tau=0}^{T-t-j_1}\beta^\tau}$ and thus $\alpha_{t-i}$ is converging as $T-t$ becomes large and thus decreases quickly if $T-t$ is small. Therefore, towards the end of life, $\alpha_t$ decreases more quickly, that is, $\frac{\partial(\alpha_{t-i} - E_{t-1}[\alpha_t])}{\partial(T-t)} < 0$, and the expected benefit of inattention is lower if

$E[\alpha_{t-i}] > E[\alpha_t]$. If $T-t$ is large, the benefit of inattention is lower if $\alpha_{t-i} > E[\alpha_t]$ and, moreover, $\frac{\partial \alpha_t}{\partial r_t + \cdots + r_{t-i+1}} < 0$. Thus, the benefit of inattention is low if $r_{t-i} + \cdots + r_{t-i-j_0+1}$ is low and thus $\alpha_{t-i}$ is high.

### C.5. *Proof of Proposition 4*

I start with the inattentive marginal propensity to consume (MPC) out of disposable wealth, $\frac{\partial C_t^{\text{in}}}{\partial W_t^{\text{in}}}$, noting that $W_t^{\text{in}} = (W_{t-i} - C_{t-i} - \sum_{k=1}^{i-1} \frac{C_{t-i+k}^{\text{in}}}{(R^d)^k})(R^d)^i = (W_{t-i} - C_{t-i})(R^d)^i - \sum_{k=1}^{i-1} C_{t-i+k}^{\text{in}}(R^d)^{i-k}$. For $\beta \approx 1$, the inattentive consumption share $\rho_t^{\text{in}}$ is the same for each period within each inattention spell as can be seen easily:

$$\rho_t^{\text{in}} = \frac{1}{1 + \sum_{\tau=0}^{T-t-j_1} \beta^{j_1+\tau}} \left(1 - \sum_{k=1}^{i-1} \rho_{t-i+k}^{\text{in}} - \sum_{k=1}^{j_1-1} \rho_{t+k}^{\text{in}}\right)$$

$$= \frac{1}{1 + T - t - j_1 + 1} \left(1 - \sum_{k=1}^{i-1} \rho_{t-i+k}^{\text{in}} - \sum_{k=1}^{j_1-1} \rho_{t+k}^{\text{in}}\right) = \rho_{t-1}^{\text{in}},$$

so that

$$\rho_t^{\text{in}} = \frac{1}{1 + T - t - j_1 + 1}\big(1 - (i-1)\rho_t^{\text{in}} - (j_1 - 1)\rho_t^{\text{in}}\big) \quad \Rightarrow \quad \rho_t^{\text{in}} = \frac{1}{T - t + i}.$$

In turn, noting that $C_t^{\text{in}} = (W_{t-i} - C_{t-i})(R^d)^i \rho_t^{\text{in}}$, one can compute $\frac{\partial C_t^{\text{in}}}{\partial W_t^{\text{in}}}$

$$\frac{\partial (W_{t-i} - C_{t-i})\big(R^d\big)^i \rho_t^{\text{in}}}{\partial \left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i-1} \frac{(W_{t-i} - C_{t-i})\big(R^d\big)^k \rho_{t-i+k}^{\text{in}}}{\big(R^d\big)^k}\right)\big(R^d\big)^i}$$

$$= \frac{\partial (W_{t-i} - C_{t-i})\rho_t^{\text{in}}}{\partial (W_{t-i} - C_{t-i})\big(1 - (i-1)\rho_t^{\text{in}}\big)} = \frac{\partial \frac{1}{T-t+i}}{\partial \left(1 - (i-1)\frac{1}{T-t+i}\right)}$$

$$= \frac{\partial \frac{1}{T-t+i}}{\partial \left(\frac{T-t+i-i+1}{T-t+i}\right)} = \frac{1}{1+T-t}.$$

In contrast, when the agent is attentive,

$$\frac{\partial C_t}{\partial W_t} = \rho_t = \frac{1}{1 + (T-t)\frac{1 + \gamma\big(\eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1 + \eta F_r(r_t)\cdots F_r(r_{t-i+1}) + \eta\lambda\big(1 - F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}.$$

The two MPCs differ by the fraction $\frac{1+\gamma(\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda(1-F_r(r_t)\cdots F_r(r_{t-i+1}))}{1+\eta F_r(r_t)\cdots F_r(r_{t-i+1})+\eta\lambda(1-F_r(r_t)\cdots F_r(r_{t-i+1}))} < 1$ whenever $\gamma < 1$ such that $\frac{\partial C_t}{\partial W_t} > \frac{\partial C_t^{\text{in}}}{\partial W_t^{\text{in}}}$.

## C.6. *Proof of Proposition 5*

Comparing the precommitted-monotone and personal-monotone portfolio share, it can be easily seen that they are not the same for any period $t \in \{1, \dots, T-1\}$.

1. The precommitted consumption share for attentive consumption is

$$\rho_t^c = \frac{1}{1 + \sum_{\tau=1}^{T-t} \beta^\tau \frac{\big(1 + \gamma\eta(\lambda - 1)\big(1 - 2F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1 + \eta(\lambda - 1)\big(1 - 2F_r(r_t)\cdots F_r(r_{t-i+1})\big)}},$$

which is lower than the personal-monotone share if $\gamma < 1$ as $\eta(\lambda - 1)(1 - 2F_r(r_t)) < \eta F_r(r_t) + \eta\lambda(1 - F_r(r_t)))$ and the difference increases if $F_r(r_t)\cdots F_r(r_{t-i+1})$ increases. $\gamma$ does not necessarily imply an increase in $\rho_t^c$ because $\eta(\lambda-1)(1-2F_r(r_t)\cdots F_r(r_{t-i+1}))$ can be negative. Thus, attentive consumption is higher only due to the differences in returns and not due to the time inconsistency any more. In contrast, inattentive consumption shares are the same for the precommitted and non-precommitted agent.

