Optimal Asset Allocation with Heterogeneous Persistent Shocks and Myopic and Intertemporal Hedging Demand

Domenica Di Virgilio
Bank of Slovenia

Fulvio Ortu
Università Bocconi & IGIER

Federico Severino
USI Lugano & Università Bocconi

Claudio Tebaldi*
Università Bocconi, IGIER & CAREFIN

April 2018

This is a pre-print of an article published in the book *Behavioral finance: the coming of age*. The final authenticated version is available online at:
`https://doi.org/10.1142/11248`.

Abstract

There is wide evidence that financial time series are the outcome of the superposition of processes with heterogeneous frequencies. This is true, in particular, for market return. Indeed, log market return can be decomposed into uncorrelated components that explain the reaction to shocks with different persistence. The instrument that allows us to do so is the Extended Wold Decomposition of Ortu, Severino, Tamoni, and Tebaldi (2017). In this paper, we construct portfolios of these components in order to maximize the utility of an agent with a fixed investment horizon. In particular, we build upon Campbell and Viceira (1999) solution of the optimal consumption-investment problem with Epstein-Zin utility, using a rebalancing interval of 2^J periods. It turns out that the optimal asset allocation involves all the persistent components of market log return up to scale J. Such components play a fundamental role in characterizing both the myopic and the intertemporal hedging demand. Moreover, the optimal policy prescribes an increasing allocation on more persistent securities when the investor's relative risk aversion rises. Finally, portfolio reallocation every 2^J periods is consistent

*We thank Carlo Favero, Lars Peter Hansen, Stavros Panageas, Pietro Reggiani, Andrea Tamoni and Pietro Veronesi for valuable insights. We also thank participants at University of Chicago, at MIPP Millenium Institute of Santiago (2016), at 8^{th} Summer Finance Conference, Arison School of Business, IDC Herzliya, Israel (2016), at The Second Israel Behavioral Finance Conference, Tel Aviv-Yaffo Academic College (2017), at 11^{th} CFE at University of London (2017) and at the 9^{th} Vienna International Conference on Mathematical Modelling (2018) for helpful comments and suggestions. Any errors or omissions are the sole responsibility of the authors.

with rational inattention. Indeed, observing assets value is costly and transaction costs make occasional rebalancing optimal.

JEL classification: G11, C32, C61.
Keywords: optimal portfolio, Epstein-Zin preferences, multiple horizons, persistence of returns, shocks heterogeneity.

1 Motivations

Economic phenomena are complex objects of study. In fact, every variable is the result of the aggregation of a multiplicity of factors that occur often unexpectedly and last for an undetermined amount of time. The superposition of these factors lets cyclicalities - as well as other meaningful patterns - arise in economic and financial time series. These issues motivated the thriving literature of business cycles detection by spectral analysis of the frequency domain. Remarkable examples of the filtering approach are provided by Stock and Watson (1999) and Baxter and King (1999).

In the realm of finance, the coexistence of sources of randomness with various durations is associated with diverse trading horizons. Investors with specific timelines require compensation from the exposures to shocks with precise frequencies. For example, long-term traders are much more concerned of political cycles than short-term investors, who are likely to be more keen on temporary mispricings. The intuition of risk premia anchored to different horizons is coherent with the Heterogeneous Market Hypothesis of Müller, Dacorogna, Davé, Olsen, Pictet, and von Weizsäcker (1997), which turned out to be fruitful for the analysis of stochastic volatility. Moreover, estimating the proper duration of shocks is crucial for long-run implications on economic dynamics, as described by Bansal and Yaron (2004) and Ortu, Tamoni, and Tebaldi (2013), among others.

From a practical perspective, exploiting different frequencies of asset returns has proven to be profitable for investment. This evidence comes from portfolio strategies that rely, for instance, on FED meetings calendar. Indeed, building on the result of Lucca and Moench (2015), Cieslak, Morse, and Vissing-Jorgensen (2018) show that market returns display biweekly cycles around FOMC meetings, which occur every six weeks. Going long (short) on the market index according to even (odd) weeks outperformed the buy-and-hold strategy by $2,4$ times from 1994 to 2016. Despite this convincing empirical evidence, portfolio theory is essentially silent on investment strategies that could optimally include the compensation to shocks with heterogeneous durations. Some tentative formalizations of multifrequency trading are provided very recently by Chaudhuri and Lo (2016) and Crouzet, Dew-Becker, and Nathanson (2017).

On the other hand, the financial literature on intertemporal asset pricing theory is vast. Among the fundamental works on multiperiod asset allocation we can quote, for instance, Brennan, Schwartz, and Lagnado (1997) and Barberis (2000). In addition, most of achievements about optimal portfolio policies in the presence of stochastic returns flowed into

Campbell (1993) and Campbell and Viceira (1999), who solved the optimal consumption-portfolio problem of an Epstein-Zin type investor by assuming autoregressive returns. Differently from one-period settings, the intertemporal dimension of the problem affects the optimal capital allocation, providing the distinction between *myopic demand* and *hedging demand*. The latter is particularly important for medium-term traders because it incorporates the agent's reaction to expected future return streams.

More generally, Campbell, Chan, and Viceira (2003) provide the methodology to implement a multivariate strategic asset allocation. However, in these models the investor is not equipped with the necessary tools to fully exploit the multi-horizon nature of market returns.

In this paper we build a portfolio optimization framework in which the agent can optimally trade assets that are associated with the heterogeneous levels of persistence of market returns. Specifically, we implement a persistence-based asset allocation in a special Campbell, Chan, and Viceira (2003) setting.

The starting point of our construction is the decomposition of market returns (in excess of a risk-free rate) into the sum of uncorrelated components associated with specific investment horizons. To achieve this goal, we apply the Extended Wold Decomposition, or *persistence-based* decomposition, of Ortu, Severino, Tamoni, and Tebaldi (2017) to the stationary time series of log excess returns. In addition, we assume that such components correspond to the returns of risky securities traded in the market. An Epstein-Zin type investor, then, maximizes her utility by optimally trading these assets. We finally illustrate the different implications of persistence heterogeneity on myopic and hedging demands.

Our market components are reminiscent of factors employed in Capital Asset Pricing Models. Since the early work of Ross (1976), factor investing has been pervasive in empirical asset pricing. Original portfolios of Fama and French (1992) three factors, based on size and value, have been enhanced with Carhart (1997) momentum factor and with a taxonomy of other stylized portfolios, as quality factors of Fama and French (2015) and Hou, Xue, and Zhang (2015) and lucky factors by Harvey and Liu (2017). Moreover, practitioners successfully contributed to this approach out of the academic world: see, for example, the overviews by Bender, Briand, Melas, and Subramanian (2013) and Podkaminer (2013). Nevertheless, differently from factors known in the financial literature, our market components involve specific investment horizons and capture shocks with frequencies associated with the horizon under consideration, building a bridge between multiperiod asset allocation and the filtering approach.

Finally, our investor is supposed to allocate her wealth every 2^J periods, where J is a reference level of persistence. The agent's choice of rebalancing her portfolio after 2^J periods of inaction is compatible with the theory of optimal inattention. Indeed, observing the value of the investment portfolio may be costly and transaction costs may induce infrequent adjustments. See, for instance, Abel, Eberly, and Panageas (2013) and Peng and Xiong (2006).

1.1 Summary of results

Given a zero-mean weakly stationary time series $\mathbf{x} = \{x_t\}_t$, the Classical Wold Decomposition allows us to write any x_t as an infinite sum of uncorrelated innovations:

$$x_t = \sum_{k=0}^{+\infty} \alpha_h \varepsilon_{t-h},$$

where $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_t$ is a unit variance white noise and α_h are the so-called impulse response functions. The Extended Wold Decomposition introduced by Ortu, Severino, Tamoni, and Tebaldi (2017), instead, decomposes x_t into uncorrelated *persistent components* $x_t^{(j)}$ associated with specific time scales j such that

$$x_t = \sum_{j=1}^{+\infty} x_t^{(j)}, \qquad x_t^{(j)} = \sum_{k=0}^{+\infty} \beta_k^{(j)} \varepsilon_{t-k2^j}^{(j)}.$$

Here each *detail process* $\boldsymbol{\varepsilon}^{(j)} = \left\{\varepsilon_t^{(j)}\right\}_t$ is an $MA(2^j - 1)$ with respect to the fundamental innovations of $\mathbf{x}$ and $\beta_k^{(j)}$ is the *multiscale impulse response* associated with scale j and time-shift $k2^j$. Moreover, fixed a maximum scale J, it is possible to write the orthogonal decomposition

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + m_t^{(J)},$$

where $m_t^{(J)}$ constitutes a *residual component*. With a small abuse of notation we denote $x_t^{(J+1)} = m_t^{(J)}$.

The derivation of the Extended Wold Decomposition stems from the application of an isometric low-pass filter. Therefore, the innovations $\boldsymbol{\varepsilon}^{(j)}$ concentrate on lower and lower frequencies as scale j increases. The whole construction is, however, developed in the time domain and so each detail process is associated with a precise time horizon. The same is true for the related persistent component. For instance, on daily basis, scale $j = 1$ involves two-day shocks, scale $j = 2$ four-day innovations (that may proxy weekly shocks) and so on.

We apply the previous decomposition to the process of market (excess) log returns associated, for instance, to S&P 500 index. We consider an Epstein-Zin investor that chooses how to distribute her wealth among $J + 1$ risky assets and a riskless security, with a periodic rebalancing of 2^J time units. Log returns of these risky assets are supposed to mimic the persistent components $x_t^{(1)}, \ldots, x_t^{(J+1)}$ of market log returns. Moreover, we assume that each $x_t^{(j)}$ follows an $AR(1)$ process on its own scale:

$$x_{t+2^j}^{(j)} = \mu_j (1 - \phi_j) + \phi_j x_t^{(j)} + s_j \varepsilon_{t+2^j}^{(j)}.$$

Finally, the fundamental innovations ε_t are i.i.d. and distributed as standard normal.

In a simplified version of the model, by denoting portfolio loadings by $\pi_t(j)$, the return over 2^J periods is

$$R_{p,t+2^J} = \sum_{j=1}^{J+1} \pi_t(j) e^{x_{t+2^J}(j)} + \left(1 - \sum_{j=1}^{J+1} \pi_t(j)\right) e^{2^J r_f}.$$

The previous assumptions allow the vector of returns $\mathbf{z}_t = \left[x_t^{(1)}, \ldots, x_t^{(J+1)}\right]'$ to follow the VAR dynamics

$$\mathbf{z}_{t+2^J} = \Phi_0 + \Phi \mathbf{z}_t + \mathbf{v}_{t+2^J},$$

where $\mathbf{v}_t$ is a multivariate white noise on the time grid $t - k2^J$ with $k \in \mathbb{Z}$.

The agent has recursive preferences but her utility depends on the current consumption and the certainty equivalent associated with the utility 2^J periods ahead:

$$\max_{\{C_t, \boldsymbol{\pi}_t\}_{t=k2^J}} \quad U_t = \left((1-\beta) C_t^{(1-\gamma)/\theta} + \beta E_t \left[U_{t+2^J}^{1-\gamma} \right]^{1/\theta} \right)^{\theta/(1-\gamma)}$$

$$sub \qquad W_{t+2^J} = R_{p,t+2^J} \left(W_t - C_t \right),$$

where $0 < \beta < 1$ is the preference discount factor, $\gamma > 0$ is the coefficient of relative risk aversion, ψ denotes the intertemporal elasticity of substitution (IES) and $\theta = (1-\gamma) / \left(1 - \psi^{-1}\right)$. Consumption C_t and wealth W_t are scalars, while the vector $\boldsymbol{\pi}_t$ contains the portfolio weights associated with the $J+1$ securities into consideration.

The previous VAR representation of returns allows us to embed our optimal consumption-investment problem into Campbell, Chan, and Viceira (2003) portfolio theory. In particular, we exploit the affine guess

$$\boldsymbol{\pi}_t = A_0 + A_1 \mathbf{z}_t$$

and, after approximating log return, budget constraint and Euler equation, we determine the optimal asset allocation, which is driven by myopic and hedging motives. The myopic demand is induced exclusively by current risk premia while the intertemporal hedging demand is driven by the ability of present risk premia to predict future changes in investment opportunities. This feature is captured by the covariance between current excess returns and future consumption-wealth ratio. In particular, we get

$$\boldsymbol{\pi}_t = A_{0,myopic} + A_{0,hedging} + \left(A_{1,myopic} + A_{1,hedging}\right) \mathbf{z}_t.$$

Although the investor's horizon is 2^J, the optimal capital allocation involves all the components of market returns, not only the one at scale J.

The orthogonality of the Extended Wold Decomposition ensures that the *myopic part* of $\pi_t(j)$ depends only on $x_t^{(j)}$. Moreover, if $\gamma = 1$ - because, for instance, the investor has logarithmic utility - the *hedging part* of $\boldsymbol{\pi}_t$ disappears. Then, for a myopic investor the weight $\pi_t(j)$ depends only on $x_t^{(j)}$. Instead, if $\gamma \neq 1$, the resulting capital allocation on the j-th component of market returns depends also on the other components. In particular, $\pi_t(j)$ depends on $x_t^{(i)}$with $i \neq j$ through the term $A_{1,hedging}\mathbf{z}_t$. Hence, the share $\pi_t(j)$ invested in the component $x_t^{(j)}$ depends on the components at scales $i \neq j$ just for hedging purposes.

We corroborate our analysis by estimating optimal weights of a portfolio investing in persistent components of S&P 500 index for different levels of risk aversions. If $\gamma = 1$ the investor is fully myopic and the weights are all equal across scales. When γ increases, the investment diversifies within persistent assets and portfolio loadings to high scales become prominent.

The article is organized as follows. The next section summarizes the Extended Wold Decomposition of Ortu, Severino, Tamoni, and Tebaldi (2017) in general terms. Section 3 shows how to apply the decomposition in order to properly set up a multeperiod asset allocation problem. Section 4 derives the approximated optimal solution of the consumption-portfolio problem and discusses the results, highlighting the role of persistence in constructing optimal loadings. Then, Section 5 is devoted to an empirical illustration of the methodology, while Section 6 concludes. The Appendix includes some complements of the theory.

2 Disentangling heterogeneous levels of persistence

As anticipated in Subsection 1.1, Ortu, Severino, Tamoni, and Tebaldi (2017) provide the methodology to decompose any zero-mean weakly stationary purely non-deterministic time series $\mathbf{x} = \{x_t\}_{t\in\mathbb{Z}}$ into uncorrelated persistent components linearly generated by shocks with increasing durations.

By the Classical Wold Decomposition Theorem it is possible to define a unit variance white noise process $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$ such that any realization x_t can be expressed as

$$x_t = \sum_{k=0}^{+\infty} \alpha_h \varepsilon_{t-h}, \qquad \sum_{h=0}^{+\infty} \alpha_h^2 < +\infty.$$

Each α_h is the impulse response function of x_t with respect to a shock occurred h periods before. We refer to $\boldsymbol{\varepsilon}$ as the process of fundamental innovations of $\mathbf{x}$.

The Extended Wold Decomposition, instead, dismantles the calendar-time occurrence of fundamental innovations by decomposing the same x_t into uncorrelated variables $x_t^{(j)}$ associated with specific levels of persistence.

