Optimality and Diversifiability of Mean Variance and Arbitrage Pricing Portfolios*

M. Hashem Pesaran
University of Cambridge and USC

Paolo Zaffaroni
Imperial College London

October 17, 2009

Abstract

This paper investigates the limit properties of mean-variance (mv) and arbitrage pricing (ap) trading strategies using a general dynamic factor model, as the number of assets diverge to infinity. It extends the results obtained in the literature for the exact pricing case to two other cases of asymptotic no-arbitrage and the unconstrained pricing scenarios. The paper characterizes the asymptotic behaviour of the portfolio weights and establishes that in the non-exact pricing cases the ap and mv portfolio weights are asymptotically equivalent and, moreover, functionally independent of the factors conditional moments. By implication, the paper sheds light on a number of issues of interest such as the prevalence of short-selling, the number of dominant factors and the granularity property of the portfolio weights.

JEL Classifications: C32, C52, C53, G11
Key Words: Large Portfolios, Factor Models, Mean-Variance Portfolio, Arbitrage Pricing, Market (Beta) Neutrality, Well Diversification

*We are grateful to Laurent Calvet, Jianqing Fan, Enrique Sentana, Eric Renault, Guofu Zhu and participants at various seminars and conferences where a preliminary version of this paper was presented.

1 Introduction

Factor models represent a parsimonious, yet effective way of modelling the conditional joint probability distribution of asset returns, and have formed the basis of the capital asset pricing model of Sharpe (1964) and Lintner (1965), and of the arbitrage pricing theory (APT) of Ross (1976). Parsimony is essential when describing the joint behaviour of a large number N of asset returns, and mainly explains the success of factor models in providing a unified framework for the analysis of large portfolios.

This paper focuses on the limiting properties of mean-variance (`mv`) and arbitrage pricing (`ap`) trading strategies when excess returns follow a general dynamic factor model (including the approximate factor model), with emphasis on the issue of portfolio diversification, as the number of assets diverge to infinity. In particular, we will be concerned with the ability of different portfolio strategies in diversifying the idiosyncratic *versus* the market (or beta) risks. In this respect, it is well known that under exact pricing the `mv` portfolio of Markowitz (1952) diversify the contribution of asset-specific characteristics, as $N \to \infty$, whilst the `ap` portfolio is specifically designed to eliminate the market risk through an orthogonality condition involving the factor loadings (see Ross (1976)). Hence, these two portfolio theories focus on different aspects of risks in portfolio management.

The paper's first contribution is to establish new results on the optimality and diversifiability properties of the `mv` and `ap` strategies in the presence of a large number of assets. It shows that the limiting properties of the two strategies critically depend on the form of the 'no-arbitrage' postulated. Three different formulations are considered: 'exact pricing' where the pricing errors defined as the *differences between expected excess returns and an affine function of factor loadings* are exactly zero, 'asymptotic no-arbitrage pricing' where pricing errors vanish asymptotically at a suitable rate as $N \to \infty$; and the 'unconstrained pricing' case where pricing errors remain even for large N. Recall that, by construction, the `ap` portfolio eliminates the common risk for any N (so long as N is larger than the number of common factors), and under any form of no-arbitrage restrictions. Turning to the `mv` strategy, it is well known that under the exact pricing case, the `mv` portfolio is well-diversified in the sense of Chamberlain (1983), namely that the idiosyncratic risk will vanish, under relatively weak conditions. However, neither portfolios will be well-diversified under the asymptotic no-arbitrage pricing case. Their associated portfolio returns will be positively correlated in the limit, with the `ap` being sub-optimal in terms of *ex ante* Sharpe ratio. Finally, in the unconstrained pricing case the two strategies are asymptotically equivalent with perfectly correlated portfolio returns.

The second main contribution of the paper is to establish the limit behaviour of the portfolio weights for the `mv` and `ap` strategies. It turns out that, although the `ap` portfolio weights are by construction functionally independent of the factors' conditional distribution for any N and any form of no-arbitrage conditions (namely `ap` portfolio is market neutral by construction), the same also applies to the `mv` portfolio weights under the asymptotic and unconstrained pricing cases when $N \to \infty$. Obviously, for the `mv` portfolio the factors' conditional distribution matters for any finite N, but it is of second-order importance as compared with the factor loadings and the idiosyncratic component conditional distribution as $N \to \infty$, which can have practical implications for portfolio management when considering many assets. Moreover, except for the case of exact pricing, the portfolio weights under the two strategies are asymptotically equivalent. Under the exact pricing case the `ap` strategy is sub-optimal and yields a zero vector for the portfolio weights. By contrast, the `mv` strategy that exploits knowledge of the common factors leads to non-zero portfolio weights with a positive Sharpe ratio bounded in N.