2. The precommitted portfolio share is

$$\alpha_t^c = \frac{\mu - r^f + \frac{\frac{1 + \gamma \sum_{\tau=1}^{T-t-j_1} \beta^\tau}{\sum_{\tau=0}^{T-t-j_1} \beta^\tau} E\left[\eta(\lambda - 1)\int_r^\infty (r_T - \tilde{r})\, dF_r(\tilde{r})\right]}{1 + \gamma\eta(\lambda - 1)\big(2F_r(r_t)\cdots F_r(r_{t-i+1}) - 1\big)}}{\sigma^2},$$

which is lower than the personal-monotone share as $\eta(\lambda - 1)(1 - 2F_r(r_t)) < \eta F_r(r_t) + \eta\lambda(1 - F_r(r_t))$ and the difference increases if $F_r(r_t)\cdots F_r(r_{t-i+1})$ increases. As precommitted marginal news utility is always lower and the gap increases in good states, there is larger variation in it, that is, it varies from $\{-\eta(\lambda - 1), \eta(\lambda - 1)\}$, which is larger than the variation in non-precommitted marginal news $\{\eta, \eta\lambda\}$, as $2\eta(\lambda - 1) > \eta(\lambda - 1)$.

3. In the precommitted equilibrium, the agent's marginal propensities to consume are different but not systematically so any more. From above, $\frac{\partial C_t^{\text{in}}}{\partial W_t^{\text{in}}}$ is the same

$$\frac{\partial(W_{t-i} - C_{t-i})\big(R^d\big)^i \rho_t^{\text{in}}}{\partial\left(W_{t-i} - C_{t-i} - \sum_{k=1}^{i-1} \frac{(W_{t-i} - C_{t-i})\big(R^d\big)^k \rho_{t-i+k}^{\text{in}}}{\big(R^d\big)^k}\right)\big(R^d\big)^i} = \frac{1}{1 + T - t}.$$

In contrast, when the agent is attentive,

$$\frac{\partial C_t}{\partial W_t} = \rho_t = \frac{1}{1 + (T - t)\frac{\big(1 + \gamma\eta(\lambda - 1)\big(1 - 2F_r(r_t)\cdots F_r(r_{t-i+1})\big)\big)}{1 + \eta(\lambda - 1)\big(1 - 2F_r(r_t)\cdots F_r(r_{t-i+1})\big)}}.$$

The two MPCs differ by the fraction $\frac{1+\gamma\eta(\lambda-1)(1-2F_r(r_t)\cdots F_r(r_{t-i+1}))}{1+\eta(\lambda-1)(1-2F_r(r_t)\cdots F_r(r_{t-i+1}))} \lesseqgtr 1$ even if $\gamma < 1$ such that $\frac{\partial C_t}{\partial W_t} \lesseqgtr \frac{\partial C_t^{\text{in}}}{\partial W_t^{\text{in}}}$.

### C.7. *Proof of Proposition 6*

I simply assume that on top of facing uncertainty over the risky return $R$ of his wealth $W > 0$, the agent will receive riskless labor income $\bar{Y} \geq 0$ and risky labor income $Y > 0$. Thus, the agent's consumption is $C = W(R^f + \alpha(R - R^f)) + \bar{Y} + Y$ and consumption utility is denoted by $u(\cdot)$. While I assume $\text{Cov}(R, Y) = 0$, I will outline the implications of $\text{Cov}(R, Y) \neq 0$. The joint distribution of labor income and the return $R$ is denoted by $F_{RY}$ (marginal and conditional distributions are denoted by $F_Y$ and $F_R$ and $F_{Y|R}$ and $F_{R|Y}$, respectively). To lighten notation, I initially set $\bar{Y} = 0$. The agent's risk premium (i.e., compensating utility differential) for investing into the risky return $R$ as opposed to receiving the unconditional expected value $E[R] = E_{RY}[R] = \int\int R\, dF_{RY}(R, Y)$ is then given by

$$\begin{aligned}
\pi = E_Y\Big[&u\big((1-\alpha)WR^f + \alpha W E[R] + Y\big) \\
&+ \eta(\lambda - 1)\int_Y^\infty \big(u\big((1-\alpha)WR^f + \alpha W E[R] + Y\big) \\
&- u\big((1-\alpha)WR^f + \alpha W E[R] + \tilde{Y}\big)\big)\, dF_Y(\tilde{Y})\Big] \\
&- E_{RY}\Big[u\big((1-\alpha)WR^f + \alpha W R + Y\big) \\
&+ \eta(\lambda - 1)\int_R^\infty \int_Y^\infty \big(u\big((1-\alpha)WR^f + \alpha W R + Y\big) \\
&- u\big((1-\alpha)WR^f + \alpha W \tilde{R} + \tilde{Y}\big)\big)\, dF_{RY}(\tilde{R}, \tilde{Y})\Big].
\end{aligned}$$

In turn, his marginal value for an additional increment of portfolio risk is

$$\begin{aligned}
\frac{\partial \pi}{\partial \alpha} = E_Y\Big[&u'\big((1-\alpha)WR^f + \alpha W E[R] + Y\big)W\big(E[R] - R^f\big) \\
&+ \eta(\lambda - 1)\int_Y^\infty \big(u'\big((1-\alpha)WR^f + \alpha W E[R] + Y\big)W\big(E[R] - R^f\big) \\
&- u'\big((1-\alpha)WR^f + \alpha W E[R] + \tilde{Y}\big)W\big(E[R] - R^f\big)\big)\, dF_Y(\tilde{Y})\Big] \\
&- E_{RY}\Big[u'\big((1-\alpha)WR^f + \alpha W R + Y\big)W\big(R - R^f\big) \\
&+ \eta(\lambda - 1)\int_R^\infty \int_Y^\infty \big(u'\big((1-\alpha)WR^f + \alpha W R + Y\big)W\big(R - R^f\big) \\
&- u'\big((1-\alpha)WR^f + \alpha W \tilde{R} + \tilde{Y}\big)W\big(\tilde{R} - R^f\big)\big)\, dF_{RY}(\tilde{R}, \tilde{Y})\Big].
\end{aligned}$$

Now, what happens if return risk becomes small, that is, $\alpha \to 0$:

$$\begin{aligned}\left.\frac{\partial \pi}{\partial \alpha}\right|_{\alpha=0} &= \eta(\lambda-1)E_Y\bigg[\int_Y^{\infty}\big(u'\big(WR^f+Y\big)W\big(E[R]-R^f\big) \\ &\quad - u'\big(WR^f+\tilde{Y}\big)W\big(E[R]-R^f\big)\big)\,dF_Y(\tilde{Y})\bigg] \\ &\quad - \eta(\lambda-1)E_{RY}\bigg[\int_R^{\infty}\int_Y^{\infty}\big(u'\big(WR^f+Y\big)W\big(R-R^f\big) \\ &\quad - u'\big(WR^f+\tilde{Y}\big)W\big(\tilde{R}-R^f\big)\big)\,dF_{RY}(\tilde{R},\tilde{Y})\bigg].\end{aligned}$$