Theorem 1 *Let $\mathbf{x}$ be a zero-mean, weakly stationary purely non-deterministic stochastic process. Then x_t decomposes as*[1]

$$x_t = \sum_{j=1}^{+\infty} \sum_{k=0}^{+\infty} \beta_k^{(j)} \varepsilon_{t-k2^j}^{(j)},$$

where

i) for any fixed $j \in \mathbb{N}$, the detail process $\boldsymbol{\varepsilon}^{(j)} = \left\{ \varepsilon_t^{(j)} \right\}_{t \in \mathbb{Z}}$ *is an $MA(2^j - 1)$ with respect to the classical Wold innovations of $\mathbf{x}$:*

$$\varepsilon_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-2^{j-1}-i} \right)$$

and $\left\{ \varepsilon_{t-k2^j}^{(j)} \right\}_{k \in \mathbb{Z}}$ is a unit variance white noise;

ii) for any $j \in \mathbb{N}$, $k \in \mathbb{N}_0$, the coefficients $\beta_k^{(j)}$ are unique and they satisfy

$$\beta_k^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+i} - \sum_{i=0}^{2^{j-1}-1} \alpha_{k2^j+2^{j-1}+i} \right),$$

hence they do not depend on t and $\sum_{k=0}^{\infty} \left(\beta_k^{(j)} \right)^2 < +\infty$ for any $j \in \mathbb{N}$;

iii) letting

$$x_t^{(j)} = \sum_{k=0}^{+\infty} \beta_k^{(j)} \varepsilon_{t-k2^j}^{(j)},$$

then, for any $j, l \in \mathbb{N}, p, q, t \in \mathbb{Z}$, $\mathbb{E}\left[x_{t-p}^{(j)} x_{t-q}^{(l)} \right]$ depends at most on $j, l, p - q$. Moreover,

$$\mathbb{E}\left[x_{t-m2^j}^{(j)} x_{t-n2^l}^{(l)} \right] = 0 \qquad \forall j \neq l, \quad \forall m, n \in \mathbb{N}_0, \quad \forall t \in \mathbb{Z}.$$

We call $x_t^{(j)}$ *persistent component* at scale j and we refer to $\beta_k^{(j)}$ as *multiscale impulse response* associated with level of persistence j and time-shift $k2^j$. Moreover, fixed a maximum scale J, it is possible to write the orthogonal decomposition

$$x_t = \sum_{j=1}^{J} x_t^{(j)} + m_{x,t}^{(J)},$$

[1] Throughout the paper, the equalities between random variables are in the L^2-norm.

where the *residual component* $m_{x,t}^{(J)}$ satisfies

$$m_{x,t}^{(J)} = \sum_{k=0}^{+\infty} \gamma_k^{(J)} \left(\frac{1}{\sqrt{2^J}} \sum_{i=0}^{2^J-1} \varepsilon_{t-k2^J-i} \right), \qquad \gamma_k^{(J)} = \frac{1}{\sqrt{2^J}} \sum_{i=0}^{2^J-1} \alpha_{k2^J+i}.$$

The support of details $\left\{\varepsilon_{t-k2^j}^{(j)}\right\}_{k\in\mathbb{Z}}$ employed in the decomposition is sparser and sparser as the scale raises, conveying the intuition of increasing duration (strengthen by the higher order of MA). Hence, scale-specific impulse responses $\beta_k^{(j)}$ capture the sensitivity of x_t with respect to underlying shocks with heterogeneous durations related, for instance, to short-, medium- or long-term economic factors. As a result, due to the dichotomic nature of the construction, each persistent component $x_t^{(j)}$ may be associated with shocks of 2^j periods. For example, on quarterly basis, scale $j = 1$ collects the impact of semiannual innovations, scale $j = 2$ involves annual disturbances and so on.

To illustrate the Extended Wold Decomposition, we plot in Figure 1 the multiscale impulse responses and the persistent components of a simulated weakly stationary $AR(2)$ process defined by $x_t = 1.2x_{t-1} - 0.3x_{t-2} + \varepsilon_t$, where we employ a Gaussian white noise. After demeaning the time series, we follow the estimation procedure described in Section 5. We observe that components at different time scales may feature contrasting behaviours that are not recognizable in the aggregate process. Indeed, impulse responses on the first time scale provide evidence of mean reversion, while the second scale reveals some degree of delayed overreaction. Scale 3 instead features the usual pattern of autoregressive impulse responses.

Finally, note that the Extended Wold Decomposition stems from the fundamental innovations $\boldsymbol{\varepsilon}$ of the original time series $\mathbf{x}$. However, the same decomposition holds in case $\boldsymbol{\varepsilon}$ is any white noise process that allows a MA representation of $\mathbf{x}$. In any case, the orthogonality of components induces a decomposition of the variance of x_t into the sum of variances at each scale. Hence, it is possible to assess the relative importance of each persistent component on the whole process.

The reverse construction is also feasible. Indeed, taken as given a white noise process $\boldsymbol{\varepsilon}$ and the dynamics $x_t^{(j)}$ on different scales, Ortu, Severino, Tamoni, and Tebaldi (2017) show how to rebuild the aggregated time series obtained by summing up such $x_t^{(j)}$.

Theorem 2 *Let $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$ be a unit variance white noise process. For any $j \in \mathbb{N}$, define the detail process $\boldsymbol{\varepsilon^{(j)}} = \left\{\varepsilon_t^{(j)}\right\}_{t\in\mathbb{Z}}$ as*

$$\varepsilon_t^{(j)} = \frac{1}{\sqrt{2^j}} \left(\sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-i} - \sum_{i=0}^{2^{j-1}-1} \varepsilon_{t-2^{j-1}-i} \right)$$

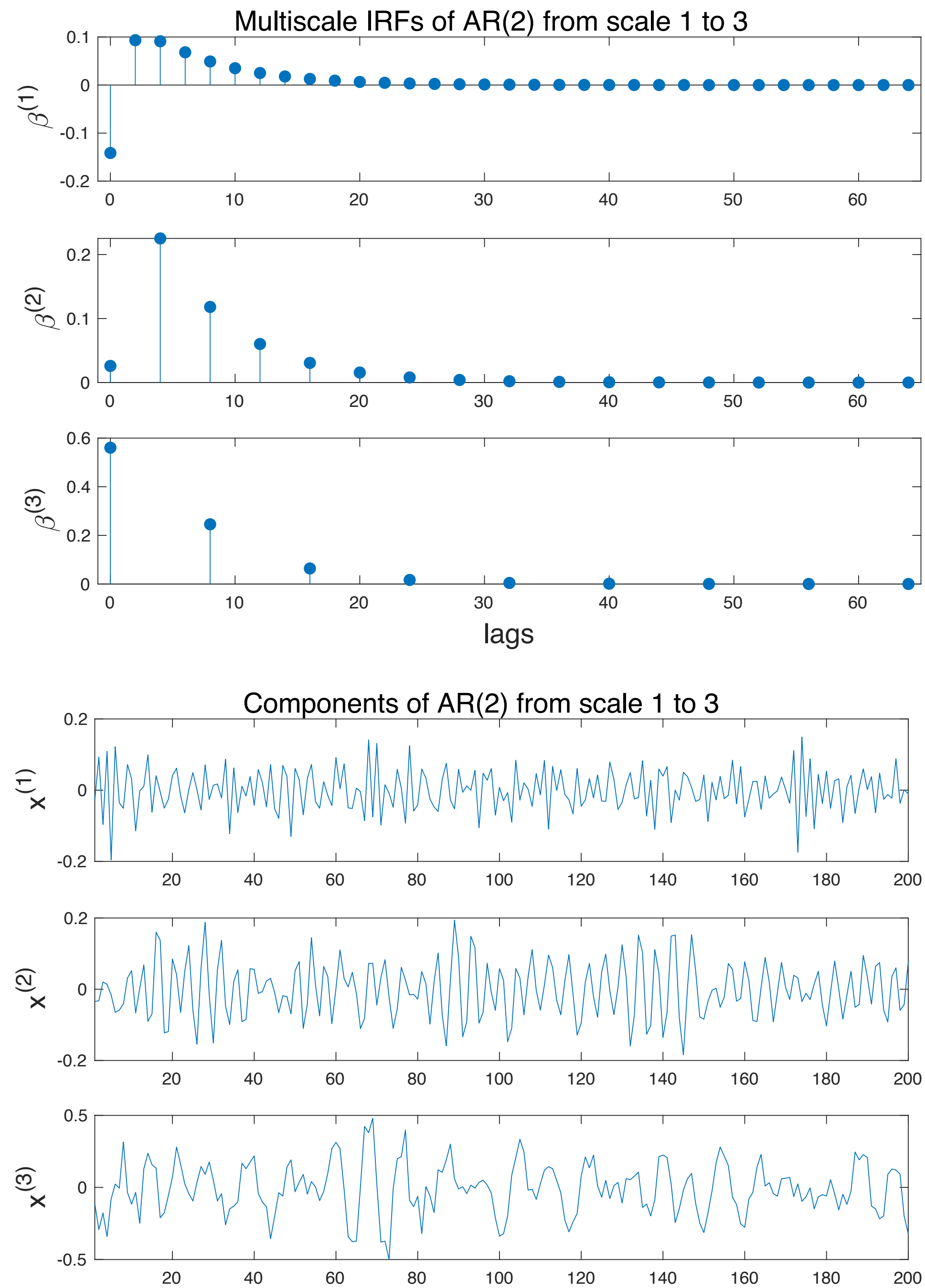

Figure 1: Estimated multiscale impulse response functions and persistent components of a demeaned weakly stationary $AR(2)$ process defined by $x_t = 1.2x_{t-1} - 0.3x_{t-2} + \varepsilon_t$.

and consider a stochastic process $\mathbf{x^{(j)}} = \left\{x_t^{(j)}\right\}_{t\in\mathbb{Z}}$ *such that*

$$x_t^{(j)} = \sum_{k=0}^{+\infty} \beta_k^{(j)} \varepsilon_{t-k2^j}^{(j)}, \qquad \sum_{j=1}^{+\infty}\sum_{k=0}^{+\infty} \left(\beta_k^{(j)}\right)^2 < +\infty.$$

Then, the process $\mathbf{x} = \{x_t\}_{t\in\mathbb{Z}}$ *defined by*

$$x_t = \sum_{j=1}^{+\infty} x_t^{(j)}$$

is zero-mean, weakly stationary purely non-deterministic and

$$x_t = \sum_{h=0}^{+\infty} \alpha_h \varepsilon_{t-h},$$

where, for any $h \in \mathbb{N}_0$*,*

$$\alpha_h = \sum_{j=1}^{+\infty} \frac{1}{\sqrt{2^j}} \beta^{(j)}_{\left\lfloor \frac{h}{2^j} \right\rfloor} \chi^{(j)}(h)$$

and

$$\chi^{(j)}(h) = \begin{cases} -1 & \text{if} \quad 2^j \left\lfloor \frac{h}{2^j} \right\rfloor \in \left\{h - 2^j + 1, \ldots, h - 2^{j-1}\right\}, \\ 1 & \text{if} \quad 2^j \left\lfloor \frac{h}{2^j} \right\rfloor \in \left\{h - 2^{j-1} + 1, \ldots, h\right\}. \end{cases}$$

This result is particularly fruitful in the context of the paper because it allows to directly specify the dynamics on different scales and to aggregate each subseries into a stationary process as a second step. In our financial application, for example, we will assume autoregressive processes at any persistence level.

3 Persistence-based capital allocation

3.1 Decomposition of market return and risk-free rate

We consider as weakly stationary process the one of market log return $\mathbf{r} = \{r_t\}_{t\in\mathbb{Z}}$.[2] The fundamental innovations of $\mathbf{r}$, denoted by $\boldsymbol{\varepsilon} = \{\varepsilon_t\}_{t\in\mathbb{Z}}$, generate the whole information structure. By fixing a maximum scale J, r_t decomposes as

$$r_t = \mu_r + \sum_{j=1}^{J} \tilde{r}_t(j) + \tilde{m}_{r,t}(J),$$

[2] Lower-case letters denote variables in logs.

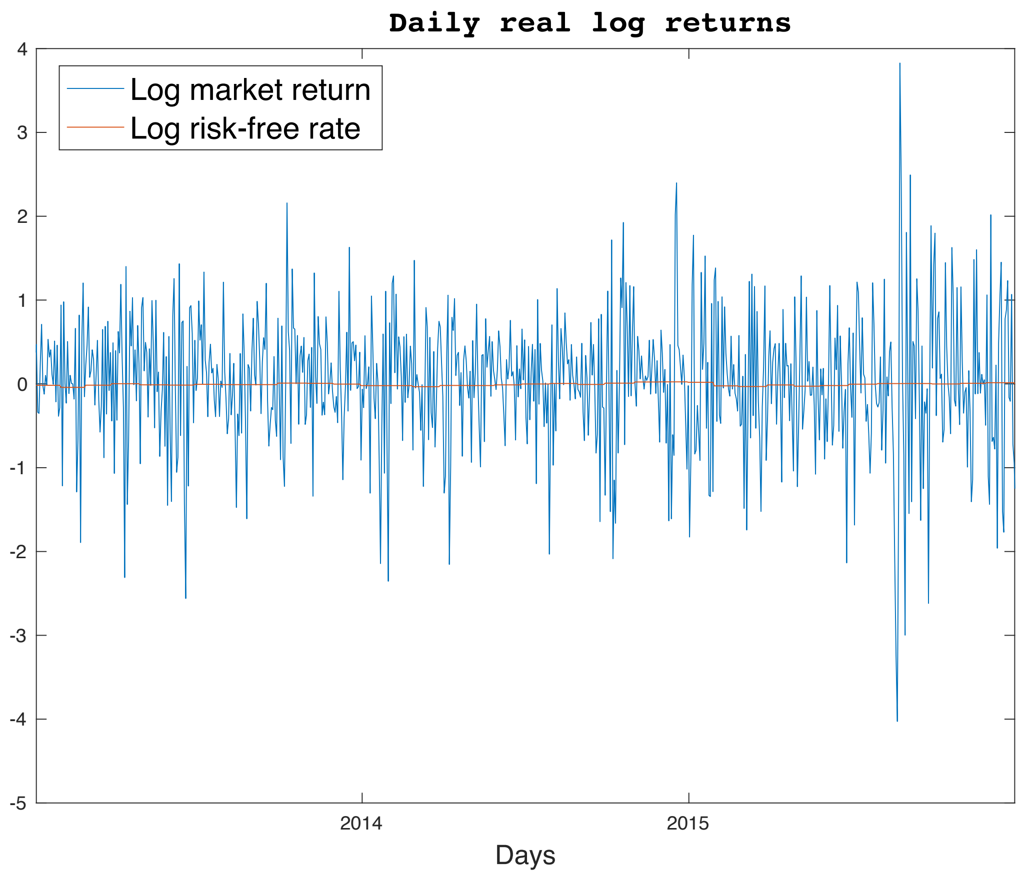

Figure 2: Realizations of daily log market and log risk-free return from January 2, 2013 to December 31, 2015. Market returns are taken from S&P 500 index, while the risk-free rate comes from three-month Treasury Bills.

where μ_r is the expectation of r_t, each $\tilde{r}_t(j)$ is the persistent component of demeaned r_t at scale j and $\tilde{m}_{r,t}(J)$ is the related residual component.[3] By defining

$$r_t(j) = \tilde{r}_t(j) + \frac{\mu_r}{2^j}, \qquad\qquad m_{r,t}(J) = \tilde{m}_{r,t}(J) + \frac{\mu_r}{2^J},$$

we find the decomposition

$$r_t = \sum_{j=1}^{J} r_t(j) + r_t(J+1), \tag{1}$$

where we name $r_t(J+1) = m_{r,t}(J)$ with a little abuse of notation. We plot in Figure 2 the daily realizations from January 2013 to December 2015 while Figure 3 depicts the variables $\tilde{r}_t(j)$ for scales $j = 1, \ldots, 8$.

In Figure 4 we show the variance explained by each persistent component of daily log returns. Due to the orthogonality of the Extended Wold Decomposition, the variance of r_t turns out to be the sum of the variances of each $r_t^{(j)}$. In this example, the first eight scales explain, together, $99,6\%$ of the variance of r_t. As we can observe, most of the weight is associated to low time scales, hence the process is mainly sensitive to short-term disturbances.