The third contribution of the paper is rather technical in nature and provides a set of primitive conditions on the asset return distributions that ensure the more familiar higher level assumptions needed for the validity of the various limit results provided in this paper. Interestingly, this includes primitive conditions that imply the various form of no-arbitrage here considered. Our assumptions are more primitive than hitherto maintained in the asset pricing literature and allow for a considerable degree of cross section dependence in factor loadings and error variances.

Our contributions relate naturally to the vast literature that exists on the CAPM and APT when there are a countably infinite number of primitive assets. Hubermann (1982) provided an alternative proof of the APT theorem (see also Connor (1984), Ingersoll (1984), Grinblatt and Titman (1987), Green and Hollifield (1992)). Chamberlain and Rothschild (1983) studied the implications of no-arbitrage for the mean-variance frontier, as N tends to infinity, in a general setting. They also extended the APT result of Ross (1976) to the case where asset returns follow an approximate factor structure. The latter extends the exact factor model by permitting certain (limited) degree of correlation across the idiosyncratic component of asset returns. Within the same approximate factor structure framework, Chamberlain (1983) clarified the relationship between exact pricing, diversification of the mean-variance efficient frontier and the mutual fund separation theorem. He also sharpened the APT result of Chamberlain and Rothschild (1983) providing bounds on the squared sum of the pricing errors. Hansen and Richard (1987) extended the analysis of Chamberlain (1983) and Chamberlain and Rothschild (1983) to a conditional Hilbert space setting, providing the foundation for a dynamic (conditional) version of APT. Subsequently, Green and Hollifield (1992) further clarified the relationships that exist between diversification (define in terms of the sup-norm of the portfolio weights rather than in terms of Euclidean norm) and mean-variance efficiency in a general setting. Employing a factor structure, these authors also provided a further generalization showing that even the approximate factor structure is too stringent for the APT to hold. Sentana (2004) considers a dynamic APT framework but with the aim of establishing the statistical properties of static and dynamic factor representing portfolios.

To summarize, the literature has considered the limiting properties of the `mv` strategy under exact pricing, and that of the `ap` under the no-arbitrage pricing cases. The close relationships that exist between the two strategies, in terms of limiting behaviour of portfolio weights and portfolio return characteristics, which are evident once the strategies are evaluated relative to the same no-arbitrage setting, appear to have been overlooked. This is particularly true of the asymptotic market neutrality property of the `mv` strategy. The present paper addresses these issues in some detail. However, it is important to note that our results assume perfect knowledge of model specification and parameters. The issues of estimation and model specification must also be addressed. But this falls outside the scope of the present paper, and need to be addressed in future research.[1]

The remainder of the paper is organized as follows. Section 2 introduces the concepts, sets out the dynamic factor model and the required assumptions, spelling out the different forms of no-arbitrage pricing analyzed. Section 2.1 discusses primitive conditions on the factor model. Section 3.1 defines the `mv` and `ap` trading strategies and formally establishes their relationship. The main results are in Section 3.2 where we show, under the various forms of no-arbitrage, the diversifiability and optimality properties of the `mv` and `ap` strategies. Section 4 elaborates and discusses the implications of the theoretical results. As a by-product, our results also have some bearing on a number of issues discussed in the literature, such as short-selling, the number of dominant factors, and the granularity property of the portfolio weights. We also provide asymptotic results on diversifiability and optimality of two sub-optimal (yet popular) strategies, namely the global minimum variance and the equally weighted portfolios.Section 5 provides a brief summary and offers some concluding remarks. Mathematical proofs are collected in three appendices.

[1]For a survey of the econometric issues associated with portfolio choice problems see Brandt (2004).

Notations: $\to_p$ and $\to_d$ denote convergence in probability and in distribution, respectively; $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbf{e}_i$ is its ith column; $\mathbf{A} > (\geq) \, 0$ means that the generic square matrix $\mathbf{A}$ is positive (semi) definite, with ith row $\mathbf{a}_i'$. $\rho(\mathbf{A})$ and $\text{tr}(\mathbf{A})$ denote the spectral norm and trace of $\mathbf{A}$, respectively. The index $i = 1, 2, ..., N$ will be used to denote an asset, and $j = 1, 2, ..., k$ to denote a factor.