In turn,

$$\begin{aligned} &u'\big(WR^f+Y\big)W\big(E[R]-R^f\big)-u'\big(WR^f+\tilde{Y}\big)W\big(E[R]-R^f\big) \\ &\quad = u'\big(WR^f+Y\big)WE[R]-u'\big(WR^f+\tilde{Y}\big)WE[R] \\ &\qquad - \big(u'\big(WR^f+Y\big)WR^f-u'\big(WR^f+\tilde{Y}\big)WR^f\big)\end{aligned}$$

and

$$\begin{aligned} &u'\big(WR^f+Y\big)W\big(R-R^f\big)-u'\big(WR^f+\tilde{Y}\big)W\big(\tilde{R}-R^f\big) \\ &\quad = u'\big(WR^f+Y\big)WR-u'\big(WR^f+\tilde{Y}\big)W\tilde{R} \\ &\qquad - \big(u'\big(WR^f+Y\big)WR^f-u'\big(WR^f+\tilde{Y}\big)WR^f\big)\end{aligned}$$

and the $R^f$ terms will simply cancel, such that

$$\begin{aligned}\left.\frac{\partial \pi}{\partial \alpha}\right|_{\alpha=0} &= \underbrace{\eta(\lambda-1)E_Y\bigg[\int_Y^{\infty}\big(u'\big(WR^f+Y\big)WE[R]-u'\big(WR^f+\tilde{Y}\big)WE[R]\big)\,dF_Y(\tilde{Y})\bigg]}_{>0 \text{ if } u''<0} \\ &\quad - \eta(\lambda-1)E_{RY}\bigg[\int_R^{\infty}\int_Y^{\infty}\big(u'\big(WR^f+Y\big)WR-u'\big(WR^f+\tilde{Y}\big)W\tilde{R}\big)\,dF_{RY}(\tilde{R},\tilde{Y})\bigg].\end{aligned}$$

The risk premium for small risks is increased for the news-utility agent (over the standard agent's risk premium that is zero) because the first integral is necessarily positive and dominates the second integral, which can be negative or positive as in the second integral the positive effect of $R$ enters on top of the negative effect of $Y$. More specifically, for each value of $\tilde{Y}$ for $\tilde{Y} > Y$, $u'(WR^f+Y)WE[R]-u'(WR^f+\tilde{Y})WE[R] > 0$ whenever $u' > 0$ and $u'' < 0$. Moreover, for any value of $Y$ and $\tilde{Y}$: I can maximize $\eta(\lambda-1)E_{RY}[\int_R^{\infty}\int_Y^{\infty}(u'(WR^f+Y)WR-u'(WR^f+\tilde{Y})W\tilde{R})\,dF_{RY}(\tilde{R},\tilde{Y})]$ for any $R$ by picking $\tilde{R}$ as low as possible, that is, $\tilde{R}=R+\varepsilon$ with $\varepsilon>0$ and $\varepsilon\to 0$. In turn, for $\lim_{\varepsilon\to 0}$, it follows that

$$\begin{aligned} &E_Y\bigg[WE[R]\int_Y^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})\bigg] \\ &\quad > E_{RY}\bigg[WR\int_R^{\infty}\int_Y^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_{RY}(\tilde{R},\tilde{Y})\bigg].\end{aligned}$$

To show this, cancel $W$ and transform the left-hand side to

$$E[R]\int_{-\infty}^{\infty}\int_{Y}^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})\,dF_Y(Y)$$
$$=E[R]\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{Y}^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})f_{RY}(R,Y)\,dY\,dR$$

and the right-hand side to

$$\begin{aligned}
&> E_{RY}\bigg[R\int_{R}^{\infty}\int_{Y}^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)f_{RY}(\tilde{R},\tilde{Y})\,d\tilde{R}\,d\tilde{Y}\bigg]\\
&= E_{RY}\bigg[R\int_{Y}^{\infty}f_Y(\tilde{Y})\int_{R}^{\infty}f_{R|Y}(\tilde{R}|\tilde{Y})\,d\tilde{R}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,d\tilde{Y}\bigg]\\
&= E_{RY}\bigg[R\int_{Y}^{\infty}f_Y(\tilde{Y})\big(1-F_{R|Y}(R|\tilde{Y})\big)\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,d\tilde{Y}\bigg]\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}R\int_{Y}^{\infty}f_Y(\tilde{Y})\big(1-F_{R|Y}(R|\tilde{Y})\big)\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,d\tilde{Y}\\
&\quad\times f_{RY}(R,Y)\,dY\,dR\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}R\underbrace{\int_{Y}^{\infty}\big(1-F_{R|Y}(R|\tilde{Y})\big)\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})}_{:=X}\\
&\quad\times f_{RY}(R,Y)\,dY\,dR\\
&:= E[RX]
\end{aligned}$$

with

$$E[X]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{Y}^{\infty}\big(1-F_{R|Y}(R|\tilde{Y})\big)\big(u'(WR^f+Y)-u'(WR^f+\tilde{Y})\big)\,dF_Y(\tilde{Y})$$
$$\times f_{RY}(R,Y)\,dY\,dR,$$

and as $E[RX]=E[R]E[X]+\mathrm{Cov}(R,X)$, the final comparison is

$$\begin{aligned}
&E[R]E_{RY}\bigg[\int_{Y}^{\infty}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})\bigg]\\
&\quad > E[R]E_{RY}\bigg[\int_{Y}^{\infty}\underbrace{\big(1-F_{R|Y}(R|\tilde{Y})\big)}_{<1}\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})\bigg]\\
&\qquad + \mathrm{Cov}\bigg(R,\int_{Y}^{\infty}\big(1-F_{R|Y}(R|\tilde{Y})\big)\big(u'\big(WR^f+Y\big)-u'\big(WR^f+\tilde{Y}\big)\big)\,dF_Y(\tilde{Y})\bigg).
\end{aligned}$$

This inequality always holds if $\mathrm{Cov}(R,Y)=0$ because $R$ and $1-F_{R|Y}(R|\tilde{Y})$ covary negatively. Moreover, if $\mathrm{Cov}(R,Y)>0$, then $\mathrm{Cov}(R,\int_Y^\infty(1-F_{R|Y}(R|\tilde{Y}))(u'(WR^f+Y)-u'(WR^f+\tilde{Y}))\,dF_Y(\tilde{Y}))$ is negative (if $Y$ is high and $\tilde{Y}>Y$, then $1-F_{R|Y}(R|\tilde{Y})$ is small (as $R$ is high) and $u'(WR^f + Y) - u'(WR^f + \tilde{Y})$ is small and positive as $u'(x) > 0$ with $x > 0$ is decreasing in $x$, i.e., $u''(x) < 0$).