Similarly, we deal with the persistent components of a short-term bond traded in the market. To ease the terminology, we refer to this security as a riskless asset. We denote the related log risk-free rate process by $\mathbf{f} = \{f_t\}_{t\in\mathbb{Z}}$ and its first moment by μ_f. In addition, we assume that $\mathbf{f}$ is a weakly stationary process driven by the same innovations $\boldsymbol{\varepsilon}$ that generate $\mathbf{r}$. With respect to these disturbances, the persistent components of demeaned f_t are the variables $\tilde{f}_t(j)$, for $j = 1, \ldots, J$, while the residual component at scale J is $\tilde{m}_{f,t}(J)$. The persistence-based decomposition of $\mathbf{f}$ is, then,

$$f_t = \mu_f + \sum_{j=1}^{J} \tilde{f}_t(j) + \tilde{m}_{f,t}(J).$$

By defining

$$f_t(j) = \tilde{f}_t(j) + \frac{\mu_f}{2^j}, \qquad\qquad m_{f,t}(J) = \tilde{m}_{f,t}(J) + \frac{\mu_f}{2^J},$$

we deduce the orthogonal decomposition of the risk-free rate

$$f_t = \sum_{j=1}^{J} f_t(j) + f_t(J+1), \tag{2}$$

where $f_t(J+1) = m_{f,t}(J)$.

Now we consider market portfolio log return in excess of log short-term rate, i.e. the process $\mathbf{x} = \{x_t\}_{t\in\mathbb{Z}}$ such that $x_t = r_t - f_t$. This time series is still driven by the shocks $\boldsymbol{\varepsilon} =$

[3] The notation $\tilde{r}_t(j)$, instead of $\tilde{r}_t^{(j)}$ is convenient for the vector structures that we will build later.

Figure 3: Estimated persistent components of daily log market return (from S&P 500 index) from January 2, 2013 to December 31, 2015 at scales $j = 1, \ldots, 8$.

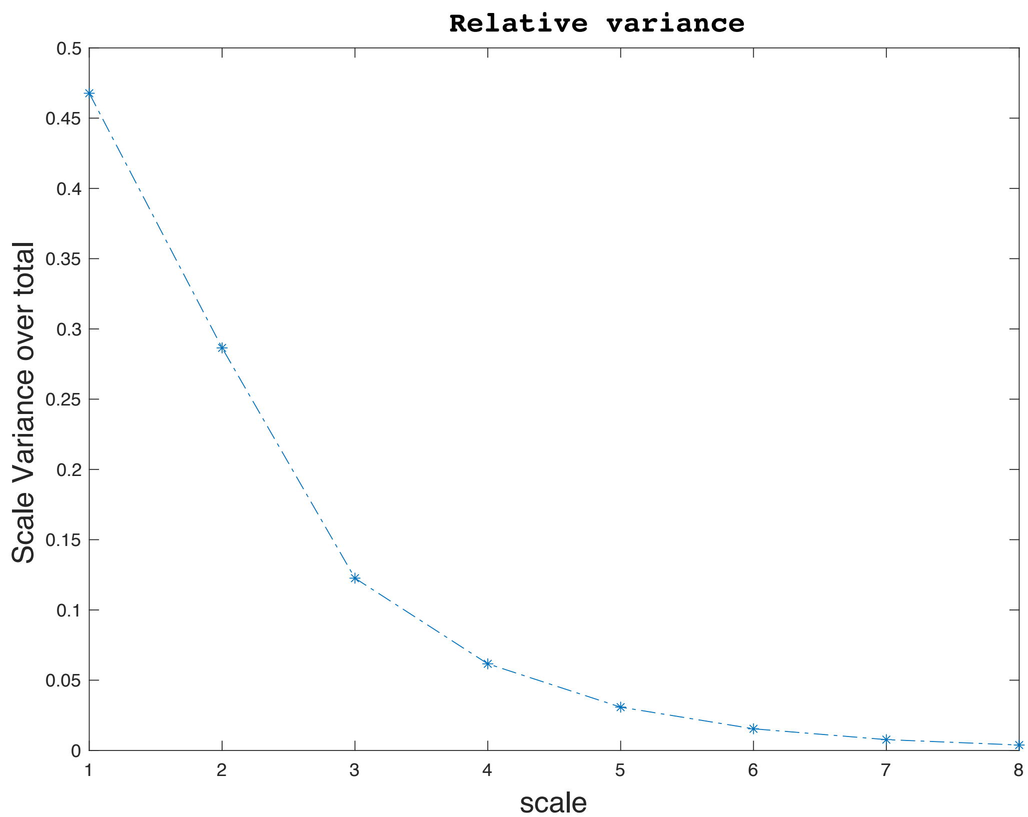

Figure 4: Variance explained by each scale (from $j = 1$ to $j = 8$) over total variance. The whole time series of daily log market returns from S&P 500 index is considered (from January 4, 1954 to December 30, 2016).

$\{\varepsilon_t\}_{t\in\mathbb{Z}}$, therefore the previous persistence-based decompositions of $\mathbf{r}$ and $\mathbf{f}$ immediately provide

$$x_t = \sum_{j=1}^{J} x_t(j) + x_t(J+1), \tag{3}$$

where $x_t(J+1) = r_t(J+1) - f_t(J+1)$ and, for all $j = 1, \ldots, J$,

$$x_t(j) = r_t(j) - f_t(j) = \tilde{r}_t(j) - \tilde{f}_t(j) + \frac{\mu_r - \mu_f}{2^j}.$$

3.2 Autoregressive persistent components

In order to exploit the persistent components of market log return and log risk-free rate in a factor-based asset allocation, we need to make assumptions on their dynamics. Hence, we adapt the assumptions of multiperiod strategic asset allocation by Campbell and Viceira (1999) and Campbell, Chan, and Viceira (2003). More specifically, we build our model on the following.

($A1$) The investor chooses how to allocate her wealth among $J+1$ *risky* assets and $J+1$ *riskless* securities. Log returns $r_t(1), \ldots, r_t(J)$ of first J risky assets mimic the persistent components of market log return, while the log return of the last risky security reproduces the residual component $r_t(J+1)$. The same reasoning applies to the first J risk-free assets, whose log returns $f_t(1), \ldots, f_t(J)$ mimic the persistent components of the log interest rate of a short-term bond, while the log return of the last safe security reproduces the residual $f_t(J+1)$.

($A2$) Investor's preferences are described by a recursive utility à la Epstein-Zin but the agent consumes and reallocates her portfolio over time-intervals of length 2^J.

($A3$) For $j = 1, \ldots, J$, each persistent component $f_t(j)$ of log risk-free rate follows an $AR(1)$ process on its own scale:

$$f_{t+2^j}(j) = \mu_{f,j}\,(1-\psi_j) + \psi_j f_t(j) + s_{f,j}\varepsilon^{(j)}_{t+2^j} \qquad \text{for } j = 1, .., J$$

where $\mu_{f,j} = \mu_f/2^j$ and $s_{f,j} = \beta^{(j)}_{f,0}$ with $\beta^{(j)}_{f,0}$ indicating the first multiscale impulse response at scale j of $\mathbf{f}$.[4]

[4]Equivalently,

$$f_{t+2^j}(j) - \mu_{f,j} = \psi_{f,j}\,(f_t(j) - \mu_{f,j}) + s_{f,j}\varepsilon^{(j)}_{t+2^j} \quad \text{for } j = 1, .., J.$$

In particular, if ψ_j is null, $f_t(j)$ is a white noise process and the usefulness of the persistence-based decomposition lies in the definition of $s_{f,j} = \beta^{(j)}_{f,0}$. This choice is crucial in order to avoid a variance mismatching between the scales and the aggregated data. This observation holds for the following dynamics, too.

In a similar way, $f_t(J+1)$ is an $AR(1)$ process:

$$f_{t+2^J}(J+1) = \mu_{f,J+1}\left(1-\psi_{J+1}\right) + \psi_{J+1} f_t(J+1) + s_{f,J+1}\left(\frac{1}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}\right),$$

where $\mu_{f,J+1} = \mu_{f,J}$ and $s_{f,J+1} = \gamma_{f,0}^{(J)}$ is the first coefficient of $f_t(J+1)$.[5]

Persistent components of market log return in excess of log short-term rate follow an $AR(1)$ on their own scale:

$$x_{t+2^j}(j) = \mu_j\left(1-\phi_j\right) + \phi_j x_t(j) + s_j \varepsilon_{t+2^j}^{(j)} \qquad \text{for } j = 1, .., J$$

with $\mu_j = \left(\mu_r - \mu_f\right)/2^j$ and $s_j \neq 0$. Here we set $s_j = \beta_0^{(j)}$, which denotes the first multiscale impulse response at scale j of $\mathbf{x}$.[6]

Similarly, $x_t(J+1)$ is an $AR(1)$:

$$x_{t+2^J}(J+1) = \mu_{J+1}\left(1-\phi_{J+1}\right) + \phi_{J+1} x_t(J+1) + s_{J+1}\left(\frac{1}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}\right),$$

where $\mu_{J+1} = \mu_J$ and $s_{J+1} = \gamma_0^{(J)}$ is the first coefficient of the residual component of $x_t(J+1)$.[7]

($A4$) Fundamental innovations ε_t are i.i.d. and normally distributed with zero mean and unit variance. Hence, innovations in the market components $\varepsilon_{t+2^j}^{(j)}$ are also i.i.d. and distributed as standard Gaussian. The same shocks drive log market return, log risk-free rate and excess log return.

[5] Equivalently,

$$f_{t+2^J}(J+1) - \mu_{f,J+1} = \psi_{J+1}\left(f_t(J+1) - \mu_{f,J+1}\right) + s_{f,J+1}\left(\frac{1}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}\right).$$

[6] An equivalent writing is

$$x_{t+2^j}(j) - \mu_j = \phi_j\left(x_t(j) - \mu_j\right) + s_j \varepsilon_{t+2^j}^{(j)} \quad \text{for } j = 1, .., J.$$

[7] Equivalently,

$$x_{t+2^J}(J+1) - \mu_{J+1} = \phi_{J+1}\left(x_t(J+1) - \mu_{J+1}\right) + s_{J+1}\left(\frac{1}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}\right).$$

The investor's portfolio selection problem consists in deciding the share $\pi_t(j)$ of her current savings to be invested in the asset mimicking the j-th component of market log return for $j = 1, 2, ...J$ and the share $\pi_t(J+1)$ to be invested in the security reproducing the residual component $r_t(J+1)$. The remaining portfolio weights are invested in the assets mimicking the components of the log risk-free rate. Precisely, the amount $\frac{1}{J+1} - \pi_t(j)$ is invested in the j-th of these assets. The resulting portfolio, whose loadings sum up to 1, generalizes the standard portfolio in which the share $1 - \sum_{j=1}^{J+1} \pi_t(j)$ is allocated in the whole riskless asset. Indeed, this special case is obtained when $J = 0$.

According to Assumption $(A3)$, scale-specific excess log market returns are supposed to be autoregressive. Empirically, once we have estimated multiscale impulse responses, details $\varepsilon_t^{(j)}$ and persistent components $r_t(j)$, we set $s_j = \beta_0^{(j)}$ and we estimate ϕ_j by OLS. The result depicted in Figure 5 is encouraging. The AR modelling does not preclude $r_t(j)$ to be a white noise process. Also in this special case, however, the role of the persistence-based decomposition is not negligible because it prescribes the choice of s_j. The same reasoning applies to log risk-free rate.

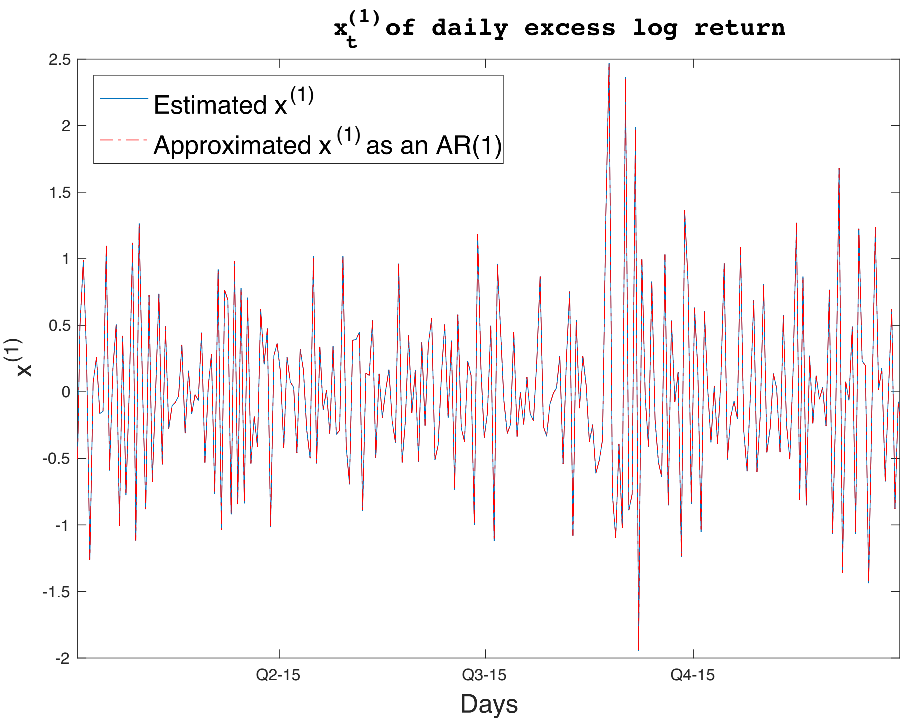

Figure 5: Comparison between the component $x_t(1)$ estimated from the time series of daily excess log return and the $AR(1)$ approximation of Assumption $(A3)$. In this case, $s_1 = 0.669$ and $\phi_1 = -0.0357$. The plot covers the period between January 2015 and December 2015.

By Assumption $(A3)$, for any scale $j = 1, \ldots, J$, the dynamics of $f_t(j)$ and $x_t(j)$ can be rewritten as

$$f_{t+2^J}(j) = \mu_{f,j}\left(1 - \psi_j^{2^{J-j}}\right) + \psi_j^{2^{J-j}} f_t(j) + s_{f,j} \sum_{i=0}^{2^{J-j}-1} \psi_j^i \varepsilon_{t+2^J-i2^j}^{(j)},$$

$$x_{t+2^J}(j) = \mu_j\left(1 - \phi_j^{2^{J-j}}\right) + \phi_j^{2^{J-j}} x_t(j) + s_j \sum_{i=0}^{2^{J-j}-1} \phi_j^i \varepsilon_{t+2^J-i2^j}^{(j)}$$

by backward substitution. Indeed, the autoregressive assumptions allow us to express the realizations at time $t + 2^J$ in terms of the ones at t for any scale under scrutiny.