2 Factor model: definitions and assumptions

We assume the N-dimensional vector $\mathbf{r}_t = (r_{1t}, r_{2t}, ..., r_{Nt})'$ of asset returns can be characterized by the following linear factor model

$$\mathbf{r}_t - r_{0,t-1}\mathbf{e} = \boldsymbol{\mu}_{t-1} + \mathbf{B}\mathbf{z}_t + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $r_{0,t}$ is the risk-free rate, $\mathbf{e} = (1, 1, ..., 1)'$ is an $N \times 1$ vector of ones, $\mathbf{z}_t = (z_{1t}, z_{2t}, ..., z_{kt})'$ is the $k \times 1$ vector of possibly latent common factors assumed with zero (conditional) mean without loss of generality as specified below, $\mathbf{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_N)' = (\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, ..., \boldsymbol{\beta}^k)$ is an $N \times k$ matrix of factor loadings with ith row $\boldsymbol{\beta}_i'$ and jth column $\boldsymbol{\beta}^j$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{Nt})'$ is an $N \times 1$ vector of idiosyncratic components, and the $N \times 1$ vector $\boldsymbol{\mu}_{t-1}$ represents the conditional mean of the vector of excess returns, $\mathbf{r}_t - r_{0,t-1}\mathbf{e}$. Throughout, it will be assumed that $k < N$ remains fixed as $N \to \infty$.

This paper considers the implications of the factor model for portfolio optimization as $N \to \infty$, at a given point in time, t. More formally, let $(\mathcal{O}_t, \mathcal{A}_t, \mathcal{P}_t)$ be the probability space for cross sectionally-invariant events common to all assets, and let $(\mathcal{O}_{i.}, \mathcal{A}_{i.}, \mathcal{P}_{i.})$ be the probability space for *time-invariant* events that are specific to asset i. Denote by $\mathcal{A}_{it}$ the sigma-field for *time- and cross sectionally-varying* events that are specific to asset i. Then set $\mathcal{A}_{(N),t} \equiv \left(\cup_{i=1}^N \mathcal{A}_{i.}\right) \cup \left(\cup_{i=1}^N \mathcal{A}_{it}\right) \cup \mathcal{A}_t$, so that $r_t \in \mathcal{A}_{(N),t}$ for any N and t. This paper establishes the asymptotic properties of portfolio weights and excess returns as $N \to \infty$ conditional on $\mathcal{A}_{t-1}$. Hence, any reference to the time index of the variables involved is redundant but it has been nevertheless maintained simply to clarify that our results apply to factor models with time-varying conditional heteroskedasticity. Also, strictly speaking, all the vector and matrices in (1) must also be indexed by N. But the index N is suppressed to simplify the exposition. For future references it is also helpful to note that (1) can be written similarly as

$$\mathbf{r}_t - r_{0,t-1}\mathbf{e} = \boldsymbol{\mu}_{t-1} + \sum_{j=1}^{k} \beta^j z_{jt} + \boldsymbol{\varepsilon}_t, \tag{2}$$

and, for $i = 1, 2, ..., N$,

$$r_{it} - r_{0,t-1} = \mu_{i,t-1} + \boldsymbol{\beta}_i' \mathbf{z}_t + \varepsilon_{it}. \tag{3}$$

We now specify a set of minimal assumptions needed for our results. These assumptions are often expressed as limits of linear and quadratic forms, for $N \to \infty$. For this reason, we also provide more primitive conditions for such limits to hold. These conditions are provided separately because, although giving sufficient conditions, appear relatively cumbersome.

Assumption 1 *(common and idiosyncratic innovations) At any given point in time t*

$$\mathbf{z}_t \mid \mathcal{A}_{(N),t-1} \sim (0, \boldsymbol{\Omega}_{t-1}), \;\; \boldsymbol{\varepsilon}_t \mid \mathcal{A}_{(N),t-1} \sim (0, \mathbf{G}_{t-1}),$$

where '$\mid \mathcal{A}_{(N),t\text{-}1}$' means conditional on $\mathcal{A}_{(N),t\text{-}1}$. Moreover the covariance matrices $\boldsymbol{\Omega}_{t-1} > 0$ and $\mathbf{G}_{t-1} > 0$ are, respectively, of dimension $k \times k$ and $N \times N$ for a fixed k and any finite $N > k$.

Hereafter, to simplify the notations we set

$$\mathbf{H}_t \equiv \mathbf{G}_t^{-1}. \tag{4}$$

Below we will refer to the ith row of $\boldsymbol{\varepsilon}_\mathbf{t}$ and $\mathbf{H}_t$ as ε_{it}, and $\mathbf{h}_{it}'$, respectively. The assumption of zero (conditional) mean for $\mathbf{z}_t$ is without loss of generality. However, note that we are not setting their covariance matrix equal to the identity matrix, as is typically done when specifying factor models, since we are allowing for conditional heteroskedasticity through $\boldsymbol{\Omega}_{t-1}$. Indeed, a perfectly equivalent setting to ours would be when $E(\mathbf{z}_t \mathbf{z}_t' \mid \mathcal{A}_{(N),t-1}) = \mathbf{I}_k$ with time-varying factor loadings $\mathbf{B}_{t-1}$. It is worth noting that Assumption 1 can also be specified with respect to the sigma-algebra spanned by the unobserved information set $\mathbf{z}_{t-s}$, $s > 0$, without affecting our main conclusions.