If I instead assume that $\bar{Y} > 0$, the above exposition does not change except that all $\tilde{Y}$ (and $Y$) are replaced by $\tilde{Y} = \bar{Y} + \tilde{Y}$ (and $Y = \bar{Y} + Y$). In turn, instead of looking at the marginal value of the risk premium for an additional increment of return risk when risk becomes small, I can look at the marginal value of the risk premium for an additional increment of permanent labor income, that is, $\bar{Y} > 0$, as risk becomes small. It can be easily seen that the risk premium is decreasing in $\bar{Y}$:

$$\left.\frac{\partial \pi}{\partial \alpha}\right|_{\alpha=0} = \underbrace{\eta(\lambda-1)E_Y\left[\int_Y^\infty \left(u'\left(WR^f + \bar{Y} + Y\right)WE[R] - u'\left(WR^f + \bar{Y} + \tilde{Y}\right)WE[R]\right)dF_Y(\tilde{Y})\right]}_{>0 \text{ if } u''<0 \text{ and decreasing in } \bar{Y} \text{ if } u'''>0}$$
$$- \eta(\lambda-1)E_{YR}\left[\int_R^\infty \int_Y^\infty \left(u'\left(WR^f + \bar{Y} + Y\right)WR - u'\left(WR^f + \bar{Y} + \tilde{Y}\right)W\tilde{R}\right)dF_{RY}(\tilde{Y}, \tilde{R})\right].$$

To see the positivity of the first term, for any $R$, $u'(WR^f + \bar{Y} + Y)WR - u'(WR^f + \bar{Y} + \tilde{Y})WR$ for $\tilde{Y} > Y > 0$ is decreasing in $\bar{Y}$ as

$$\frac{\partial\left(u'\left(WR^f + \bar{Y} + Y\right)WR - u'\left(WR^f + \bar{Y} + \tilde{Y}\right)WR\right)}{\partial \bar{Y}}$$
$$= u''\left(WR^f + \bar{Y} + Y\right)WR - u''\left(WR^f + \bar{Y} + \tilde{Y}\right)WR < 0$$

if $u'' < 0$ and $u''' > 0$ (such that $u''(y) > u''(x)$ if $y > x$ and thus $u''(x) - u''(y) < 0$). Moreover, the first term of the risk premium dominates the second term:

$$\frac{\partial \left.\frac{\partial \pi}{\partial \alpha}\right|_{\alpha=0}}{\partial \bar{Y}} = \eta(\lambda-1)E_Y\left[\int_Y^\infty \underbrace{\left(u''\left(WR^f + \bar{Y} + Y\right)WE[R] - u''\left(WR^f + \bar{Y} + \tilde{Y}\right)WE[R]\right)}_{<0 \text{ if } u'''>0} dF_Y(\tilde{Y})\right]$$
$$- \eta(\lambda-1)E_{YR}\left[\int_R^\infty \int_Y^\infty \underbrace{\left(u''\left(WR^f + \bar{Y} + Y\right)WR - u''\left(WR^f + \bar{Y} + \tilde{Y}\right)W\tilde{R}\right)}_{<0 \text{ or } >0} dF_{RY}(\tilde{Y}, \tilde{R})\right].$$

For the second term, using the same argument as above, I can rewrite

$$\frac{\partial \left.\frac{\partial \pi}{\partial \alpha}\right|_{\alpha=0}}{\partial \bar{Y}} = E[R]E_{RY}\left[\int_Y^\infty \underbrace{\left(u''\left(WR^f + \bar{Y} + Y\right) - u''\left(WR^f + \bar{Y} + \tilde{Y}\right)\right)}_{<0} dF_Y(\tilde{Y})\right]$$
$$- E[R]E_{RY}\left[\int_Y^\infty \underbrace{\left(1 - F_{R|Y}(R|\tilde{Y})\right)}_{<1}\right.$$
$$\left.\times \underbrace{\left(u''\left(WR^f + \bar{Y} + Y\right) - u''\left(WR^f + \bar{Y} + \tilde{Y}\right)\right)}_{<0} dF_Y(\tilde{Y})\right]$$

$$- \operatorname{Cov}\Bigg(R, \int_Y^\infty \big(1 - F_{R|Y}(R|\tilde{Y})\big) \times \underbrace{\big(u''\big(WR^f + \bar{Y} + Y\big) - u''\big(WR^f + \bar{Y} + \tilde{Y}\big)\big)}_{<0} \, dF_Y(\tilde{Y})\Bigg)$$

$$< 0.$$

This inequality always holds if $\operatorname{Cov}(R, Y) = 0$ because $R$ and $1 - F_{R|Y}(R|\tilde{Y})$ covary negatively and $\int_Y^\infty (u''(WR^f + \bar{Y} + Y) - u''(WR^f + \bar{Y} + \tilde{Y}))\, dF_Y(\tilde{Y}) < 0$. If $\operatorname{Cov}(R, Y) > 0$, then $\operatorname{Cov}(R, \int_Y^\infty (1 - F_{R|Y}(R|\tilde{Y}))(u''(WR^f + Y) - u''(WR^f + \tilde{Y}))\, dF_Y(\tilde{Y}))$ is again positive (if $Y$ is high and $\tilde{Y} > Y$, then $(1 - F_{R|Y}(R|\tilde{Y}))$ is small (as $R$ is high) and $u''(WR^f + Y) - u''(WR^f + \tilde{Y})$ is negative but closer to zero as $u''(x) < 0$ and increasing in $x$, i.e., $u'''(x) > 0$).