Let $\mathbf{f}_t$ be the vector that collects the $J+1$ components of log risk-free rate and $\mathbf{x}_t$ the vector of components of excess log return. We define the vector $\mathbf{z}_t$ of length $2J+2$ by stacking $\mathbf{f}_t$ and $\mathbf{x}_t$, that is

$$\mathbf{f}_t = \begin{bmatrix} \vdots \\ f_t(j) \\ \vdots \end{bmatrix}, \qquad \mathbf{x}_t = \begin{bmatrix} \vdots \\ x_t(j) \\ \vdots \end{bmatrix}, \qquad \mathbf{z}_t = \begin{bmatrix} \mathbf{f}_t \\ \mathbf{x}_t \end{bmatrix}.$$

Thanks to previous relations, it is possible to express the dynamics of $\mathbf{z}_t$ through the VAR representation:

$$\mathbf{z}_{t+2^J} = \Phi_0 + \Phi \mathbf{z}_t + \mathbf{v}_{t+2^J},$$

where

- Φ_0 is a vector of length $2J+2$, with entries

$$\begin{array}{lcll} \Phi_0(j) & = & \mu_{f,j}\left(1 - \psi_j^{2^{J-j}}\right) & \text{for} \quad j = 1, \ldots, J, \\ \Phi_0(J+1) & = & \mu_{f,J+1}\left(1 - \psi_{J+1}\right), & \\ \Phi_0(j) & = & \mu_{j-J-1}\left(1 - \phi_{j-J-1}^{2^{2J-j+1}}\right) & \text{for} \quad j = J+2, \ldots, 2J+1, \\ \Phi_0(2J+2) & = & \mu_{J+1}\left(1 - \phi_{J+1}\right). & \end{array}$$

- Φ is a $(2J+2) \times (2J+2)$ diagonal matrix, whose general diagonal term is

$$\begin{array}{lcll} \Phi(j,j) & = & \psi_j^{2^{J-j}} & \text{for} \quad j = 1, \ldots, J, \\ \Phi(J+1, J+1) & = & \psi_{J+1}, & \\ \Phi(j,j) & = & \phi_{j-J-1}^{2^{2J-j+1}} & \text{for} \quad j = J+2, \ldots, 2J+1, \\ \Phi(2J+2, 2J+2) & = & \phi_{J+1}. & \end{array}$$

- $\mathbf{v}_{t+2^J}$ is a vector of length $2J+2$ such that

$$\begin{array}{lcll}
\mathbf{v}_{t+2^J}(j) & = & s_{f,j}\sum_{i=0}^{2^{J-j}-1}\psi_j^i\varepsilon_{t+2^J-i2^j}^{(j)} & \text{for} \quad j=1,\dots,J, \\
\mathbf{v}_{t+2^J}(J+1) & = & \frac{s_{f,J+1}}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}, & \\
\mathbf{v}_{t+2^J}(j) & = & s_{j-J-1}\sum_{i=0}^{2^{2J-j+1}-1}\phi_{j-J-1}^i\varepsilon_{t+2^J-i2^{j-J-1}}^{(j-J-1)} & \text{for} \quad j=J+2,\dots,2J+1, \\
\mathbf{v}_{t+2^J}(2J+2) & = & \frac{s_{J+1}}{\sqrt{2^J}}\sum_{i=0}^{2^J-1}\varepsilon_{t+2^J-i}. &
\end{array}$$

Let $\Sigma_{\mathbf{v}}$ denote the covariance matrix of $\mathbf{v}_{t+2^J}$ conditional on time t. Then, $\Sigma_{\mathbf{v}}$ is a block matrix

$$\Sigma_{\mathbf{v}} = var_t\left(\mathbf{v}_{t+2^J}\right) = \begin{bmatrix} \Sigma_{\mathbf{f}} & \Sigma_{\mathbf{fx}} \\ \Sigma_{\mathbf{xf}} & \Sigma_{\mathbf{x}} \end{bmatrix},$$

where $\Sigma_{\mathbf{f}}$ and $\Sigma_{\mathbf{x}}$ are the conditional covariance matrices of $\mathbf{f}_{t+2^J}$ and $\mathbf{x}_{t+2^J}$ respectively and

$$\Sigma_{\mathbf{fx}}(p,q) = cov_t\left(f_{t+2^J}(p), x_{t+2^J}(q)\right), \qquad p,q=1,\dots,J+1.$$

By the properties of details $\varepsilon_t^{(j)}$, $\Sigma_{\mathbf{f}}$, $\Sigma_{\mathbf{x}}$ and $\Sigma_{\mathbf{fx}}$ are diagonal matrices. Hence, $\Sigma_{\mathbf{fx}} = \Sigma_{\mathbf{xf}}$. We denote by $\boldsymbol{\sigma}_{\mathbf{f}}^2$ the $(J+1)$-vector which collects all diagonal terms of $\Sigma_{\mathbf{f}}$, that is

$$\boldsymbol{\sigma}_{\mathbf{f}}^2(j) = s_{f,j}^2\frac{1-\psi_j^{2^{J-j+1}}}{1-\psi_j^2} \qquad \text{for} \quad j=1,\dots,J$$

and $\boldsymbol{\sigma}_{\mathbf{f}}^2(J+1) = s_{f,J+1}^2$. Similarly, $\boldsymbol{\sigma}_{\mathbf{x}}^2$ is the vector that contains the diagonal terms of $\Sigma_{\mathbf{x}}$:

$$\boldsymbol{\sigma}_{\mathbf{x}}^2(j) = s_j^2\frac{1-\phi_j^{2^{J-j+1}}}{1-\phi_j^2} \qquad \text{for} \quad j=1,\dots,J$$

and $\boldsymbol{\sigma}_{\mathbf{x}}^2(J+1) = s_{J+1}^2$. Finally, $\boldsymbol{\sigma}_{\mathbf{fx}}$ includes the diagonal terms of $\Sigma_{\mathbf{fx}}$:

$$\boldsymbol{\sigma}_{\mathbf{fx}}(j) = s_{f,j}s_j\frac{1-(\psi_j\phi_j)^{2^{J-j}}}{1-\psi_j\phi_j} \qquad \text{for} \quad j=1,\dots,J$$

with $\boldsymbol{\sigma}_{\mathbf{fx}}(J+1) = s_{f,J+1}s_{J+1}$.

Moreover, the orthogonality properties of details $\varepsilon_t^{(j)}$ guarantee that $\mathbf{v}_t$ defines a *multivariate white noise* on the time grid $t-k2^J$ with $k\in\mathbb{Z}$. This feature, ensured by the Extended Wold Decomposition, is crucial for translating one-period rebalancing to 2^J-period reallocation. Although persistent components contemporaneously capture economic innovations with heterogeneous durations, their orthogonality every 2^J time units allows us to properly set a multiperiod investment problem.

3.3 Optimization problem and approximation of portfolio return

We assume in (A2) that the investor has preferences captured by a utility function à la Epstein-Zin (see Epstein and Zin (1989), Epstein and Zin (1991) and Weil (1989)), though it displays the following peculiarity: her utility depends on current consumption and the certainty equivalent associated with the utility 2^J periods ahead, that is

$$\max_{\{C_t,\boldsymbol{\pi}_t\}_{t=k2^J}} \quad U_t = \left((1-\beta)C_t^{(1-\gamma)/\theta} + \beta E_t\left[U_{t+2^J}^{1-\gamma}\right]^{1/\theta}\right)^{\theta/(1-\gamma)}$$

$$sub \qquad W_{t+2^J} = R_{p,t:t+2^J}\left(W_t - C_t\right),$$

where $0 < \beta < 1$ is the preference discount factor, $\gamma > 0$ is the coefficient of relative risk aversion, ψ denotes the intertemporal elasticity of substitution (IES) and $\theta = (1-\gamma)/\left(1-\psi^{-1}\right)$. Consumption C_t is a scalar, while the vector $\boldsymbol{\pi}_t$ of $\mathbb{R}^{J+1}$ contains the portfolio weights associated with the $J+1$ risky securities into consideration. Moreover, W_t is the investor's wealth and $R_{p,t:t+2^J}$ denotes the return on investor's portfolio in 2^J periods. Similarly, the 2^J-period log returns of the risky and the risk-free assets - as well as the related excess log returns - are denoted by

$$r_{t:t+2^J}(j), \qquad f_{t:t+2^J}(j), \qquad x_{t:t+2^J}(j).$$

Following the previous description, the portfolio return is given by

$$R_{p,t:t+2^J} = e^{r_{p,t:t+2^J}} = \sum_{j=1}^{J+1}\left(\frac{1}{J+1} - \pi_t(j)\right)e^{f_{t:t+2^J}(j)} + \sum_{j=1}^{J+1}\pi_t(j)e^{r_{t:t+2^J}(j)}.$$

We generalize the approximation of log returns provided by Campbell, Chan, and Viceira (2003) to our portfolio on the period from t to $t+2^J$, obtaining

$$\begin{aligned} r_{p,t:t+2^J} &\simeq \sum_{j=1}^{J+1}\left(\frac{1}{J+1} - \pi_t(j)\right)f_{t:t+2^J}(j) + \sum_{j=1}^{J+1}\pi_t(j)r_{t:t+2^J}(j) + \frac{1}{2}var_t\left(r_{p,t:t+2^J}\right) \\ &= \frac{f_{t:t+2^J}}{J+1} + \boldsymbol{\pi}_t'\mathbf{x}_{t:t+2^J} + \frac{1}{2}var_t\left(r_{p,t:t+2^J}\right). \end{aligned}$$

See the derivation in Appendix A. Hence, the log portfolio return is approximated by

$$r_{p,t:t+2^J} \simeq \frac{f_{t:t+2^J}}{J+1} + \boldsymbol{\pi}_t'\mathbf{x}_{t:t+2^J} + 2^J\left\{\frac{1}{2}\boldsymbol{\pi}_t'\left(\boldsymbol{\sigma}_{\mathbf{x}}^2 - \Sigma_{\mathbf{x}}\boldsymbol{\pi}_t\right) + \frac{J}{J+1}\left[\boldsymbol{\pi}_t'\boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2}\frac{var_t\left(f_{t+2^J}\right)}{J+1}\right]\right\},$$

where $var_t\left(f_{t+2^J}\right)$ is simply obtained by summing up the entries of the vector $\boldsymbol{\sigma}_{\mathbf{f}}^2$. Note that, by setting $J=0$, we precisely retrieve Campbell, Chan, and Viceira (2003) approximation rule.

Moreover, we introduce a further simplifying assumption:

$$r_{t:t+2^J}(j) = 2^J r_{t+2^J}(j), \qquad f_{t:t+2^J}(j) = 2^J f_{t+2^J}(j), \qquad x_{t:t+2^J}(j) = 2^J x_{t+2^J}(j), \tag{4}$$

that allows us to compute the 2^J-period return by employing just the one-period return at time $t+2^J$. Hence, the previous approximation formula becomes

$$r_{p,t:t+2^J} \simeq 2^J \left\{ \frac{f_{t+2^J}}{J+1} + \boldsymbol{\pi}_t' \mathbf{x}_{t+2^J} + \frac{1}{2}\boldsymbol{\pi}_t' \left(\boldsymbol{\sigma}_{\mathbf{x}}^2 - \Sigma_{\mathbf{x}} \boldsymbol{\pi}_t\right) + \frac{J}{J+1}\left[\boldsymbol{\pi}_t' \boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2}\frac{var_t\left(f_{t+2^J}\right)}{J+1}\right]\right\}.$$

The value function per unit of wealth, namely $V_t = U_t/W_t$, is given by

$$V_t = (1-\beta)^{-\frac{\psi}{1-\psi}} \left(\frac{C_t}{W_t}\right)^{\frac{1}{1-\psi}}.$$

The Euler Equation associated with this optimization problem implies that the return $R_{i,t:t+2^J}$ on any asset i must satisfy the condition

$$E_t\left[\left\{\beta\left(\frac{C_{t+2^J}}{C_t}\right)^{-\frac{1}{\psi}}\right\}^{\theta} R_{p,t:t+2^J}^{-(1-\theta)} R_{i,t:t+2^J}\right] = 1. \tag{5}$$

When $i = p$, the Euler Equation rewrites as

$$E_t\left[\left\{\beta\left(\frac{C_{t+2^J}}{C_t}\right)^{-\frac{1}{\psi}} R_{p,t:t+2^J}\right\}^{\theta}\right] = 1$$

or, equivalently,

$$E_t\left[e^{\theta \log\beta - \frac{\theta}{\psi}\Delta_J c_{t+2^J} + \theta r_{p,t:t+2^J}}\right] = 1, \tag{6}$$

where $\Delta_J c_{t+2^J} \equiv c_{t+2^J} - c_t$ and c_t is log consumption at time t. Also in other occurrences we will employ the notation Δ_J for the first-difference operator over the period from t to $t+2^J$.

4 Solution method

4.1 Log-linearisation of Euler Equation and budget constraint

Our proposed solution method follows Campbell and Viceira (1999) and builds on the log-linear approximations of Euler Equation and intertemporal budget constraint previously proposed by Campbell (1993). In order to get a log-linear approximation of the Euler Equation (6), we take a second order Taylor approximation around the conditional mean of $\left\{\Delta_J c_{t+2^J}, r_{p,t:t+2^J}\right\}$ and then use the property of logs: $\log(1+\epsilon) \simeq \epsilon$ when ϵ is small enough. As a result, we get the log-linear approximate Euler Equation

$$0 = \theta \log \beta - \frac{\theta}{\psi} E_t \left[\Delta_J c_{t+2^J}\right] + \theta E_t \left[r_{p,t:t+2^J}\right] + \frac{1}{2} var_t \left(\frac{\theta}{\psi} \Delta_J c_{t+2^J} - \theta r_{p,t:t+2^J}\right).$$

Reordering terms, we get the equilibrium relationship between expected log consumption growth and expected log return on wealth

$$E_t \left[\Delta_J c_{t+2^J}\right] \simeq \psi \log \beta + v_{p,t} + \psi E_t \left[r_{p,t:t+2^J}\right], \tag{7}$$

where

$$v_{p,t} = \frac{1}{2} \frac{\theta}{\psi} var_t \left(\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J}\right). \tag{8}$$

We repeat the procedure and take a log-linear approximation of Euler Equation (1) in the case $r_{i,t:t+2^J} = r_{t:t+2^J}(j)$ for $j = 1, \ldots, J+1$. Then, we subtract the resulting log-linear equation for the case $r_{i,t:t+2^J} = f_{t:t+2^J}(j)$ from the equation obtained when $r_{i,t:t+2^J} = r_{t:t+2^J}(j)$ and we deduce the equation for log risk premia:

$$\begin{aligned}
E_t &\left[2^J x_{t+2^J}(j)\right] + \frac{1}{2} var_t \left(2^J x_{t+2^J}(j)\right) \\
&= 2^J cov_t \left(\frac{\theta}{\psi} \Delta_J c_{t+2^J} + (1-\theta)\, r_{p,t:t+2^J}, r_{t+2^J}(j)\right) \\
&\quad - 2^J cov_t \left(\frac{\theta}{\psi} \Delta_J c_{t+2^J} + (1-\theta)\, r_{p,t:t+2^J}, f_{t+2^J}(j)\right) \\
&\quad - \frac{1}{2} \left\{var_t \left(2^J r_{t+2^J}(j)\right) - var_t \left(2^J f_{t+2^J}(j)\right) - var_t \left(2^J x_{t+2^J}(j)\right)\right\}.
\end{aligned}$$

We also employed assumptions in Equations (4). By using the identity

$$\Delta_J c_{t+2^J} = \left(c_{t+2^J} - w_{t+2^J}\right) - (c_t - w_t) + \Delta_J w_{t+2^J} \tag{9}$$

and the log-linear approximation of budget constraint[8]

$$\Delta_J w_{t+2^J} \approx r_{p,t:t+2^J} + \left(1 - \frac{1}{\rho}\right)(c_t - w_t) + k, \tag{10}$$

[8]Here k and ρ are endogenous parameters depending on the mean optimal consumption-wealth ratio, which is determined once the model is solved (see Campbell (1993)):

$$\rho = 1 - e^{\mathbb{E}[c_t - w_t]}, \qquad k = \log(\rho) - \left(1 - \frac{1}{\rho}\right)\mathbb{E}\left[c_t - w_t\right].$$

Numerically, we start by setting $\rho = \beta$, as suggested by Campbell and Viceira (1999), and we compute the optimal consumption-wealth stream based on this value. Then, if $\mathbb{E}\left[c_t - w_t\right]$ is positive, we define a new ρ, as described above, and we repeat the procedure. In case the tolerance between two consequent values of ρ is smaller than 10^{-3}, we stop the recursion and we consider the consumption policy associated with the last value of optimal ρ.

we find the equation

$$
\begin{aligned}
E_t\left[x_{t+2^J}(j)\right] + \frac{2^J}{2} var_t\left(2^J x_{t+2^J}(j)\right) & \\
= \frac{\theta}{\psi}\left(\sigma_{r(j),c-w,t} - \sigma_{f(j),c-w,t}\right) + \gamma\left(\sigma_{r(j),p,t} - \sigma_{f(j),p,t}\right) & \\
- \frac{2^J}{2}\left\{var_t\left(r_{t+2^J}(j)\right) - var_t\left(f_{t+2^J}(j)\right) - var_t\left(x_{t+2^J}(j)\right)\right\}. &
\end{aligned} \tag{11}
$$

4.2 Solution for optimal portfolio

The goal now is to write the covariance terms at the right-hand side of Equation (11) as explicit functions of $\boldsymbol{\pi}_t$ and to solve for it. Observe that

$$\sigma_{r(j),p,t} = 2^J\left\{\boldsymbol{\pi}_t'\left(\Sigma_{\mathbf{x}}^{[j]} + \Sigma_{\mathbf{fx}}^{[j]}\right) + \frac{\boldsymbol{\sigma}_{\mathbf{fx}}(j)}{J+1} + \frac{\boldsymbol{\sigma}_{\mathbf{f}}^2(j)}{J+1}\right\}$$

and

$$\sigma_{f(j),p,t} = 2^J\left\{\boldsymbol{\pi}_t'\Sigma_{\mathbf{fx}}^{[j]} + \frac{\boldsymbol{\sigma}_{\mathbf{f}}^2(j)}{J+1}\right\},$$

where $\Sigma_{\mathbf{x}}^{[j]}$ denotes the j-th column of $\Sigma_{\mathbf{x}}$. By taking the difference $\sigma_{r(j),p,t} - \sigma_{f(j),p,t}$ and stacking the equations over j, we get

$$\boldsymbol{\sigma}_{\mathbf{x},p,t} = 2^J\left\{\Sigma_{\mathbf{x}}\boldsymbol{\pi}_t + \frac{\boldsymbol{\sigma}_{\mathbf{fx}}}{J+1}\right\}.$$

Moreover, note that $\sigma_{r(j),c-w,t}$ and $\sigma_{f(j),c-w,t}$ depend on the endogenous consumption-wealth ratio which, in turns, depends on $\boldsymbol{\pi}_t$ through $r_{p,t:t+2^J}$. See the log-linear approximation of $\Delta_J w_{t+2^J}$ in Equation (10) and $\Delta_J c_{t+2^J}$ in Equation (9). As a consequence of the concatenation of endogenous terms $\boldsymbol{\pi}_t$ and $c_t - w_t$, in order to find a solution for $(\boldsymbol{\pi}_t, c_t - w_t)$ we make a guess on the optimal portfolio and consumption rules.