Concerning the idiosyncratic components, we allow the entries of $\boldsymbol{\varepsilon}_t$ to be cross sectionally weakly dependent at a given point in time, to be made more precise below. Thus we permit the off-diagonal elements of $\mathbf{G}_{t-1}$ to be non-zero so long as they are subject to certain bounded conditions. The results that follow do not depend on a particular specification of the volatility model characterizing the asset returns. Moreover, the factors can either be observable or non-observable. As a consequence, $\boldsymbol{\Omega}_{t-1}$ and $\mathbf{G}_{t-1}$ could belong to the multivariate stochastic volatility class as well as to the generalized autoregressive conditional heteroskedasticity class of volatility models.

Regarding the factor loadings, we consider the case where the elements of $\mathbf{B}$ are random variates satisfying the following limit condition:

Assumption 2 *(factor loadings) As $N \to \infty$*

$$N^{-1}\mathbf{B}'\mathbf{e} \to_p \boldsymbol{\mu}_\beta = (\mu_{\beta 1}, \mu_{\beta 2}, ..., \mu_{\beta k})' \neq \mathbf{0}, \tag{5}$$

$$N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B} \to_p \mathbf{D}_t > 0. \tag{6}$$

For any $i = 1, 2, ..., N$, each $\boldsymbol{\beta}_i$ is mutually independent from any ε_{it} and $\mathbf{h}_{it}$. (7)

Condition (6) is satisfied if all the k common factors are pervasive or strong in the sense of Connor (1984) or Chudik, Pesaran, and Tosetti (2009), although these concepts are typically defined in terms of the limiting behaviour of $N^{-1}\mathbf{B}'\mathbf{B}$ or the column matrix norm of $\mathbf{B}$, and do not involve the $\mathbf{H}_t$ matrix. But it is clear that under certain bounded conditions on the eigenvalues of $\mathbf{H}_t$, to be discussed below, $N^{-1}\mathbf{B}'\mathbf{B}$ and $N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B}$ will have the same limiting behaviour. A simple special case is when $\boldsymbol{\beta}_i$ are *i.i.d.* with finite second order moments and $\mathbf{H}_t$ is diagonal satisfying $N^{-1}\mathbf{e}'\mathbf{H}_t\mathbf{e} \to_p a_t$. In this case $\mathbf{D}_t = a_t(\text{cov}(\boldsymbol{\beta}_i) + \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta')$. Obviously, when the mean of the jth factor loading $\mu_{\beta j}$ is non zero, then $N^{-1}\boldsymbol{\beta}^{j\prime}\boldsymbol{\beta}^j \geq (N^{-1}\boldsymbol{\beta}^{j\prime}\mathbf{e})^2 \to_p (\mu_{\beta j})^2 > 0$, and the jth factor is strong since $\boldsymbol{\beta}^{j\prime}\boldsymbol{\beta}^j$ diverges to infinity at an appropriate rate. But the reverse is not necessarily true and the factor, z_{jt}, can be strong even if $\mu_{\beta j} = 0$, so long as $N^{-1}\Sigma_{i=1}^N \left|\beta_{ij}\right| \to_p K > 0$, where K is a fixed finite constant.[2] Assumption 2 is an ergodicity assumption over the cross section. It is much weaker than the *i.i.d.* assumption typically made when considering random factor loadings. The results can be generalized further to the case of heterogeneous yet non-random $\boldsymbol{\beta}_i$. General pimitive conditions on factor loadings will be discussed below.

Assumption 3 *(conditional mean returns)*

$$\boldsymbol{\mu}_{t-1} \equiv E(\mathbf{r}_t - r_{0,t-1}\mathbf{e} \mid \mathcal{A}_{(N),t-1}) = \mathbf{B}\boldsymbol{\lambda}_{t-1} + \mathbf{v}_{t-1}, \tag{8}$$

where $\boldsymbol{\lambda}_{t-1}$ defines the vector of factor risk premia, and $\mathbf{v}_t = (v_{1t}, v_{2t}..., v_{Nt})$ is the vector of pricing errors satisfying:

For any $i = 1, 2, ..., N,$ each v_{it} is mutually independent from any ε_{it}, and $\mathbf{h}_{i,t}$. (9)

Note that the linear projection (8) holds in population, implying that the individual entries of $\mathbf{v}_t$ and $\boldsymbol{\lambda}_t$ do *not* depend on N. Instead, rather than

[2] Note that for z_{jt} to be strong it is not sufficient that $\Sigma_{i=1}^N \left|\beta_{ij}\right|$ diverges to infinity but it must do so at rate $O_p(N)$. See Chudik, Pesaran, and Tosetti (2009) for further details.