## APPENDIX D: EXTENSIONS AND QUANTITATIVE IMPLICATIONS

### D.1. *Signals About the Market*

Instantaneous utility is either prospective news utility over the realization of $R$ or prospective news utility over the signal $\hat{r} = r + \varepsilon$ with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ with the function $\nu$ given by $\nu(x) = \eta x$ for $x > 0$ and $\nu(x) = \eta\lambda x$. Prospective news utility over the signal is given by

$$\gamma\beta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu\big(u\big(e^{\hat{r}-y}\big) - u\big(e^{x-y}\big)\big)\, dF_{r+\varepsilon}(x)\, dF_\varepsilon(y),$$

because the agent separates uncertainty that has been realized, represented by $\hat{r}$ and $F_{r+\varepsilon}(x)$, from uncertainty that has not been realized, represented by $F_\varepsilon(y)$. The agent's expected news utility from looking up the return conditional on the signal $\hat{r}$ is given by

$$\gamma\beta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu\big(u\big(e^{\hat{r}-y}\big) - u\big(e^{x}\big)\big)\, dF_r(x)\, dF_\varepsilon(y).$$

Thus, the agent expects more favorable news utility over the return when having received a favorable signal. As can be easily seen, expected news utility from checking the return is always less than news utility from knowing merely the signal. The reason is simply that the agent expects to experience news utility, which is negative on average, over both the signal $\hat{r}$ and the error $\varepsilon$. Thus, he always prefers to ignore his return when prospective news utility is concerned. But, the difference between the two is smaller when the agent receives a more favorable signal. The reason is that the expected news disutility from $\varepsilon$ is less high up on the concave utility curve, that is, when $\hat{r}$ is high.

### D.2. *Numerical Solution of the Portfolio-Choice Model*

The agent lives for $t = \{1, \dots, T\}$ periods and is endowed with initial wealth $W_1$. Each period, the agent optimally decides how much to consume $C_t$ out of his cash-on-hand $X_t$ and how to invest the remaining funds $A_t = X_t - C_t$. The agent has access to a risk-free investment with return $R^f$ and a risky investment with i.i.d. return $R_t$. The risky investment's share is denoted by $\alpha_t$ such that the portfolio return in period $t$ is given by

$R_t^p = R^f + \alpha_{t-1}(R_t - R^f)$. Additionally, the agent receives labor income in each period $t$ given by $Y_t = P_t N_t^T = P_t e^{s_t^T}$ with $s_t^T \sim N(0, \sigma_Y^2)$ stochastic up until retirement $T - Ret$ and $P_t$ a deterministic profile. Accordingly, the agent's maximization problem in each period $t$ is given by

$$\max_{C_t}\left\{u(C_t) + n\big(C_t, F_{C_t}^{t-1}\big) + \gamma\sum_{\tau=1}^{T-t}\beta^\tau \boldsymbol{n}\big(F_{C_{t+\tau}}^{t,t-1}\big) + E_t\left[\sum_{\tau=1}^{T-t}\beta^\tau U_{t+\tau}\right]\right\}$$

subject to the budget constraint

$$X_t = (X_{t-1} - C_{t-1})R_t^p + Y_t = A_{t-1}\big(R^f + \alpha_{t-1}\big(R_t - R^f\big)\big) + P_t e^{s_t^T}.$$

I solve the model by numerical backward induction. The maximization problem in any period $t$ is characterized by the following first-order condition:

$$u'(C_t) = \frac{\Psi_t' + \gamma\Phi_t'\big(\eta F_{A_t}^{t-1}(A_t) + \eta\lambda\big(1 - F_{A_t}^{t-1}(A_t)\big)\big)}{1 + \eta F_{C_t}^{t-1}(C_t) + \eta\lambda\big(1 - F_{C_t}^{t-1}(C_t)\big)}$$

with

$$\Phi_t' = \beta E_t\left[R_{t+1}^p \frac{\partial C_{t+1}}{\partial X_{t+1}} u'(C_{t+1}) + R_{t+1}^p\left(1 - \frac{\partial C_{t+1}}{\partial X_{t+1}}\right)\Phi_{t+1}'\right]$$

and

$$\begin{aligned}\Psi_t' = \beta E_t\Bigg[&R_{t+1}^p \frac{\partial C_{t+1}}{\partial X_{t+1}} u'(C_{t+1}) \\ &+ \eta(\lambda - 1)\int_{C_{t+1}}^{\infty}\left(R_{t+1}^p \frac{\partial C_{t+1}}{\partial X_{t+1}} u'(C_{t+1}) - c\right) dF^t_{R_{t+1}^p \frac{\partial C_{t+1}}{\partial X_{t+1}} u'(C_{t+1})}(c) \\ &+ \gamma\eta(\lambda - 1)\int_{A_{t+1}}^{\infty}\left(R_{t+1}^p \frac{\partial A_{t+1}}{\partial X_{t+1}}\Phi_{t+1}' - x\right) dF^t_{R_{t+1}^p \frac{\partial A_{t+1}}{\partial X_{t+1}}\Phi_{t+1}'}(x) \\ &+ R_{t+1}^p\left(1 - \frac{\partial C_{t+1}}{\partial X_{t+1}}\right)\Psi_{t+1}'\Bigg].\end{aligned}$$

Note that I denote $\Psi_t = \beta E_t[\sum_{\tau=0}^{T-t}\beta^\tau U_{t+1+\tau}]$, $\Phi_t = \beta E_t[\sum_{\tau=0}^{T-t}\beta^\tau u(C_{t+1+\tau})]$, $\Psi_t' = \frac{\partial \Psi_t}{\partial A_t}$, and $\Phi_t' = \frac{\partial \Phi_t}{\partial A_t}$. In turn, the optimal portfolio share can be determined by the following first-order condition:

$$\gamma\frac{\partial \Phi_t}{\partial \alpha_t}\big(\eta F_{A_t}^{t-1}(A_t) + \eta\lambda\big(1 - F_{A_t}^{t-1}(A_t)\big)\big) + \frac{\partial \Psi_t}{\partial \alpha_t} = 0$$

or equivalently by choosing $\alpha_t$ to maximize $\gamma\sum_{\tau=1}^{T-t}\beta^\tau \boldsymbol{n}(F_{C_{t+\tau}}^{t,t-1}) + \Psi_t$ (maximizing the transformed function $u^{-1}(\cdot)$ yields more robust results). I use a couple of tricks that considerably improve the numerical solution and thus make a structural estimation feasible. In particular, I assume quadrature weights and nodes for the numerical integration, use the endogenous-grid method of Carroll (2001), precompute the marginal value functions in the numerical solution procedures for the optimal consumption and portfolio share function to then use linear interpolation, and transform the marginal value functions $\frac{\partial \Phi_t, \Psi_t}{\partial A_t}$

and $\frac{\partial \Phi_t, \Psi_t}{\partial \alpha_t}$ by $A_t^{\theta}$ and $A_t^{1-\theta}$, respectively, to reduce nonlinearities. More details on the numerical backward induction solution of a news-utility life-cycle model are provided in Pagel (2017).