Campbell and Viceira (1999) assume that weights assigned to market portfolio are affine in the equity premium, while optimal log consumption-wealth ratio is quadratic in the equity premium. Accordingly, using the vector of state variables $\mathbf{z}_t$, we guess

$$
\begin{aligned}
\boldsymbol{\pi}_t &= A_0 + A_1\mathbf{z}_t \\
c_t - w_t &= b_0 + B_1'\mathbf{z}_t + \mathbf{z}_t' B_2 \mathbf{z}_t,
\end{aligned}
$$

where A_0 is a vector of length $J+1$, A_1 is a $(J+1)\times(2J+2)$ matrix, b_0 is a scalar, B_1 is a vector of length $2J+2$ and B_2 is a square matrix of order $2J+2$. As suggested by Campbell, Chan, and Viceira (2003), we can assume B_2 to be symmetric in order to reduce the dimensionality of the problem, with no loss of generality.

By exploiting the guess on $c_t - w_t$, the covariances $\sigma_{r(j),c-w,t}$ and $\sigma_{f(j),c-w,t}$ rewrite as

$$\begin{aligned}\sigma_{r(j),c-w,t} \quad = \quad & B_1'\left(\Sigma_{\mathbf{v}}^{[J+j+1]}+\Sigma_{\mathbf{v}}^{[j]}\right)+\Phi_0' B_2\left(\Sigma_{\mathbf{v}}^{[J+j+1]}+\Sigma_{\mathbf{v}}^{[j]}\right)+\left(\Sigma_{\mathbf{v}}^{[J+j+1]}+\Sigma_{\mathbf{v}}^{[j]}\right)' B_2\Phi_0 \\ & +\mathbf{z}_t'\Phi' B_2\left(\Sigma_{\mathbf{v}}^{[J+j+1]}+\Sigma_{\mathbf{v}}^{[j]}\right)+\left(\Sigma_{\mathbf{v}}^{[J+j+1]}+\Sigma_{\mathbf{v}}^{[j]}\right)' B_2\Phi\mathbf{z}_t,\end{aligned}$$

$$\sigma_{f(j),c-w,t} = B_1'\Sigma_{\mathbf{v}}^{[j]}+\Phi_0' B_2\Sigma_{\mathbf{v}}^{[j]}+\left(\Sigma_{\mathbf{v}}^{[j]}\right)' B_2\Phi_0+\mathbf{z}_t'\Phi' B_2\Sigma_{\mathbf{v}}^{[j]}+\left(\Sigma_{\mathbf{v}}^{[j]}\right)' B_2\Phi\mathbf{z}_t.$$

Therefore

$$\sigma_{r(j),c-w,t}-\sigma_{f(j),c-w,t} = \left(\Sigma_{\mathbf{v}}^{[J+j+1]}\right)' B_1+\left(\Sigma_{\mathbf{v}}^{[J+j+1]}\right)'\left(B_2+B_2'\right)\left(\Phi_0+\Phi\mathbf{z}_t\right).$$

Stacked, these equations provide

$$\begin{aligned}\boldsymbol{\sigma}_{\mathbf{x},c-w,t} &= \boldsymbol{\sigma}_{\mathbf{r},c-w,t}-\boldsymbol{\sigma}_{\mathbf{f},c-w,t} \\ &= \left[\left(\Sigma_{\mathbf{v}}H_x'\right)' B_1+\left(\Sigma_{\mathbf{v}}H_x'\right)'\left(B_2+B_2'\right)\Phi_0\right]+\left[\left(\Sigma_{\mathbf{v}}H_x'\right)'\left(B_2+B_2'\right)\Phi\right]\mathbf{z}_t \\ &= \Lambda_0+\Lambda_1\mathbf{z}_t,\end{aligned}$$

where Λ_0 is a vector of length $J+1$, Λ_1 is a $(J+1)\times(2J+2)$ matrix and H_x is a selection matrix which selects the vector of excess returns $\mathbf{x}_t$ from $\mathbf{z}_t$.[9] Furthermore, up to multiplying by 2^J, the third term in the expression of log risk premia at the right-hand side of Equation (11) is equivalent to

$$\begin{aligned}-\frac{1}{2}\Bigg(var_t\left(r_{t+2^J}(j)\right)-var_t\left(f_{t+2^J}(j)\right)-var_t\left(x_{t+2^J}(j)\right)\Bigg) \\ &= var_t\left(f_{t+2^J}(j)\right)-cov_t\left(r_{t+2^J}(j),f_{t+2^J}(j)\right) \\ &= var_t\left(f_{t+2^J}(j)\right)-cov_t\left(x_{t+2^J}(j)+f_{t+2^J}(j),f_{t+2^J}(j)\right) \\ &= -cov_t\left(x_{t+2^J}(j),f_{t+2^J}(j)\right) \\ &= -\sigma_{x(j),f(j),t}\end{aligned}$$

and, piling over j, we can write

$$\left[-\sigma_{x(j),f(j),t}\right]_{j=1,..,J+1} = -\boldsymbol{\sigma}_{\mathbf{fx}}.$$

[9] H_x is the $(J+1)\times(2J+2)$ matrix defined by $\left[\begin{array}{ccc|c} 0 & \dots & 0 & \\ \vdots & \ddots & \vdots & I \\ 0 & \dots & 0 & \end{array}\right]$.

Finally, by stacking over j the log risk premia (the left-hand side of Equation (11)) and their expression at the right-hand side of the same equality, we get

$$E_t\left[\mathbf{x}_{t+2^J}\right]+\frac{2^J}{2}var_t\left(\mathbf{x}_{t+2^J}\right)=H_x\Phi_0+H_x\Phi\mathbf{z}_t+\frac{2^J}{2}\boldsymbol{\sigma}^2_{\mathbf{x}}$$

and

$$\frac{\theta}{\psi}\boldsymbol{\sigma}_{\mathbf{x},c-w,t}+\gamma\boldsymbol{\sigma}_{\mathbf{x},p,t}-2^J\boldsymbol{\sigma}_{\mathbf{fx}}=\frac{\theta}{\psi}\left(\Lambda_0+\Lambda_1\mathbf{z}_t\right)+2^J\gamma\left(\Sigma_{\mathbf{x}}\boldsymbol{\pi}_t+\frac{\boldsymbol{\sigma}_{\mathbf{fx}}}{J+1}\right)-2^J\boldsymbol{\sigma}_{\mathbf{fx}}.$$

Then, from the Euler Equation

$$E_t\left[\mathbf{x}_{t+2^J}\right]+\frac{2^J}{2}var_t\left(\mathbf{x}_{t+2^J}\right)=\frac{\theta}{\psi}\boldsymbol{\sigma}_{\mathbf{x},c-w,t}+2^J\gamma\left(\Sigma_{\mathbf{x}}\boldsymbol{\pi}_t+\frac{\boldsymbol{\sigma}_{\mathbf{fx}}}{J+1}\right)-2^J\boldsymbol{\sigma}_{\mathbf{fx}},$$

we obtain

$$\boldsymbol{\pi}_t=\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[E_t\left[\mathbf{x}_{t+2^J}\right]+\frac{2^J}{2}var_t\left(\mathbf{x}_{t+2^J}\right)+2^J\left(1-\frac{\gamma}{J+1}\right)\boldsymbol{\sigma}_{\mathbf{fx}}\right]}_{myopic\ demand}+\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[-\frac{\theta}{\psi}\boldsymbol{\sigma}_{\mathbf{x},c-w,t}\right]}_{hedging\ demand},$$

that is

$$\boldsymbol{\pi}_t=\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[H_x\Phi_0+H_x\Phi\mathbf{z}_t+\frac{2^J}{2}\boldsymbol{\sigma}^2_{\mathbf{x}}+2^J\left(1-\frac{\gamma}{J+1}\right)\boldsymbol{\sigma}_{\mathbf{fx}}\right]}_{myopic\ demand}+\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[-\frac{\theta}{\psi}\left(\Lambda_0+\Lambda_1\mathbf{z}_t\right)\right]}_{hedging\ demand}.$$

The strategic allocation in the risky assets, $\boldsymbol{\pi}_t$, displays two components. The first one is driven by the features of current investment opportunity set, such as the current risk premia and the covariance between risky and riskless assets. The second one, instead, is determined by future changes of the investment opportunity set, to the extent to which these are predictable through the covariance between current optimal consumption-wealth ratio and risk premia. Coherently, the first term of $\boldsymbol{\pi}_t$ is called *myopic demand* while the second one is referred to as *hedging demand*. Indeed, the first term depends on contemporary motives while the second one incorporates the hedging purposes of the investor, who wants to protect himself from unfavourable future changes in investment opportunities.

By collecting terms in an alternative way, we find

$$\boldsymbol{\pi}_t=\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[H_x\Phi_0+\frac{2^J}{2}\boldsymbol{\sigma}^2_{\mathbf{x}}+2^J\left(1-\frac{\gamma}{J+1}\right)\boldsymbol{\sigma}_{\mathbf{fx}}-\frac{\theta}{\psi}\Lambda_0\right]}_{A_0}+\underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left[-\frac{\theta}{\psi}\Lambda_1+H_x\Phi\right]}_{A_1}\mathbf{z}_t.$$

As a result, the optimal asset allocation is affine in the vector of the state variables $\mathbf{z}_t$. Moreover, by setting $J=0$, we retrieve the solution of Campbell, Chan, and Viceira (2003).

Observe that A_0 and A_1 depend on exogenous parameters, driving either asset returns or preferences, and on the endogenous coefficients of consumption policy through Λ_0 and Λ_1. In particular, for all $j = 1, \ldots, J$,

$$
\begin{aligned}
A_0(j) &= \frac{1-\phi_j^2}{2^J \gamma s_j^2 \left(1-\phi_j^{2^{J-j+1}}\right)} \left\{ \mu_j \left(1-\phi_j^{2^{J-j}}\right) \right. \\
&\quad \left. + \frac{2^J s_j^2}{2} \frac{1-\phi_j^{2^{J-j+1}}}{1-\phi_j^2} + 2^J \left(1-\frac{\gamma}{J+1}\right) \boldsymbol{\sigma}_{\mathbf{fx}}(j) \right\} \\
&\quad - \frac{\gamma-1}{2^J \gamma s_j^2} \frac{1}{1-\psi} \frac{1-\phi_j^2}{1-\phi_j^{2^{J-j+1}}} \Lambda_0(j) \\
A_0(J+1) &= \frac{1}{2^J \gamma s_{J+1}^2} \left\{ \mu_{J+1} \left(1-\phi_{J+1}\right) + \frac{2^J s_{J+1}^2}{2} + 2^J \left(1-\frac{\gamma}{J+1}\right) \boldsymbol{\sigma}_{\mathbf{fx}}(J+1) \right\} \\
&\quad - \frac{\gamma-1}{2^J \gamma s_{J+1}^2} \frac{1}{1-\psi} \Lambda_0(J+1) \\
A_1(j, J+1+j) &= \frac{1}{2^J \gamma s_j^2} \frac{1-\phi_j^2}{1-\phi_j^{2^{J-j+1}}} \phi_j^{2^{J-j}} - \frac{\gamma-1}{2^J \gamma s_j^2} \frac{1}{1-\psi} \frac{1-\phi_j^2}{1-\phi_j^{2^{J-j+1}}} \Lambda_1(j, J+1+j) \\
A_1(J+1, 2J+2) &= \frac{1}{2^J \gamma s_{J+1}^2} \phi_{J+1} - \frac{\gamma-1}{2^J \gamma s_{J+1}^2} \frac{1}{1-\psi} \Lambda_1(J+1, 2J+2) \\
A_1(j, i) &= -\frac{\gamma-1}{2^J \gamma s_j^2} \frac{1}{1-\psi} \frac{1-\phi_j^2}{1-\phi_j^{2^{J-j+1}}} \Lambda_1(j, i) \qquad \text{for} \quad i \neq J+1+j \\
A_1(J+1, i) &= -\frac{\gamma-1}{2^J \gamma s_{J+1}^2} \frac{1}{1-\psi} \Lambda_1(J+1, i) \qquad \text{for} \quad i \neq 2J+2.
\end{aligned}
$$

It is interesting to see that, when the scales i, j are different, $A_1(i, j)$ does not contain any myopic demand, but only the hedging term. In other words, the share $\pi_t(j)$ of portfolio invested in asset j depends on the assets at the scales $i \neq j$ just for hedging purposes. The resulting asset allocation can, then, be written as

$$\pi_t(j) = A_0(j) + A_1(j, j) x_t(j) + \sum_{i \neq j} A_1(j, i) x_t(i).$$

4.3 Discussion

The optimal asset allocation is driven by myopic and hedging reasons:

$$
\begin{aligned}
\boldsymbol{\pi}_t \quad = \quad & \underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left(H_x\Phi_0 + H_x\Phi\mathbf{z}_t + 2^J\left(1-\frac{\gamma}{J+1}\right)\boldsymbol{\sigma}_{\mathbf{fx}} + \frac{2^J}{2}\boldsymbol{\sigma}_{\mathbf{x}}^2\right)}_{myopic\ demand} \\
& \underbrace{-\frac{1}{2^J}\left(1-\frac{1}{\gamma}\right)\frac{1}{1-\psi}\Sigma_{\mathbf{x}}^{-1}\left(\Lambda_0 + \Lambda_1\mathbf{z}_t\right)}_{hedging\ demand}.
\end{aligned}
$$

Consequently, the coefficients of optimal asset allocation can be decomposed as

$$
\boldsymbol{\pi}_t = A_{0,myopic} + A_{0,hedging} + \left(A_{1,myopic} + A_{1,hedging}\right)\mathbf{z}_t,
$$

where

$$
\begin{aligned}
A_0 \quad &= \quad \underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}\left(H_x\Phi_0 + \frac{2^J}{2}\sigma_{\mathbf{x}}^2 + 2^J\left(1-\frac{\gamma}{J+1}\right)\boldsymbol{\sigma}_{\mathbf{fx}}\right)}_{A_{0,myopic}} \underbrace{-\frac{\gamma-1}{2^J\gamma}\frac{1}{1-\psi}\Sigma_{\mathbf{x}}^{-1}\Lambda_0}_{A_{0,hedging}} \\
A_1 \quad &= \quad \underbrace{\frac{1}{2^J\gamma}\Sigma_{\mathbf{x}}^{-1}H_x\Phi}_{A_{1,myopic}} \underbrace{-\frac{\gamma-1}{2^J\gamma}\frac{1}{1-\psi}\Sigma_{\mathbf{x}}^{-1}\Lambda_1}_{A_{1,hedging}}.
\end{aligned}
$$

Since the persistent components $x_t(i)$ and $x_t(l)$ are uncorrelated at any scale $i \neq l$, then the matrix $\Sigma_{\mathbf{x}}^{-1}$ is diagonal. Moreover, from Assumption (A3) Φ is diagonal as well and, therefore, the *myopic part* of $\pi_t(j)$ depends only on $x_t(j)$ and not on $x_t(l)$ for any $l \neq j$. Moreover, if $\gamma = 1$, the *hedging part* of $\boldsymbol{\pi}_t$ disappears, in line with Giovannini and Weil (1989). For instance, this happens when the investor has logarithmic utility, namely $\gamma = \psi = 1$.