(8) the APT literature typically starts from (see Ross (1976), Hubermann (1982), Ingersoll (1984) among others)

$$\boldsymbol{\mu}_{t-1} = \mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1} + \hat{\mathbf{v}}_{t-1}, \tag{10}$$

where

$$\hat{\boldsymbol{\lambda}}_{t-1} = \left(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{H}_{t-1}\boldsymbol{\mu}_{t-1} \tag{11}$$

is the (generalized) least squares estimator and $\hat{\mathbf{v}}_{t-1}$ is the regression residual, which satisfies $\mathbf{B}'\mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1} = \mathbf{0}$ or, alternatively,

$$\hat{\mathbf{v}}_{t-1} = [\mathbf{I}_N - \mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}]\boldsymbol{\mu}_{t-1} = [\mathbf{I}_N - \mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}]\mathbf{v}_{t-1}. \tag{12}$$

Clearly (10) can always be obtained, for non-singular $\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}$, although the statistical properties of $\hat{\boldsymbol{\lambda}}_{t-1}, \hat{\mathbf{v}}_{t-1}$ for large N will require formal assumptions on (8). Moreover, individual entries of $\hat{\boldsymbol{\lambda}}_{t-1}$, and $\hat{\mathbf{v}}_{t-1}$ will now depend on N and our notion of factor risk premia, $\boldsymbol{\lambda}_{t-1}$, will be equivalent to $\boldsymbol{\lambda}_{\infty,t-1}$ of Ingersoll (1984, Theorem 3), namely the limit of $\hat{\boldsymbol{\lambda}}_{t-1}$ as $N \to \infty$. (Here the time index has been added to Ingersoll's notation). It is more convenient to base the analysis on $\boldsymbol{\lambda}_{t-1}$ and $\mathbf{v}_{t-1}$ directly since our aim is to precisely characterize the limit of the portfolio weights and other aspects of the distribution of the `mv` portfolios return, unlike the APT for which assumptions on $\hat{\mathbf{v}}_{t-1}$ suffice, as discussed below. For future references it is important always to bear in mind that $\mathbf{B}'\mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1} = \mathbf{0}$ holds for any $N > k$, but the magnitude of $\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}$ depends on N and the degree of pricing errors assumed under different arbitrage conditions. See Assumption 5 below.

Furthermore, given the focus of our analysis, in what follows we take the specification of $\boldsymbol{\mu}_{t-1}$, especially its $\mathbf{v}_{t-1}$ component, as given. Using (8) in (1) the factor model can be written as

$$\mathbf{r}_t - r_{0,t-1}\mathbf{e} = \mathbf{B}\mathbf{f}_t + \mathbf{u}_t, \tag{13}$$

where $\mathbf{f}_t = \mathbf{z}_t + \boldsymbol{\lambda}_{t-1}$, and $\mathbf{u}_t = \boldsymbol{\varepsilon}_t + \mathbf{v}_{t-1}$. Therefore, $\boldsymbol{\lambda}_{t-1}$ and $\mathbf{v}_{t-1}$ can also be viewed as the predictable components of the common and the idiosyncratic factors, respectively. As illustrated below, arbitrage opportunities arise when one or the other of these components are non-zero for a finite N or as $N \to \infty$. The conditional variance covariance matrices of the two components are given by $\boldsymbol{\Omega}_{t-1}$ and $\mathbf{G}_{t-1}$, as before.

To derive the limiting behavior of the portfolios (as $N \to \infty$) under alternative arbitrage opportunities we further require the following assumption:

Assumption 4 *(limit conditions) At any given point in time, t, as $N \to \infty$,* *we have*

$$N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B} \to_p \mathbf{F}_t \geq 0, \tag{14}$$

$$N^{-1}\mathbf{e}'\mathbf{H}_t\mathbf{e} \to_p a_t > 0, \tag{15}$$

$$\mathbf{B}'\mathbf{H}_t\mathbf{e}_i = O_p(1), \ \ \mathbf{e}'\mathbf{H}_t\mathbf{e}_i = O_p(1), \ \textit{for any } i, \tag{16}$$

where all the limits are finite almost surely (a.s.). We also require that $a_t\boldsymbol{\mu}'_\beta\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta < 1$ a.s.