### D.3. *Structural Estimation Details*

*Data*

The SCF data are a statistical survey of incomes, balance sheets, pensions, and other demographic characteristics of families in the United States, sponsored by the Federal Reserve Board in cooperation with the Treasury Department. The SCF was conducted in six survey waves from 1992 to 2007 but did not survey households consecutively. Therefore, I construct a pseudo-panel by averaging participation and shares of all households at each age following the risk-free and risky-asset definitions of Flavin and Nakagawa (2008). In addition to the age effects of interest, the data are contaminated by potential time and cohort effects, which constitutes an identification problem as time minus age equals cohort. In the portfolio-choice literature, it is standard to solve the identification problem by acknowledging age and time effects (as tradable and non-tradable wealth vary with age and contemporaneous stock-market happenings are likely to affect participation and shares) while omitting cohort effects (Campbell and Viceira (2002)). In contrast, in the consumption literature, it is standard to omit time effects but acknowledge cohort effects (Gourinchas and Parker (2002)). I employ a new method, recently invented by Schulhofer-Wohl (2013), that solves the age-time-and-cohort identification problem with the mere assumption that age, time, and cohort effects are linearly related. I first estimate a pooled OLS model, whereby I jointly control for age, time, and cohort effects and identify the model with a random assumption about its trend. If I would estimate this arbitrary trend together with the structural parameters, I would obtain consistent estimates using data that are uncontaminated by time and cohort effects. This application of Schulhofer-Wohl (2013) to household portfolio data is important, as portfolio profiles are highly dependent on which assumptions the identification is based on (as was made clear by Ameriks and Zeldes (2004)). In practice, the trend is estimated close to 1, so I fix it at 1 to not bias the empirical results in favor of my model. Therefore, I match a hump-shaped profile consistent with the evidence in Ameriks and Zeldes (2004).

The CEX data are a survey of the consumption expenditures, incomes, balance sheets, and other demographic characteristics of families in the United States, sponsored by the Bureau of Labor Statistics. The news-utility agent's consumption profile is hump-shaped and roughly matches the empirical consumption profile estimated from CEX data, displayed in Figure 6.

*Calibration*

For an annual investment period, the literature suggests fairly tight ranges for the parameters of the log-normal return; I match these by estimating $\hat{\mu} - \hat{r}^f = 6.33\%$, $\hat{\sigma} = 19.4\%$, and the log risk-free rate, $\hat{r}^f = 0.86\%$, using value-weighted CRSP return data. Moreover, the life-cycle consumption literature suggests fairly tight ranges for the parameters determining stochastic labor income: labor income is log-normal, characterized by shocks with variance $\sigma_Y$, a probability of unemployment $p$, and a trend $G$ that I roughly match by estimating $\hat{\sigma}_Y = 0.1$, $p = 1\%$, and $\hat{G}_t$ from the SCF data. I abstract from permanent income shocks because Carroll (2001, 1997), and more recently Heathcote, Perri, and Violante (2010), argued that household income processes are well approxi-

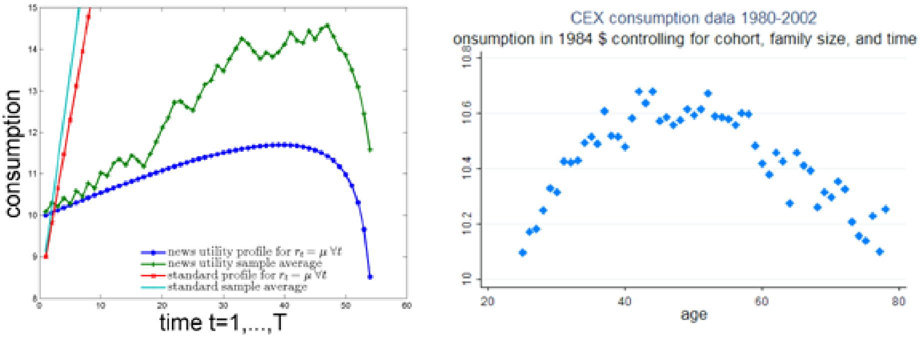


FIGURE 6.—Life-cycle consumption profiles for a news-utility or standard agent for whom $r_t = \mu\ \forall t$, sample average profiles of 10,000 news-utility or standard agents, and empirical consumption profile. The model's period length is calibrated to an annual frequency and the environmental parameters can be found in Table I. The preference parameter values are $\beta = 0.98$, $\eta = 1$, $\eta(\lambda - 1) = 2$, and $\gamma = 0.8$.

mated by a deterministic trend and a transitory shock. Labor income is correlated with the risky asset with a coefficient of approximately 0.2 following Viceira (2001) among others. Moreover, because 25 is chosen as the beginning of life by Gourinchas and Parker (2002), I choose $\hat{Ret} = 11$ and $\hat{T} = 54$ in accordance with the average retirement age in the United States according to the OECD and the average life expectancy in the United States according to the UN.

*Estimation—Identification*

Are the empirical life-cycle participation and portfolio shares profiles able to identify the preference parameters? I am interested in five preference parameters, namely, $\beta$, $\theta$, $\eta$, $\lambda$, and $\gamma$. As shown, both participation and portfolio shares are determined by the first-order condition $\gamma \frac{\partial \Phi_t}{\partial \alpha_t}(\eta F_{A_t}^{t-1}(A_t) + \eta\lambda(1 - F_{A_t}^{t-1}(A_t))) + \frac{\partial \Psi_t}{\partial \alpha_t} = 0$ for which I observe the average of all households. $\frac{\partial \Phi_t}{\partial \alpha_t}$ represents future marginal consumption utility, as in the standard model, and is determined by $\beta$ and $\theta$, which can be separately identified in a finite-horizon model. $\frac{\partial \psi_t}{\partial \alpha_t}$ represents future marginal consumption and news utility and is thus determined by something akin to $\eta(\lambda - 1)$. $\eta F_{A_t}^{t-1}(A_t) + \eta\lambda(1 - F_{A_t}^{t-1}(A_t))$ represents the weighted sum of the cumulative distribution function of savings, $A_t$, of which merely the average determined by $\eta 0.5(1 + \lambda)$ is observed. Thus, I have two equations in two unknowns and can separately identify $\eta$ and $\lambda$. Finally, $\gamma$ enters the first-order condition distinctly from all other parameters, and I conclude that the model is identified, which can also be verified by deriving the Jacobian that has full rank.