Then, for a myopic investor $\pi_t(j)$ depends only on $x_t(j)$. On the contrary, if $\gamma \neq 1$, the resulting capital allocation in the j-th component of market returns depends also on the other components with different degree of persistence.

Note that $\pi_t(j)$ depends on $x_t(j)$ and, moreover, on $x_t(i)$, with $i \neq j$, through the term $A_{1,hedging}\mathbf{z}_t$. Hence, we can claim that the share $\pi_t(j)$ invested in the component $x_t(j)$ depends on the components at scales $i \neq j$ just for hedging purposes.

Finally, although the investor's horizon is 2^J, the optimal capital allocation involves all the components of market returns, not only the one with persistence J, even if he consumes and rebalances her portfolio every 2^J periods.

4.4 Solution for optimal consumption

In order to assert that our guess is indeed a solution to the investor's optimization problem, we still need to show that the optimal consumption-wealth ratio $c_t - w_t$ is quadratic in the vector of state variables $\mathbf{z}_t$. To reach this goal, we exploit two useful expressions for expected consumption growth:

$$\begin{aligned} E_t\left[\Delta_J c_{t+2^J}\right] &\simeq \psi\log\beta + v_{p,t} + \psi E_t\left[r_{p,t:t+2^J}\right] && (12)\\ E_t\left[\Delta_J c_{t+2^J}\right] &= E_t\left[c_{t+2^J} - w_{t+2^J}\right] - (c_t - w_t) + E_t\left[\Delta_J w_{t+2^J}\right]. && (13)\end{aligned}$$

Equation (12) is obtained from the log-linear approximation of the Euler Equation - see Equations (7) and (8) - while Equation (13) follows from the accounting identity (9).

In particular, if we substitute the log-linear approximation of the budget constraint of Equation (10) into the accounting identity, we get

$$\Delta_J c_{t+2^J} \approx \left(c_{t+2^J} - w_{t+2^J}\right) - \frac{1}{\rho}\left(c_t - w_t\right) + r_{p,t:t+2^J} + k \qquad (14)$$

and, as a consequence, Equation (13) rewrites as

$$E_t\left[\Delta_J c_{t+2^J}\right] = E_t\left[c_{t+2^J} - w_{t+2^J}\right] - \frac{1}{\rho}\left(c_t - w_t\right) + E_t\left[r_{p,t:t+2^J}\right] + k. \qquad (15)$$

Combining Equations (12) and (15), we obtain

$$c_t - w_t = -\rho\left(\psi\log\beta + v_{p,t}\right) + \rho(1-\psi)E_t\left[r_{p,t:t+2^J}\right] + E_t\left[c_{t+2^J} - w_{t+2^J}\right] + k. \qquad (16)$$

In order to prove that $c_t - w_t$ is quadratic in $\mathbf{z}_t$ we only need to show that $E_t\left[r_{p,t:t+2^J}\right]$ and $v_{p,t}$ are quadratic in $\mathbf{z}_t$ because we already know (by using the guess at $t+2^J$) that $E_t\left[c_{t+2^J} - w_{t+2^J}\right]$ is quadratic in the vector of state variables. As for the expected portfolio return, it holds

$$\begin{aligned} E_t\left[r_{p,t:t+2^J}\right] &= 2^J\Bigg\{\frac{E_t\left[f_{t+2^J}\right]}{J+1} + \boldsymbol{\pi}_t' E_t\left[\mathbf{x}_{t+2^J}\right] + \frac{1}{2}\boldsymbol{\pi}_t'\left(\boldsymbol{\sigma}_{\mathbf{x}}^2 - \Sigma_{\mathbf{x}}\boldsymbol{\pi}_t\right)\\ &\quad + \frac{J}{J+1}\left[\boldsymbol{\pi}_t'\boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2}\frac{var_t\left(f_{t+2^J}\right)}{J+1}\right]\Bigg\}\\ &= 2^J\Bigg\{\frac{\boldsymbol{\iota}'}{J+1}H_f\left(\Phi_0 + \Phi\mathbf{z}_t\right) + \left(A_0 + A_1\mathbf{z}_t\right)' H_x\left(\Phi_0 + \Phi\mathbf{z}_t\right)\\ &\quad + \frac{1}{2}\left(A_0 + A_1\mathbf{z}_t\right)'\boldsymbol{\sigma}_{\mathbf{x}}^2 - \frac{1}{2}\left(A_0 + A_1\mathbf{z}_t\right)'\Sigma_{\mathbf{x}}\left(A_0 + A_1\mathbf{z}_t\right)\\ &\quad + \frac{J}{J+1}\left[\left(A_0 + A_1\mathbf{z}_t\right)'\boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2}\frac{var_t\left(f_{t+2^J}\right)}{J+1}\right]\Bigg\}\\ &= \Gamma_0 + \Gamma_1\mathbf{z}_t + \Gamma_2 vec(\mathbf{z}_t\mathbf{z}_t'), && (17)\end{aligned}$$

where the second equality follows from replacing $\boldsymbol{\pi}_t$ with its guess, $\boldsymbol{\iota}$ is a $(J+1)$-vector of ones, H_f is a matrix which selects the short-term interest rate components from the vector

$\mathbf{z}_t$, $\Gamma_0 \in \mathbb{R}$ and Γ_1, Γ_2 are vectors:

$$
\begin{aligned}
\Gamma_0 &= 2^J \left\{ \frac{\boldsymbol{\iota}'}{J+1} H_f \Phi_0 + A_0' H_x \Phi_0 + \frac{1}{2} A_0' \boldsymbol{\sigma}_{\mathbf{x}}^2 - \frac{1}{2} A_0' \Sigma_{\mathbf{x}} A_0 + \frac{J}{J+1} \left[A_0' \boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2} \frac{var_t\left(f_{t+2^j}\right)}{J+1} \right] \right\} \\
\Gamma_1 &= 2^J \left\{ \frac{\boldsymbol{\iota}'}{J+1} H_f \Phi + \Phi_0' H_x' A_1 + A_0' H_x \Phi + \frac{1}{2} \left(\boldsymbol{\sigma}_{\mathbf{x}}^2 \right)' A_1 - A_0' \Sigma_{\mathbf{x}} A_1 + \frac{J}{J+1} \boldsymbol{\sigma}_{\mathbf{fx}}' A_1 \right\} \\
\Gamma_2 &= 2^J \left\{ vec(A_1' H_x \Phi)' - \frac{1}{2} vec\left(A_1' \Sigma_{\mathbf{x}} A_1 \right)' \right\}.
\end{aligned}
$$

Hence, we are left to prove that $v_{p,t}$ is quadratic in $\mathbf{z}_t$. Recall that

$$
v_{p,t} = \frac{1}{2} \frac{\theta}{\psi} var_t \left(\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} \right).
$$

In order to compute this variance, it is convenient to write the expression for the innovation

$$
\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} - E_t \left[\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} \right]
$$

since

$$
var_t \left(\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} \right) = var_t \left(\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} - E_t \left[\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} \right] \right).
$$

From Equation (14), we observe that

$$
\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} = \left(c_{t+2^J} - w_{t+2^J} \right) - \frac{1}{\rho} \left(c_t - w_t \right) + (1-\psi) r_{p,t:t+2^J} + k
$$

and, therefore,

$$
\begin{aligned}
&\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} - E_t \left[\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} \right] \\
&\qquad = c_{t+2^J} - w_{t+2^J} - E_t \left[c_{t+2^J} - w_{t+2^J} \right] + (1-\psi) \left(r_{p,t:t+2^J} - E_t \left[r_{p,t:t+2^J} \right] \right).
\end{aligned}
$$

Moreover, from previous calculations it is immediate to see that

$$
r_{p,t:t+2^J} - E_t \left[r_{p,t:t+2^J} \right] = 2^J \left\{ \frac{\boldsymbol{\iota}'}{J+1} H_f \mathbf{v}_{t+2^J} + A_0' H_x \mathbf{v}_{t+2^J} + \mathbf{z}_t' A_1' H_x \mathbf{v}_{t+2^J} \right\},
$$

while, by exploiting the guess at time $t + 2^J$ on consumption policy, we have

$$
\begin{aligned}
c_{t+2^J} - w_{t+2^J} &= b_0 + B_1' \mathbf{z}_{t+2^J} + \mathbf{z}_{t+2^J}' B_2 \mathbf{z}_{t+2^J} \\
&= b_0 + B_1' \mathbf{z}_{t+2^J} + \Phi_0' B_2 \Phi_0 + \Phi_0' B_2 \Phi \mathbf{z}_t + \Phi_0' B_2 \mathbf{v}_{t+2^J} \\
&\quad + \mathbf{z}_t' \Phi' B_2 \Phi \mathbf{z}_t + \mathbf{z}_t' \Phi' B_2 \Phi_0 + \mathbf{z}_t' \Phi' B_2 \mathbf{v}_{t+2^J} + \mathbf{v}_{t+2^J}' B_2 \Phi_0 \\
&\quad + \mathbf{v}_{t+2^J}' B_2 \Phi \mathbf{z}_t + vec\left(B_2\right)' vec\left(\mathbf{v}_{t+2^J}' \mathbf{v}_{t+2^J} \right).
\end{aligned}
$$

Since Assumptions (A3) and (A4) ensure that

$$E_t\left[\mathbf{v}_{t+2^J}\right] = \mathbf{0},$$

$$E_t\left[\mathbf{z}_{t+2^J}\right] = E_t\left[\Phi_0 + \Phi\mathbf{z}_t + \mathbf{v}_{t+2^J}\right] = \Phi_0 + \Phi\mathbf{z}_t$$

and

$$E_t\left[vec\left(B_2\right)' vec\left(\mathbf{v}'_{t+2^J}\mathbf{v}_{t+2^J}\right)\right] = vec\left(B_2\right)' \Sigma_{\mathbf{v}},$$

then

$$\begin{aligned} E_t\left[c_{t+2^J} - w_{t+2^J}\right] &= b_0 + B'_1\Phi_0 + B'_1\Phi\mathbf{z}_t + \Phi'_0 B_2\Phi_0 + \Phi'_0 B_2 \Phi\mathbf{z}_t \\ &\quad +\mathbf{z}'_t\Phi' B_2\Phi\mathbf{z}_t + \mathbf{z}'_t\Phi' B_2\Phi_0 + vec\left(B_2\right)'\Sigma_{\mathbf{v}}, \end{aligned} \tag{18}$$

$$\begin{aligned} c_{t+2^J} - w_{t+2^J} - E_t\left[c_{t+2^J} - w_{t+2^J}\right] &= B'_1\mathbf{v}_{t+2^J} + \Phi'_0 B_2\mathbf{v}_{t+2^J} + \mathbf{z}'_t\Phi' B_2\mathbf{v}_{t+2^J} \\ &\quad +\mathbf{v}'_{t+2^J} B_2\Phi_0 + \mathbf{v}'_{t+2^J} B_2\Phi\mathbf{z}_t \\ &\quad +vec\left(B_2\right)' vec\left(\mathbf{v}'_{t+2^J}\mathbf{v}_{t+2^J}\right) - vec\left(B_2\right)'\Sigma_{\mathbf{v}}. \end{aligned}$$

As a consequence,

$$\begin{aligned} &\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J} - E_t\left[\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J}\right] \\ = \; & c_{t+2^J} - w_{t+2^J} - E_t\left[c_{t+2^J} - w_{t+2^J}\right] + (1-\psi)\left(r_{p,t:t+2^J} - E_t\left[r_{p,t:t+2^J}\right]\right) \\ = \; & \left[B'_1 + \Phi'_0(B_2 + B'_2) + 2^J(1-\psi)A'_0 H_x + 2^J\frac{1-\psi}{J+1}\boldsymbol{\iota}' H_f\right]\mathbf{v}_{t+2^J} \\ & +\mathbf{z}'_t\left[\Phi'(B_2 + B'_2) + 2^J(1-\psi)A'_1 H_x\right]\mathbf{v}_{t+2^J} + vec(B_2)' vec(\mathbf{v}_{t+2^J}\mathbf{v}'_{t+2^J}) \\ = \; & \left\{\Pi_1 + \mathbf{z}'_t\Pi_2\right\}\mathbf{v}_{t+2^J} + vec(B_2)' vec(\mathbf{v}_{t+2^J}\mathbf{v}'_{t+2^J}), \end{aligned}$$

where we define the vector Π_1 and the matrix Π_2 as

$$\Pi_1 = B'_1 + \Phi'_0(B_2 + B'_2) + 2^J(1-\psi)A'_0 H_x + 2^J\frac{1-\psi}{J+1}\boldsymbol{\iota}' H_f$$

$$\Pi_2 = \Phi'(B_2 + B'_2) + 2^J(1-\psi)A'_1 H_x.$$

Therefore,

$$\begin{aligned} var_t\left(\Delta_J c_{t+2^J} - \psi r_{p,t:t+2^J}\right) &= \Pi_1\Sigma_{\mathbf{v}}\Pi'_1 + \left(2\Pi_1\Sigma_{\mathbf{v}}\Pi'_2\right)\mathbf{z}_t + \\ &\quad +vec\left(\Pi_2\Sigma_{\mathbf{v}}\Pi'_2\right) vec\left(\mathbf{z}_t\mathbf{z}'_t\right) + \\ &\quad +vec(B_2)' var_t\left(vec(\mathbf{v}_{t+2^J}\mathbf{v}'_{t+2^J})\right) vec(B_2), \end{aligned}$$

since $\mathbf{v}_{t+2^J}$ is conditionally normally distributed - and so all third moments are zero. This proves that $v_{p,t}$ is quadratic in the state vector $\mathbf{z}_t$ and, in particular,

$$v_{p,t} = V_0 + V_1\mathbf{z}_t + V_2 vec\left(\mathbf{z}_t\mathbf{z}'_t\right), \tag{19}$$

with $V_0 \in \mathbb{R}$ and the vectors V_1, V_2 defined by

$$\begin{aligned}
V_0 &= \frac{\theta}{2\psi}\left[\Pi_1 \Sigma_{\mathbf{v}} \Pi_1' + vec(B_2)' var_t \left(vec(\mathbf{v}_{t+2^J} \mathbf{v}_{t+2^J}')\right) vec(B_2)\right] \\
V_1 &= \frac{\theta}{2\psi} 2\Pi_1 \Sigma_{\mathbf{v}} \Pi_2' \\
V_2 &= \frac{\theta}{2\psi} vec\left(\Pi_2 \Sigma_{\mathbf{v}} \Pi_2'\right)'.
\end{aligned}$$

As a result, we have shown that the optimal consumption-wealth ratio is quadratic in $\mathbf{z}_t$, as conjectured. In order to solve for the coefficients of the optimal consumption rule b_0, B_1 and B_2, we simply substitute the expressions of $E_t\left[r_{p,t:t+2^J}\right]$, $E_t\left[c_{t+2^J} - w_{t+2^J}\right]$ and $v_{p,t}$ provided by Equations (17), (18) and (19) respectively into the expression of $c_t - w_t$ given by Equation (16). Hence, we deduce that

$$c_t - w_t = \Xi_0 + \Xi_1 \mathbf{z}_t + \Xi_2 vec\left(\mathbf{z}_t \mathbf{z}_t'\right),$$

where $\Xi_0 \in \mathbb{R}$ and Ξ_1, Ξ_2 are the vectors

$$\begin{aligned}
\Xi_0 &= \rho[-\psi \log\beta + k - V_0 + (1-\psi)\Gamma_0 + b_0 + B_1'\Phi_0 \\
&\quad + vec(B_2)' vec\left(\Phi_0 \Phi_0'\right) + vec(B_2)' vec(\Sigma_{\mathbf{v}})] \\
\Xi_1 &= \rho[-V_1 + (1-\psi)\Gamma_1 + B_1'\Phi + 2\Phi_0'(B_2' + B_2)\Phi] \\
\Xi_2 &= \rho[-V_2 + (1-\psi)\Gamma_2 + vec(\Phi' B_2 \Phi)'].
\end{aligned}$$

The last equations further clarify that $c_t - w_t$ is quadratic in the state vector, as conjectured. Observe that Ξ_0, Ξ_1 and Ξ_2 depend on b_0, B_1 and B_2. Thus, for the solution to be consistent, it must be

$$\begin{aligned}
b_0 &= \Xi_0 \\
B_1 &= \Xi_1' \\
vec\,(B_2) &= \Xi_2'.
\end{aligned}$$

5 Estimation of multiscale impulse responses and persistent components

We briefly recap the estimation strategy of Ortu, Severino, Tamoni, and Tebaldi (2017) for the coefficients $\beta_k^{(j)}$ in the Extended Wold Decomposition of a zero-mean weakly stationary time series $\mathbf{x} = \{x_t\}_t$ such that

$$x_t = \Phi(\mathbf{L})\varepsilon_t = \sum_{h=0}^{+\infty} \alpha_h \varepsilon_{t-h}, \qquad \sum_{h=0}^{+\infty} \alpha_h^2 < +\infty,$$

with the operator $\Phi(\mathbf{L})$ invertible.