The common feature of the limits presented in Assumption 4 is that they involve, possibly weighted, averages of the elements of $\mathbf{H}_t$. In particular, they impose implicitly an upper bound on the speed with which the maximum eigenvalue of $\mathbf{H}_t$ (which coincide with the smallest eigenvalue of $\mathbf{G}_t$) could diverge to infinity. This is clearly seen from condition (15): assuming for illustrative purposes that $\mathbf{H}_t$ is diagonal, with $h_{ii,t}$ as its $(i,i)^{th}$ entry, then (15) allows $max_{1\leq i\leq N}h_{ii,t} = o_p(N)$. Condition (14) requires a further constraint on the speed of divergence of $max_{1\leq i\leq N}h_{ii,t}$ which can now be at most $o_p(N^{\frac{1}{2}})$. Even this case is much weaker than $max_{1\leq i\leq N}h_{ii,t} \leq K < \infty$, for some constant K, implied by the approximate factor model often adopted in the literature (see Chamberlain and Rothschild (1983)). Green and Hollyfield (1992) were the first to note that, insofar as optimal asset allocation is concerned, a degree of cross-sectional dependence stronger than the one implied by the approximate factor structure is permitted. When $\mathbf{H}_t$ is non-diagonal, the previous discussion applies to its largest eigenvalue. Condition (14), although not necessarily implied by (6), does not require $\mathbf{F}_t > 0$ and it is a technical assumption needed by our Lemma B in the Appendix. Condition (16) imposes a finite upper bound on each of the columns of $\mathbf{H}_t$ and is therefore much stronger than (6) and (15) that are expressed in terms of averages. In particular, the second condition of (16) is satisfied by an approximate factor structure. The condition $a_t\boldsymbol{\mu}'_\beta\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta < 1$ rules out the case where $k = 1$ and the factor loading coefficient β_{i1} is homogeneous across i. In this case $\boldsymbol{\mu}_\beta = \mu_{\beta 1}\mathbf{e}$, and $\mathbf{D}_t = a_t\mu_{\beta 1}^2$ and the condition is clearly violated. The homogeneous slope case with $k > 1$ is ruled out under Assumption 2, since in this case $\mathbf{D}_t = a_t\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta$ which is not a positive definite matrix contrary to (6).

We now present a set of assumptions on the pricing errors $\mathbf{v}_t$ that, in turn, characterize the consequences of no-arbitrage pricing, as well as the unconstrained case where asymptotic arbitrage possibilities are permitted.

Assumption 5 *(arbitrage conditions) At any given point in time, t, either one of the following three set of conditions hold, where all the limits below are a.s. finite*:

(i) exact no-arbitrage pricing: for any $N > k$

$$\mathbf{v}_t = \mathbf{0}. \tag{17}$$

(ii) asymptotic no-arbitrage pricing: as $N \to \infty$

$$N^{1/2}\mathbf{e}_i'\mathbf{H}_t\mathbf{v}_t = O_p(1) \text{ for any } i, \tag{18}$$
$$N^{-1/2}\mathbf{e}'\mathbf{H}_t\mathbf{v}_t \to_p c_t, \tag{19}$$
$$\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t \to_p d_t. \tag{20}$$
$$\mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{v}_t = O_p(1). \tag{21}$$

(iii) unconstrained pricing: as $N \to \infty$

$$\mathbf{e}_i'\mathbf{H}_t\ \mathbf{v}_t = O_p(1) \textit{ for any } i, \tag{22}$$
$$N^{-1}\mathbf{e}'\mathbf{H}_t\mathbf{v}_t \to_p c_t, \tag{23}$$
$$N^{-1}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t \to_p d_t. \tag{24}$$
$$N^{-1}\mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{v}_t = O_p(1). \tag{25}$$

When either (ii) or (iii) hold, then

$$d_t a_t - c_t^2 > 0, \quad \textit{almost surely.} \tag{26}$$

When the exact no-arbitrage pricing condition (17) holds, model (1) can be obtained, when $k = 1$, as the traditional CAPM of Sharpe (1964) and Lintner (1965) or, when $k > 1$, as a version of the intertemporal CAPM of Merton (1973).

Turning to the asymptotic no-arbitrage pricing set-up, conditions (19)-(20) represent the strongest prediction of the APT of Ross (1976), when a risk free asset is available. In particular, the APT implies [see Ingersoll (1984, Theorem 1) for instance] that for any $N > k$

$$\hat{\mathbf{v}}_t'\mathbf{H}_t\hat{\mathbf{v}}_t = \mathbf{v}_t'[\mathbf{H}_t - \mathbf{H}_t\mathbf{B}(\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t]\mathbf{v}_t = O_p(1), \tag{27}$$

where $\hat{\mathbf{v}}_t$ is defined in (10). Since (6) holds, (27) requires $\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t = O_p(1)$ and $\mathbf{v}_t'\mathbf{H}_t\mathbf{B} = O_p(N^{\frac{1}{2}})$ which in turn, by the mutual independence between the v_{it} and $\boldsymbol{\beta}_i$ and given (5), implies $\mathbf{v}_t'\mathbf{H}_t\mathbf{e} = O_p(N^{\frac{1}{2}})$. Our conditions (18)-(19)-(20) represent slightly stronger version of these, here needed to establish convergence (of certain linear and quadratic forms) than simply establishing upper bounds such as (27).