*Estimation—Methods of Simulated Moments Procedure*

I focus on matching portfolio shares rather than participation, because the optimization procedure would prioritize them anyhow as they have lower variances than participation. I can then use participation as an out-of-sample test. The empirical profile is the average of the portfolio shares at each age $a \in [1, T]$ across all household observations $i$. More precisely, it is $\bar{\alpha}_a = \frac{1}{n_a}\sum_{i=1}^{n_a} \bar{\alpha}_{i,a}$ with $\bar{\alpha}_{i,a}$ being the household $i$'s portfolio share at age $a$ of which $n_a$ are observed. The theoretical population analogue to $\bar{\alpha}_a$ is denoted by $\alpha_a(\theta, \Xi)$ and the simulated approximation is denoted by $\hat{\alpha}_a(\boldsymbol{\theta}, \Xi)$. Moreover, I define

$$\hat{g}(\boldsymbol{\theta}, \Xi) = \hat{\alpha}_a(\boldsymbol{\theta}, \Xi) - \bar{\alpha}_a.$$

In turn, if $\boldsymbol{\theta}_0$ and $\Xi_0$ are the true parameter vectors, the procedure's moment conditions imply that $E[g(\theta_0, \Xi_0)] = 0$. Let $W$ denote a positive definite weighting matrix. In that case,

$$q(\boldsymbol{\theta}, \Xi) = \hat{g}(\boldsymbol{\theta}, \Xi) W^{-1} \hat{g}(\boldsymbol{\theta}, \Xi)'$$

is the weighted sum of squared deviations of the simulated moments from their corresponding empirical moments. I assume that $W$ is a robust weighting matrix rather than the optimal weighting matrix to avoid small-sample bias. More precisely, I assume that $W$ corresponds to the inverse of the variance-covariance matrix of each point of $\bar{\alpha}_a$, which I denote by $\Omega_g^{-1}$ and consistently estimate from the sample data. Taking $\hat{\Xi}$ as given, I minimize $q(\boldsymbol{\theta}, \hat{\Xi})$ with respect to $\boldsymbol{\theta}$ to obtain $\hat{\boldsymbol{\theta}}$ the consistent estimator of $\boldsymbol{\theta}$ that is asymptotically normally distributed with standard errors:

$$\Omega_\theta = \left(G_\theta' W G_\theta\right)^{-1} G_\theta' W \left[\Omega_g + \Omega_g^s + G_\Xi \Omega_\Xi G_\Xi'\right] W G_\theta \left(G_\theta' W G_\theta\right)^{-1}.$$

Here, $G_\theta$ and $G_\Xi$ denote the derivatives of the moment functions $\frac{\partial g(\boldsymbol{\theta}_0, \Xi_0)}{\partial \boldsymbol{\theta}}$ and $\frac{\partial g(\boldsymbol{\theta}_0, \Xi_0)}{\partial \Xi}$, $\Omega_g$ denotes the variance-covariance matrix of the second-stage moments, as above, that corresponds to $E[g(\boldsymbol{\theta}_0, \Xi_0) g(\boldsymbol{\theta}_0, \Xi_0)']$, and $\Omega_g^s = \frac{n_a}{n_s} \Omega_g$ denotes the sample correction with $n_s$ being the number of simulated observations at each age $a$. As $\Omega_g$, I can estimate $\Omega_\Xi$ directly and consistently from sample data. For the minimization, I employ a Nelder–Mead algorithm. For the standard errors, I numerically estimate the gradient of the moment function at its optimum. If I omit the first-stage correction and simulation correction, the expression becomes $\Omega_\theta = (G_\theta' \Omega_g^{-1} G_\theta)^{-1}$. Thus, I can test for overidentification by comparing $\hat{g}(\hat{\boldsymbol{\theta}}, \hat{\Xi})$ to a chi-squared distribution with $T - 5$ degrees of freedom.

### D.4. *Structural Estimation Results*

*Fitted Values*

It can be seen in Figure 7 that the portfolio share function during retirement is more smooth than the function before retirement. The reason is that after retirement the model can be solved analytically, whereas before retirement the model has to be solved numerically on a grid with a number of grid points that make a structural estimation feasible;

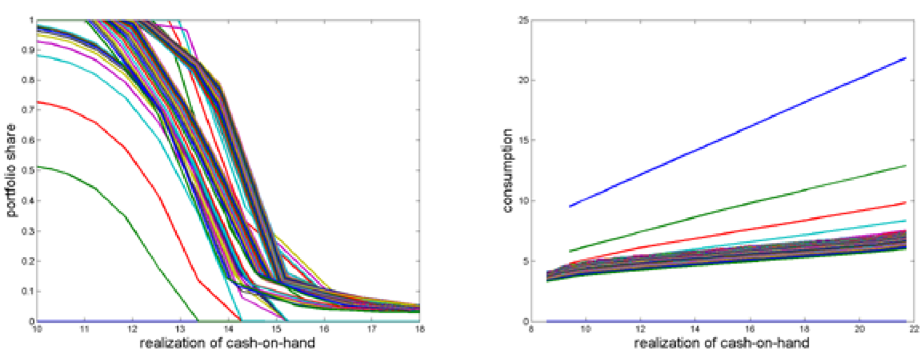


FIGURE 7.—Portfolio share and consumption functions during and before retirement for each $t \in \{1, \dots, T\}$, that is, at each age 25 to 78. The environmental and preference parameters equal the estimates in Table I and the initial savings level equals $A_{t-1} = 11.9$ and the initial portfolio share equals $\alpha_{t-1} = 0.79$ for each $t \in \{1, \dots, T\}$.

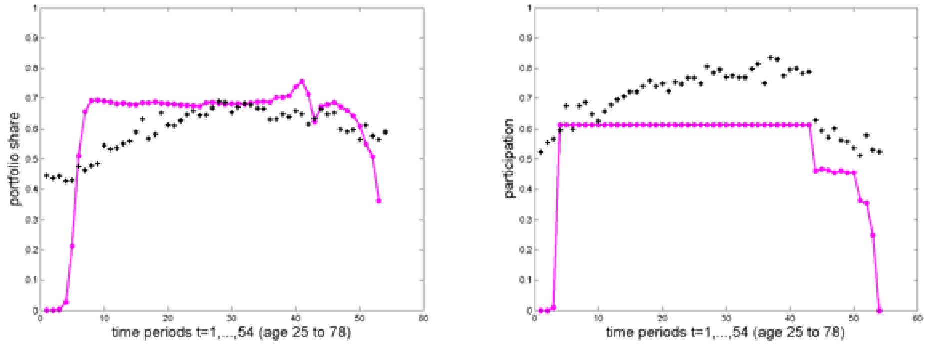


FIGURE 8.—Average simulated portfolio share and participation life-cycle profiles for 10,000 agents. The environmental and preference parameters equal the estimates in Table I.