We approximate x_t by an autoregressive of a suitable order N, determined by the Bayesian Information Criterion (BIC), namely

$$x_t = \sum_{k=1}^{N} b_k x_{t-k} + \eta_t.$$

By OLS we estimate the regression coefficients b_k and the unit variance white noise $\varepsilon_t = \eta_t \mathbb{E}\left[\eta_t^2\right]^{-1/2}$. From the relation

$$x_t = \eta_t + \sum_{k=1}^{N} b_k \sum_{h=0}^{+\infty} \alpha_h \varepsilon_{t-k-h} = \mathbb{E}\left[\eta_t^2\right]^{1/2} \varepsilon_t + \sum_{n=1}^{+\infty} \left(\sum_{h=\max\{n-N,0\}}^{n-1} \alpha_h b_{n-h} \right) \varepsilon_{t-n},$$

we get the impulse response functions

$$\alpha_0 = \mathbb{E}\left[\left(x_t - \sum_{k=1}^{N} b_k x_{t-k}\right)^2\right]^{1/2}, \qquad \alpha_n = \sum_{h=\max\{n-N,0\}}^{n-1} \alpha_h b_{n-h} \quad \forall n \in \mathbb{N}.$$

Then, multiscale impulse responses $\beta_k^{(j)}$ can be computed by employing Theorem 1. In turns, persistent components can be obtained by using as innovations the residuals of the initial autoregressive regression.

5.1 Estimated portfolio weights

We consider daily data from January 4, 1954 to December 30, 2016. Market portfolio returns are `vwretd` of CRSP S&P 500 index, while risk-free rates are taken from FRED DTB3, which contains the (not seasonally adjusted) secondary market rates of three-month Treasury Bills. Inflation data are derived from monthly inflation of CRSP database on a compound basis. We employ inflation for the computation of real log risk-free rate (log risk-free rate minus log inflation) and real log market return (log return on the S&P index minus log inflation). The BIC criterium applied to real log market return recommends to employ an $AR(2)$ process for the construction.

To make our portfolio analysis we fix a maximum scale $J = 4$. This choice roughly corresponds to a monthly horizon because scale 4 captures shocks with an approximative duration of sixteen working days. As already signalled in the Appendix of Campbell, Chan, and Viceira (2003), the solution algorithm for multivariate strategic allocation is computationally intensive, due to the employed approximations and the dimensionality of the linear systems involved. Moreover, in our application the algorithm is not always stable. Therefore, we focus our discussion on averages results and we provide an heuristic illustration.

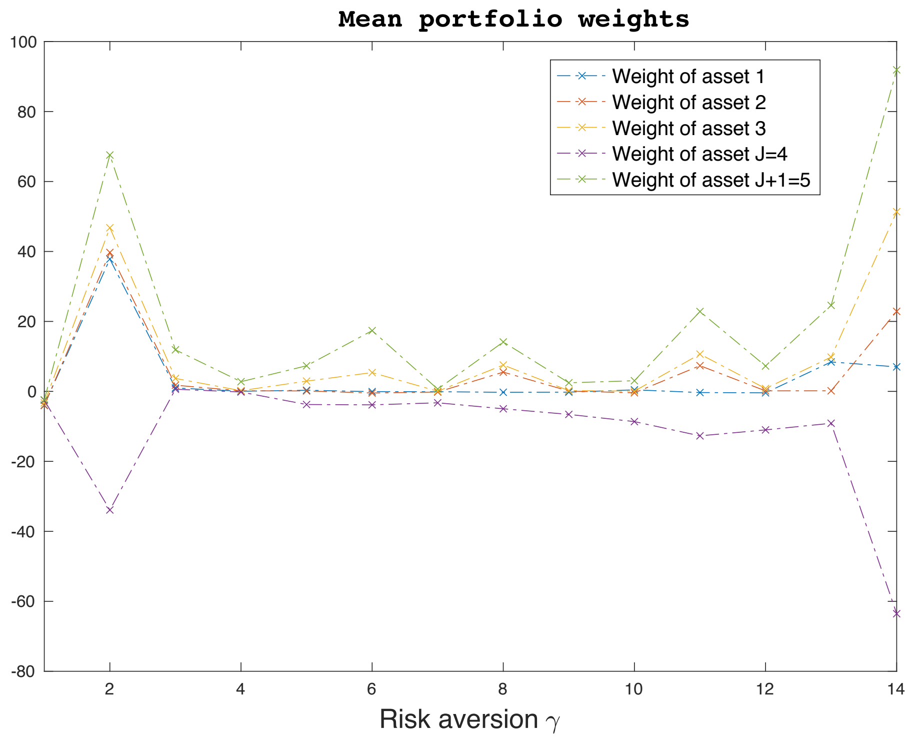

Figure 6: Mean optimal portfolio weights $\pi(j)$ for $j = 1, ..., 5$ with respect to different degrees of relative risk aversion $\gamma = 1, \ldots, 14$ and fixed $\psi = 0.99$.

First, we set the IES $\psi = 0.99$ and we run the algorithm for increasing values of relative risk aversion γ. We estimate the optimal asset allocation $\boldsymbol{\pi}_t$ of an investor that rebalances her portfolio monthly as described in Section 3. Then, we take the average of portfolio weights over the whole time series. The strategic choice occurs every 2^J working days and we have 2^J possible starting dates. Therefore, we take the mean of portfolio weights across these realizations, too. The outcome is depicted in Figure 6.

We see from the graph that the allocation generally increases (in absolute value) with the scale and with the relative risk aversion. The highest loadings are associated with the $(J+1)$-th asset, which collects the sensitivity to shocks with any persistence higher than J. The investment in the J-th security is the one to be negative for almost all degrees of risk aversion. Hence, the optimal portfolio rule prescribes a relevant short-selling of this asset, that partly counterbalances the large purchase of the $(J+1)$-th security. These observations suggest that the $(J+1)$-th asset shares the same interpretation of the optimal growth portfolio discussed, for instance, by Merton (1969). See also the more recent summary by Christensen (2012).

If $\gamma = 1$ the investor is fully myopic and the weights are similar across scales. When γ increases, the investment diversifies within persistent assets and the allocation to high scales becomes prominent.

In order to understand more deeply the dependence of portfolio loadings from persistent components we consider the entries of the matrix A_1 for raising values of risk aversion. For this analysis we employ the last eighth of the sample. We first plot the optimal portfolio weights in Figure 7. Observe that the average weights on the two more persistent assets almost overlap. Then, as an example, we focus on the dependences of the fourth optimal weight $\pi(4)$ on each persistent security. Specifically, we depict in Figure 8 the entries $(4, j)$ of the matrix A_1 for $j = J+2, \ldots, 2J+2$. This graph, in fact, complements the outcomes of Figure 7 which refer to the whole $\boldsymbol{\pi}_t = A_0 + A_1\mathbf{z}_t$.

At a first glance, we note that the fourth loading is mostly dependent on the persistent component at scale 4. This feature is shared also by the other weights (not plotted here), which are mainly sensitive to the security at the same scale. Nevertheless, when the agent's risk aversion raises, the relative importance of this security shrinks visibly. This behaviour conveys the intuition that a less risk-averse investor bases her valuations on scale-specific factors, while a more risk-averse trader is likey to include the other components in her strategy. This phenomenon is consistent with the hedging nature of persistent components presented in Subsection 4.3.

However, the behaviours of Figure 8 are not always present. For instance, we can set the IES $\psi = 0.5$ and solve the asset allocation problem on the whole sample. Results about mean portfolio weights and the sensitivity of the first loading with respect to the persistent securities are plotted in Figures 9 and 10 respectively. In this example, the weight on the first persistent asset equally depends on the risky securities at any scale. Also the other portfolio weights feature the same property. Indeed, low IES induces smooth consumption streams and hedging is remarkable at any level of risk aversion.

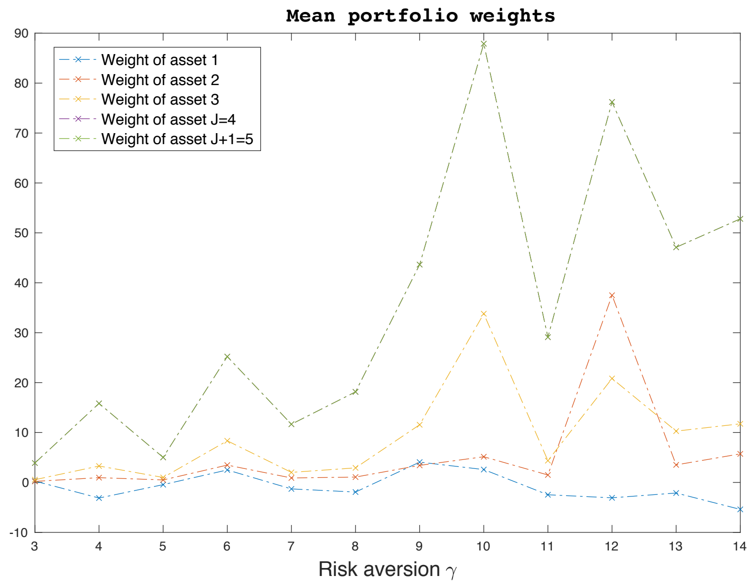

Figure 7: Mean optimal portfolio weights $\pi(j)$ for $j = 1, ..., 5$ with respect to different degrees of relative risk aversion $\gamma = 3, \ldots, 14$ and fixed $\psi = 0.99$. Here the last eighth of the sample is employed.

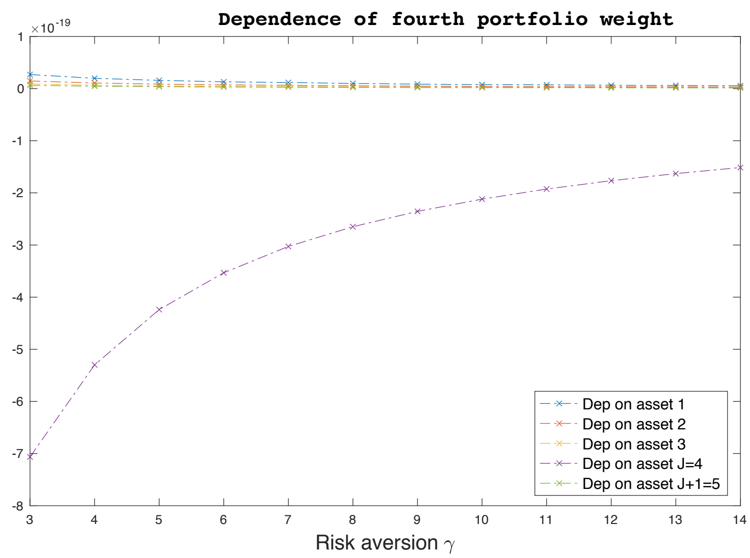

Figure 8: Dependence of the fourth (mean) optimal portfolio weight $\pi(4)$ from each persistent asset with respect to different degrees of relative risk aversion $\gamma = 3, \ldots, 14$ and fixed $\psi = 0.99$. Here the last eighth of the sample is employed.

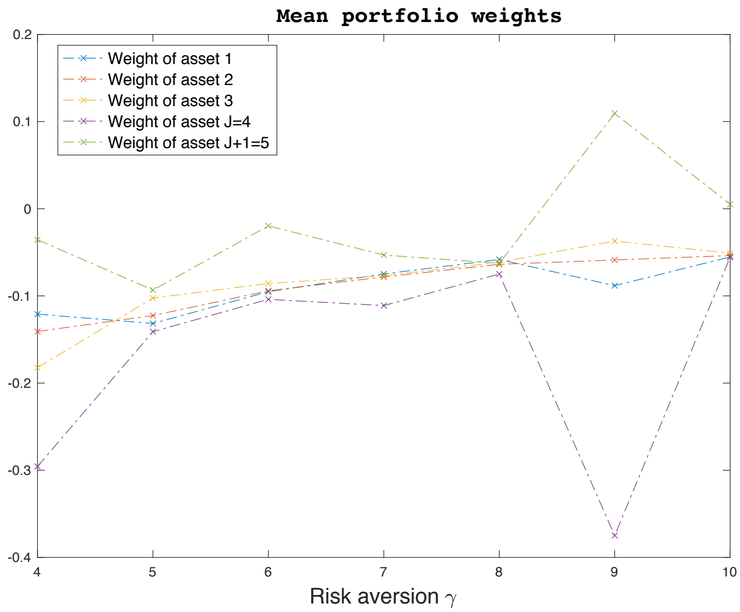

Figure 9: Mean optimal portfolio weights $\pi(j)$ for $j = 1, ..., 5$ with respect to different degrees of relative risk aversion $\gamma = 4, \ldots, 10$ and fixed $\psi = 0.5$.

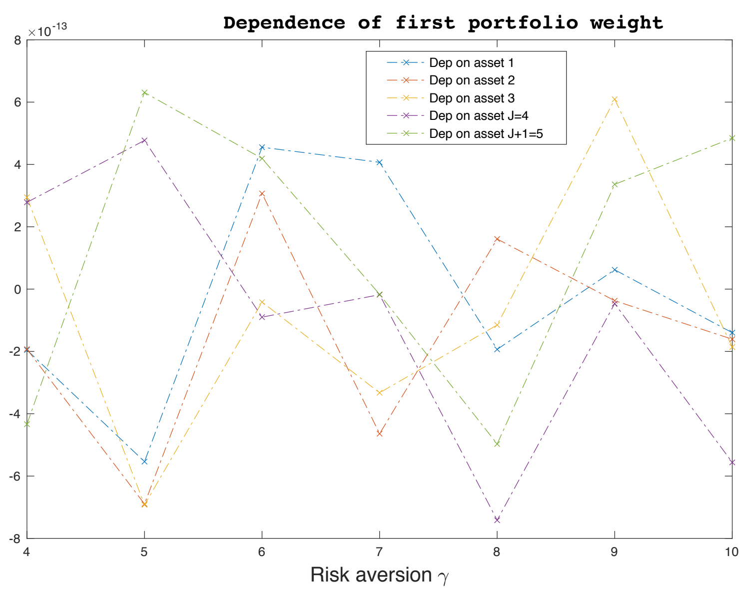

Figure 10: Dependence of the first (mean) optimal portfolio weight $\pi(1)$ from each persistent asset with respect to different degrees of relative risk aversion $\gamma = 4, \ldots, 10$ and fixed $\psi = 0.5$.