Clearly, when (17) holds then (18)-(19)-(20) trivially apply but the latter also apply when the elements of $\mathbf{v}_t$, converge to zero quicker than $N^{-\frac{1}{2}}$. Finally, concerning the unrestricted case, conditions (23) and (24) also require the elements of $\mathbf{v}_t$ not to grow, if at all, too fast as compared with N. The limit c_t in condition (23) is bounded, in absolute value, by $(a_t\, d_t)^{\frac{1}{2}}$. The limit d_t in condition (24) is finite whenever (15) holds and $N^{-1}\mathbf{v}_t'\mathbf{v}_t$ has a finite limit. Condition (26) implies that both c_t and d_t are non-zero. But note that this is not needed in the case where $\mathbf{v}_t$ is a non-degenerate random variable. It is easily seen that $a_t d_t - c_t^2 = (N^{-1}\mathbf{e}'\mathbf{H}_t\mathbf{e})\,(N^{-1}\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t)\,(1-\rho_{t\tilde{v}\tilde{e}}^2) \geq 0$, where $\rho_{t\tilde{v}\tilde{e}}$ is the correlation coefficient of $\mathbf{H}_t^{1/2}\mathbf{e}$ and $\mathbf{H}_t^{1/2}\mathbf{v}_t$.

The factor model, (1), implies the well-known form of the asset return conditional variance-covariance matrix:

$$E\left[(\mathbf{r}_t - r_{0,t-1}\mathbf{e} - \boldsymbol{\mu}_{t-1})(\mathbf{r}_t - r_{0,t-1}\mathbf{e} - \boldsymbol{\mu}_{t-1})' \mid \mathcal{A}_{(N),t-1}\right] = \boldsymbol{\Sigma}_{t-1} = \mathbf{B}\boldsymbol{\Omega}_{t-1}\mathbf{B}' + \mathbf{G}_{t-1}. \tag{28}$$

Thus model (1) nests the various factor models with time-varying conditional second moment proposed in the econometrics literature. See among many others Diebold and Nerlove (1989), King, Sentana, and Wadhwani (1994), Chib, Nardari, and Shephard (2002), Fiorentini, Sentana, and Shephard (2004), Connor, Korajczyk, and Linton (2006), and Doz and Renault (2006). These papers, which focus on estimation of volatility factor models, in particular when $\mathbf{z}_t$ is not observable, all assume constant conditional first-order moments. On the other hand, the finance literature dealing with factor models-based asset allocation assumes homoskedastic factors whereby $\boldsymbol{\Omega}_{t-1} = \boldsymbol{\Omega}$, often normalized to be equal to the identity matrix (see among many others Pesaran and Timmermann (1995) and Kandel and Stambaugh (1996)). A few contributions analyze asset allocation problems allowing for volatility dynamics but impose constant conditional means (see, for instance, Aguilar and West (2000) and Fleming, Kirby, and Ostdiek (2001)). Only recently, a limited number of studies have considered time variations in both the first and second conditional moments of asset returns (see for instance Johannes, Polson, and Stroud (2002) and Han (2006)). Model (1) nests all of the above specifications.

2.1 Primitive assumptions

Assumptions 2, 4 and 5 involve high level assumptions on certain linear and quadratic forms to have certain limiting values. We now present a set of primitive regularity conditions that ensure the limits, as $N \to \infty$, of the bilinear and quadratic forms stated in the above assumptions. Such conditions often involve the expectation operator, conditional on the sigma-algebra of cross sectionally-invariant events, $\mathcal{A}_t$, of functions of $\mathbf{B}, \mathbf{H}_t$, and $\mathbf{v}_t$, defined as follows.[3] For a generic random variable X, define the cross-sectional expectation operator as $E_t(\cdot) \equiv E(\cdot \mid \mathcal{A}_t)$, and the cross-sectional variance operator as $\text{var}_t(X) \equiv E\left\{[X - E_t(X)]^2 \mid \mathcal{A}_t\right\}$.[4] Although, high level assumptions provide important simplifications and have been used in the literature, we believe it is important that the implications of such assumptions for the primitives of the asset pricing model (the factor loadings, the factor covariances, the parameters of the idiosyncratic errors) are explored and investigated. To this end first recall that $\mathbf{B} = (\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, ..., \boldsymbol{\beta}^k)$ and $\boldsymbol{\mu}_\beta = (\mu_{\beta 1}, \mu_{\beta 2}, ..., \mu_{\beta k})'$ and, denoting by $\otimes$ denotes the Kronecker product, set