I chose a grid with 20 points, but a more finely-spaced grid does not materially alter the quantitative results. The kinks in the portfolio share functions before retirement stem from the 3-point grid for labor income that I can also choose to be more fine (holding the dispersion of labor income constant) without affecting the functions quantitatively. In any case, the solution before retirement is reasonably smooth and is not subject to any instabilities; therefore, I conclude that numerical inaccuracies are not a problem. Moreover, the participation and consumption functions are very smooth during and before retirement. In Figure 8, it can be seen that portfolio share and participation are hump-shaped and fit the empirical patterns.

*Attitudes Over Consumption and Wealth Gambles*

I calculate the required gain G for a range of losses L to make each agent indifferent between accepting or rejecting a 50/50 either win G or lose L gamble at a wealth level of 300,000 as in Rabin (2001) and Chetty and Szeidl (2007). First, I want to match risk attitudes towards gambles regarding immediate consumption, which are determined solely by $\eta$ and $\lambda$ because it can be reasonably assumed that utility over immediate consumption is linear. Thus, $\eta \approx 1$ and $\lambda \in [2, 3]$ are suggested by the laboratory evidence on loss aversion over immediate consumption, most importantly, the endowment effect literature. $\eta \approx 1$ and $\lambda \approx 3$ imply that the equivalent Kahneman and Tversky (1979) coefficient of loss aversion is around 2. The reason is that classical prospect theory effectively consists of news utility only, whereas the news-utility agent experiences consumption and news utility (and consumption utility works in favor of any small-scale gamble).[33] Furthermore, it can be seen in Table II that news-utility preferences generate reasonable attitudes towards small and large gambles over wealth. In contrast, the standard agent is risk neutral for small and medium stakes and risk averse for large stakes only. I elicit the agents' risk attitudes towards wealth gambles by assuming that each of them is presented with the gamble after all consumption in that period has taken place. The news-utility agent will

[33]For illustration, I borrow a concrete example from Kahneman, Knetsch, and Thaler (1990), in which the authors distributed a good (mugs or pens) to half of their subjects and asked those who received the good about their willingness to accept (WTA) and those who did not receive it about their willingness to pay (WTP) if they traded the good. The median WTA is \$5.25, whereas the median WTP is \$2.75. Accordingly, I infer $(1+\eta)u(mug) = (1+\eta\lambda)2.25$ and $(1+\eta\lambda)u(mug) = (1+\eta)5.25$, which imply that $\lambda \approx 3$ when $\eta \approx 1$. I obtain a similar result for the pen experiment.

TABLE II

RISK ATTITUDES OVER SMALL AND LARGE CONSUMPTION AND WEALTH GAMBLES[a]

| Loss (L) | Standard | News-Utility Contemp. | News-Utility Prospective |
|---|---|---|---|
| 10 | 10 | 20 | 19 |
| 200 | 200 | 400 | 378 |
| 1000 | 1003 | 2010 | 1898 |
| 5000 | 5085 | 10,256 | 9677 |
| 50,000 | 36,301 | 132,000 | 123,340 |
| 100,000 | 120,380 | 375,000 | 345,260 |

[a] For each loss L, required gain G to make each agent indifferent between accepting and rejecting a 50–50 gamble win G or lose L at a wealth level of 300,000.

only experience prospective news utility over the wealth gambles' outcomes that is determined by the prospective news discount factor $\gamma$. $\gamma$ implies attitudes towards intertemporal consumption tradeoffs that are similar to those implied by a hyperbolic-discounting parameter of the same value. Empirical estimates for the hyperbolic-discounting parameter in a variety of contexts typically range between 0.7 and 0.8 (e.g., Laibson, Repetto, and Tobacman (2012)). Thus the experimental and field evidence on peoples' attitudes towards intertemporal consumption tradeoffs suggests that $\gamma \approx 0.8$ and $\beta \approx 1$ are plausible estimates. Moreover, my estimates match the parameter values obtained by a structural estimation of a life-cycle consumption model in Pagel (2017).

## REFERENCES


AMERIKS, J., AND S. ZELDES (2004): "How Do Household Portfolio Shares Vary With age?" Working Paper, Columbia Business School. [24]

CAMPBELL, J., AND L. VICEIRA (2002): *Strategic Asset Allocation*. Oxford: Oxford University Press. [1,24]

CARROLL, C. (2001): "Theoretical Foundations of Buffer Stock Saving," Working Paper, Johns Hopkins University. [23,24]

CARROLL, C. D. (1997): "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112, 1–56. [24]

CHETTY, R., AND A. SZEIDL (2007): "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics*, 122 (2), 831–877. [27]

FLAVIN, M., AND S. NAKAGAWA (2008): "A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence," *American Economic Review*, 98 (1), 474–495. [24]

GOURINCHAS, P.-O., AND J. A. PARKER (2002): "Consumption Over the Life Cycle," *Econometrica*, 70, 47–89. [24,25]

HEATHCOTE, J., F. PERRI, AND G. VIOLANTE (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006," *Review of Economic Dynamics*, 13 (1), 15–51. [24]

KAHNEMAN, D., J. KNETSCH, AND R. THALER (1990): "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98 (6), 1325–1348. [27]

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–291. [27]

LAIBSON, D., A. REPETTO, AND J. TOBACMAN (2012): "Estimating Discount Functions With Consumption Choices Over the Lifecycle," Working Paper. [28]

PAGEL, M. (2017): "Expectations-Based Reference-Dependent Life-Cycle Consumption," *Review of Economic Studies*, 84 (2), 885–934. [8,24,28]

RABIN, M. (2001): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68 (5), 1281–1292. [27]

SCHULHOFER-WOHL, S. (2013): "The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models," Minneapolis Fed Working Paper 707. [24]

VICEIRA, L. (2001): "Optimal Portfolio Choice for Long-Horizon Investors," *Journal of Finance*, 56 (2), 433–470. [25]

*Co-editor Itzhak Gilboa handled this manuscript.*

*Manuscript received 18 May, 2016; final version accepted 27 December, 2017; available online 27 December, 2017.*