6 Conclusions

This work constitutes a first attempt to formalize strategic asset allocation when shocks are heterogeneous in terms of duration. The sensitivity of returns to such innovations is captured by scale-specific components, retrieved by the Extended Wold Decomposition. When investors are allowed to trade securities whose returns mimic the ones of persistent components, we provide an approximately optimal way to allocate wealth (and decide, in turn, consumption). The main contribution is the embedding of persistence-based returns into the classical multiperiod portfolio optimization framework, which employs peculiar tools such as autoregressive dynamics and log-linearisation techniques.

The empirical implementation suffers from weak algorithmic stability that, however, is also present in standard intertemporal multivariate asset allocation. Moreover, an important question is the feasibility of the replication of persistent components through traded securities in the market. This aspect is fundamental to make our results usable by practitioners. We will devote future research to deeply understand and solve these issues.

A Derivation of portfolio log return approximation

The log return on the portfolio $r_{p,t:t+2^J}$ is a discrete-time approximation of its continuous-time counterpart. In this section we show how to obtain the following approximation

$$r_{p,t:t+2^J} \simeq \frac{f_{t:t+2^J}}{J+1} + \boldsymbol{\pi}_t' \mathbf{x}_{t:t+2^J} + 2^J \left\{ \frac{1}{2} \boldsymbol{\pi}_t' \left(\boldsymbol{\sigma}_{\mathbf{x}}^2 - \Sigma_{\mathbf{x}} \boldsymbol{\pi}_t \right) + \frac{J}{J+1} \left[\boldsymbol{\pi}_t' \boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2} \frac{var_t \left(f_{t+2^J} \right)}{J+1} \right] \right\}.$$

We start by assuming the dynamics for the value processes associated with the components of risk-free asset (vector $\mathbf{B}_t$) and market portfolio (vector $\mathbf{P}_t$):

$$\frac{d\mathbf{B}_t}{\mathbf{B}_t} = \boldsymbol{\mu}_{b,t} dt + \boldsymbol{\sigma}_b d\mathbf{W}_t \qquad \text{(A.1)}$$

$$\frac{d\mathbf{P}_t}{\mathbf{P}_t} = \boldsymbol{\mu}_t dt + \boldsymbol{\sigma} d\mathbf{W}_t \qquad \text{(A.2)}$$

where $\boldsymbol{\mu}_{b,t}$ and $\boldsymbol{\mu}_t$ are the drift vectors, $\boldsymbol{\sigma}_b$ and $\boldsymbol{\sigma}$ are the diffusion matrices and $\mathbf{W}_t$ is a $J+1$-dimensional standard Brownian motion. We can obtain the corresponding log values by employing Ito's Lemma:

$$d\log B_{j,t} = \left(\frac{dB_{j,t}}{B_{j,t}} \right) - \frac{1}{2} \left(\boldsymbol{\sigma}_{bj} \boldsymbol{\sigma}_{bj}' \right) dt \qquad \text{(A.3)}$$

$$d\log P_{j,t} = \left(\frac{dP_{j,t}}{P_{j,t}} \right) - \frac{1}{2} \left(\boldsymbol{\sigma}_{j} \boldsymbol{\sigma}_{j}' \right) dt \qquad \text{(A.4)}$$

where $\boldsymbol{\sigma}_j$ ($\boldsymbol{\sigma}_{bj}$) represents the j-th row of the diffusion matrix $\boldsymbol{\sigma}$ ($\boldsymbol{\sigma}_b$), for $j = 1, ..., J+1$.

Let V_t be the portfolio value at time t and denote by $\boldsymbol{\iota}$ a vector of $J+1$ ones. We have

$$V_t = \boldsymbol{\pi}_t' \mathbf{P}_t + \left(\frac{1}{J+1} \boldsymbol{\iota}' - \boldsymbol{\pi}_t' \right) \mathbf{B}_t$$

and we set

$$r_{p,t:t+2^J} = d\log V_t = \frac{dV_t}{V_t} - \frac{1}{2} \left(\frac{dV_t}{V_t} \right)^2.$$

In particular,

$$\begin{aligned}
\frac{dV_t}{V_t} &= \boldsymbol{\pi}_t' \frac{d\mathbf{P}_t}{\mathbf{P}_t} + \left(\frac{1}{J+1} \boldsymbol{\iota}' - \boldsymbol{\pi}_t' \right) \frac{d\mathbf{B}_t}{\mathbf{B}_t} \\
&= \boldsymbol{\pi}_t' \left(d\log \mathbf{P}_t + \frac{1}{2} \left[\boldsymbol{\sigma}_j \boldsymbol{\sigma}_j' \right] dt \right) + \left(\frac{1}{J+1} \boldsymbol{\iota}' - \boldsymbol{\pi}_t' \right) \left(d\log \mathbf{B}_t + \frac{1}{2} \left[\boldsymbol{\sigma}_{bj} \boldsymbol{\sigma}_{bj}' \right] dt \right) \\
&= \boldsymbol{\pi}_t' \left(d\log \mathbf{P}_t - d\log \mathbf{B}_t \right) + \frac{1}{J+1} \boldsymbol{\iota}' d\log \mathbf{B}_t + \frac{1}{2} \boldsymbol{\pi}_t' \left(\left[\boldsymbol{\sigma}_j \boldsymbol{\sigma}_j' \right] - \left[\boldsymbol{\sigma}_{bj} \boldsymbol{\sigma}_{bj}' \right] \right) dt \\
&\quad + \frac{1}{2} \frac{1}{J+1} \boldsymbol{\iota}' \left[\boldsymbol{\sigma}_{bj} \boldsymbol{\sigma}_{bj}' \right] dt,
\end{aligned}$$

where brackets $[\cdot]$ denote a vector with $\boldsymbol{\sigma}_j\boldsymbol{\sigma}_j'$ (or $\boldsymbol{\sigma}_{bj}\boldsymbol{\sigma}_{bj}'$) entries. Moreover

$$\begin{aligned}\left(\frac{dV_t}{V_t}\right)^2 &= \boldsymbol{\pi}_t'\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)'\boldsymbol{\pi}_t \\ &\quad +\frac{1}{J+1}\boldsymbol{\iota}'\left(d\log\mathbf{B}_t\right)\left(d\log\mathbf{B}_t\right)'\frac{1}{J+1}\boldsymbol{\iota} \\ &\quad +2\boldsymbol{\pi}_t'\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)\left(\frac{1}{J+1}\boldsymbol{\iota}' d\log\mathbf{B}_t\right) + o\left(dt\right),\end{aligned}$$

where all $o\left(dt\right)$ terms vanish because they involve either dt^2 or $dtd\mathbf{W}_t$. By combining Equations (A.1)-(A.4), we obtain

$$\begin{aligned}d\log\mathbf{P}_t - d\log\mathbf{B}_t &= \boldsymbol{\mu}_t dt + \boldsymbol{\sigma} d\mathbf{W}_t - \frac{1}{2}\left[\boldsymbol{\sigma}_j\boldsymbol{\sigma}_j'\right]dt \\ &\quad -\boldsymbol{\mu}_{b,t}dt - \boldsymbol{\sigma}_b d\mathbf{W}_t + \frac{1}{2}\left[\boldsymbol{\sigma}_{bj}\boldsymbol{\sigma}_{bj}'\right]dt \\ &\approx \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)d\mathbf{W}_t,\end{aligned}$$

where the last line follows after ignoring dt terms. Then,

$$\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)' = \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)' dt,$$

$$\left(d\log\mathbf{P}_t - d\log\mathbf{B}_t\right)\left(\frac{1}{J+1}\boldsymbol{\iota}' d\log\mathbf{B}_t\right) = \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\boldsymbol{\sigma}_b'\boldsymbol{\iota}\frac{1}{J+1}dt.$$

Consequently, by employing the notations $\mathbf{x}_{t:t+2^J} = d\log\mathbf{P}_t - d\log\mathbf{B}_t$ and $\mathbf{f}_{t:t+2^J} = d\log\mathbf{B}_t$ and by setting $dt = 2^J$, it follows that

$$\begin{aligned}r_{p,t:t+2^J} = d\log V_t = \boldsymbol{\pi}_t'\mathbf{x}_{t+2^J} + \frac{1}{J+1}\boldsymbol{\iota}'\mathbf{f}_{t+2^J} + 2^J\Bigg\{&\frac{1}{2}\boldsymbol{\pi}_t'\left(\left[\boldsymbol{\sigma}_j\boldsymbol{\sigma}_j'\right] - \left[\boldsymbol{\sigma}_{bj}\boldsymbol{\sigma}_{bj}'\right]\right) \\ -\frac{1}{2}&\left[\boldsymbol{\pi}_t'\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)'\boldsymbol{\pi}_t + 2\boldsymbol{\pi}_t'\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\boldsymbol{\sigma}_b'\boldsymbol{\iota}\frac{1}{J+1}\right] \\ +\frac{1}{2}&\frac{1}{J+1}\boldsymbol{\iota}'\left[\boldsymbol{\sigma}_{bj}\boldsymbol{\sigma}_{bj}'\right] - \frac{1}{2}\frac{1}{J+1}\boldsymbol{\iota}'\boldsymbol{\sigma}_b\boldsymbol{\sigma}_b'\boldsymbol{\iota}\frac{1}{J+1}\Bigg\}.\end{aligned}$$

By using the notation in the VAR, we have

$$\begin{aligned}\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)' &= \Sigma_{\mathbf{x}} \\ \boldsymbol{\sigma}_b\boldsymbol{\sigma}_b' &= \Sigma_{\mathbf{f}} \\ \left[\boldsymbol{\sigma}_{bj}\boldsymbol{\sigma}_{bj}'\right] &= \boldsymbol{\sigma}_{\mathbf{f}}^2 \\ \left[\boldsymbol{\sigma}_j\boldsymbol{\sigma}_j'\right] &= \boldsymbol{\sigma}_{\mathbf{x}}^2 + \boldsymbol{\sigma}_{\mathbf{f}}^2 + 2\boldsymbol{\sigma}_{\mathbf{fx}} \\ \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}_b\right)\boldsymbol{\sigma}_b'\boldsymbol{\iota} &= \boldsymbol{\sigma}_{\mathbf{fx}}\end{aligned}$$

and so

$$
\begin{aligned}
r_{p,t:t+2^J} &\simeq \frac{f_{t:t+2^J}}{J+1} + \boldsymbol{\pi}_t' \mathbf{x}_{t:t+2^J} + 2^J \Bigg\{ \frac{1}{2} \boldsymbol{\pi}_t' \left(\boldsymbol{\sigma}_{\mathbf{x}}^2 + \boldsymbol{\sigma}_{\mathbf{f}}^2 + 2\boldsymbol{\sigma}_{\mathbf{fx}} - \boldsymbol{\sigma}_{\mathbf{f}}^2 \right) \\
&\quad - \frac{1}{2} \left[\boldsymbol{\pi}_t' \Sigma_{\mathbf{x}} \boldsymbol{\pi}_t + 2\boldsymbol{\pi}_t' \boldsymbol{\sigma}_{\mathbf{fx}} \frac{1}{J+1} \right] + \frac{1}{2} \frac{1}{J+1} \underbrace{\boldsymbol{\iota}' \; \boldsymbol{\sigma}_{\mathbf{f}}^2}_{var_t\left(f_{t+2^J}\right)} - \frac{1}{2(J+1)^2} \underbrace{\boldsymbol{\iota}' \boldsymbol{\sigma}_{\mathbf{f}}^2}_{var_t\left(f_{t+2^J}\right)} \Bigg\} \\
&= \frac{f_{t:t+2^J}}{J+1} + \boldsymbol{\pi}_t' \mathbf{x}_{t:t+2^J} + 2^J \left\{ \frac{1}{2} \boldsymbol{\pi}_t' \left(\boldsymbol{\sigma}_{\mathbf{x}}^2 - \Sigma_{\mathbf{x}} \boldsymbol{\pi}_t \right) + \frac{J}{J+1} \left[\boldsymbol{\pi}_t' \boldsymbol{\sigma}_{\mathbf{fx}} + \frac{1}{2} \frac{var_t\left(f_{t+2^J}\right)}{J+1} \right] \right\}.
\end{aligned}
$$

References

A.B. Abel, J.C. Eberly, and S. Panageas. Optimal inattention to the stock market with information costs and transactions costs. *Econometrica*, 81(4):1455–1481, 2013.

R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509, 2004.

N. Barberis. Investing for the long run when returns are predictable. *Journal of Finance*, 55(1):225–264, 2000.

M. Baxter and R.G. King. Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4):575–593, 1999.

J. Bender, R. Briand, D. Melas, and R.A. Subramanian. Foundations of factor investing. *MCSI Research Insight*, 2013.

M.J. Brennan, E.S. Schwartz, and R. Lagnado. Strategic asset allocation. *Journal of Economic Dynamics and Control*, 21(8):1377–1403, 1997.

J.Y. Campbell. Intertemporal asset pricing without consumption data. *American Economic Review*, pages 487–512, 1993.

J.Y. Campbell and L.M. Viceira. Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics*, 114(2):433–495, 1999.

J.Y. Campbell, Y.L. Chan, and L.M. Viceira. A multivariate model of strategic asset allocation. *Journal of Financial Economics*, 67(1):41–80, 2003.

M.M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1): 57–82, 1997.

S.E. Chaudhuri and A.W. Lo. Spectral portfolio theory. *Working paper*, 2016.

M.M. Christensen. On the history of the growth-optimal portfolio. *Machine Learning for Financial Engineering, Imperial College Press, London*, pages 1–80, 2012.

A. Cieslak, A. Morse, and A. Vissing-Jorgensen. Stock returns over the fomc cycle. *Journal of Finance, forthcoming*, 2018.

N. Crouzet, I. Dew-Becker, and C.G. Nathanson. A model of multi-frequency trade. *Northwestern University Working Paper*, 2017.

L.G. Epstein and S.E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, pages 937–969, 1989.

L.G. Epstein and S.E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99 (2):263–286, 1991.

E.F. Fama and K.R. French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465, 1992.

E.F. Fama and K.R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

A. Giovannini and P. Weil. Risk aversion and intertemporal substitution in the capital asset pricing model, 1989.

C.R. Harvey and Y. Liu. Lucky factors. *Working paper*, 2017.

K. Hou, C. Xue, and L. Zhang. Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705, 2015.

D.O. Lucca and E. Moench. The pre-fomc announcement drift. *Journal of Finance*, 70(1): 329–371, 2015.

R.C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, pages 247–257, 1969.

U.A. Müller, M.M. Dacorogna, R.D. Davé, R.B. Olsen, O.V. Pictet, and J.E. von Weizsäcker. Volatilities of different time resolutions: Analyzing the dynamics of market components. *Journal of Empirical Finance*, 4(2):213–239, 1997.

F. Ortu, A. Tamoni, and C. Tebaldi. Long-run risk and the persistence of consumption shocks. *Review of Financial Studies*, 26(11):2876–2915, 2013.

F. Ortu, F. Severino, A. Tamoni, and C. Tebaldi. A persistence-based wold-type decomposition for stationary time series. $Available\ at\ https : //papers.ssrn.com/sol3/papers.cfm?abstract\_id = 1973049$, 2017.

L. Peng and W. Xiong. Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3):563–602, 2006.

E.L. Podkaminer. Risk factors as building blocks for portfolio diversification: The chemistry of asset allocation. *Investment Risk and Performance Newsletter*, 2013(1), 2013.

S.A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13 (3):341–360, 1976.

J.H. Stock and M.W. Watson. Business cycle fluctuations in us macroeconomic time series. *Handbook of macroeconomics*, 1:3–64, 1999.

P. Weil. The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(3):401–421, 1989.