$$\begin{aligned}
&\boldsymbol{\mu}_\beta^j = E\boldsymbol{\beta}^j, \, \boldsymbol{\Omega}_{\beta j} = E(\boldsymbol{\beta}^j - \boldsymbol{\mu}_\beta^j)(\boldsymbol{\beta}^j - \boldsymbol{\mu}_\beta^j)' = \mathbf{P}_{\beta j}\mathbf{P}_{\beta j}', \\
&\boldsymbol{\Sigma}_{\beta jl} = E\left(\boldsymbol{\beta}^j\boldsymbol{\beta}^{j\prime} \otimes \boldsymbol{\beta}^l\boldsymbol{\beta}^{l\prime}\right), \\
&\boldsymbol{\Omega}_{\beta jl} = E\left(\boldsymbol{\beta}^j\boldsymbol{\beta}^{j\prime} \otimes \boldsymbol{\beta}^l\boldsymbol{\beta}^{l\prime}\right) - E\left(\boldsymbol{\beta}^j \otimes \boldsymbol{\beta}^l\right) E\left(\boldsymbol{\beta}^{j\prime} \otimes \boldsymbol{\beta}^{l\prime}\right),
\end{aligned}$$

where we assume $\boldsymbol{\Omega}_{\beta j} > 0$ (namely that the jth factor loadings are not homogeneous across i).

Assumption 6 *(factor loadings)* For all $1 \leq j, l \leq k$, as $N \to \infty$

$$N^{-1}\boldsymbol{\mu}_\beta^{j\prime}\mathbf{e} \to \mu_{\beta j}, \tag{29}$$
$$N^{-1}E_t\left(\boldsymbol{\beta}^{j\prime}\mathbf{H}_t\boldsymbol{\beta}^l\right) \to_p d_{t,jl}, \tag{30}$$
$$N^{-1}E_t\left(\boldsymbol{\beta}^{j\prime}\mathbf{H}_t^2\boldsymbol{\beta}^l\right) \to_p f_{t,jl}, \tag{31}$$
$$0 < N^{-1}\boldsymbol{\mu}_\beta^{j\prime}\boldsymbol{\mu}_\beta^j = O(1), \tag{32}$$
$$\rho(\mathbf{P}_{\beta j}) = o(N^{\frac{1}{4}}) \tag{33}$$
$$\text{tr}(\boldsymbol{\Omega}_{\beta j}) = O(N), \tag{34}$$
$$\rho(\boldsymbol{\Omega}_{\beta j}^{-1}) = O(1), \tag{35}$$
$$\rho(\boldsymbol{\Omega}_{\beta jl}) = O(1), \, j \neq l, \tag{36}$$
$$\text{tr}(\boldsymbol{\Sigma}_{\beta jl}) = O(N^2). \tag{37}$$

Here $\{d_{t,jl}\}$ and $\{f_{t,jl}\}$ denote the $(j,l)^{th}$ element of $\mathbf{D}_t$ and $\mathbf{F}_t$, respectively. When the factor loadings are identically distributed across assets then (29) is immediately satisfied since $\boldsymbol{\mu}_\beta^j = \mu_{\beta j}\mathbf{e}$. The factor loadings cannot have zero

[3] For a formal characterization of $\mathcal{A}_t$, see Section 2.

[4] Obviously, for any function of the factor loadings, $\mathbf{B} = \mathbf{B}(\omega_N)$, $\boldsymbol{\omega}_N \in \mathcal{O}_N$, and a measurable mapping, $\mathbf{B}(\cdot)$, then $E_t(\cdot)$ coincides with the unconditional expectations operator $E(\cdot)$. The same applies to any time-invariant element of the model.

mean for (32) to follow. Since $\rho(\mathbf{\Omega}_{\beta j}) = \rho^2(\mathbf{P}_{\beta j})$ condition (33) limits the degree of cross-sectional dependence across assets for the j^{th} factor loading which, however, can still be substantial since the maximum eigenvalue of $\mathbf{\Omega}_{\beta j}$ does not need to be bounded. Condition (35) involves the minimum eigenvalue of $\mathbf{\Omega}_{\beta j}$ and (34) ensures that only a limited proportion of its eigenvalues can diverge with N. The rate imposed by condition (37) takes care of the fact that $\mathbf{\Sigma}_{\beta jl}$ is of dimension $N^2 \times N^2$.

Let

$$\boldsymbol{\mu}_{ht} = E_t\left[\text{vech}(\mathbf{H}_t)\right], \, \mathbf{\Omega}_{ht} = E_t\left[(\text{vech}(\mathbf{H}_t) - \boldsymbol{\mu}_{ht})(\text{vech}(\mathbf{H}_t) - \boldsymbol{\mu}_{ht})'\right],$$

and

$$\boldsymbol{\mu}_{h^2,t} = E_t[\text{vech}(\mathbf{H}_t^2)], \, \mathbf{\Omega}_{h^2,t} = E_t\left[(\text{vech}(\mathbf{H}_t^2) - \boldsymbol{\mu}_{h^2,t})(\text{vech}(\mathbf{H}_t^2) - \boldsymbol{\mu}_{h^2,t})'\right],$$

where vec$(\mathbf{A})$ stacks in a column vector all the elements of the generic matrix $\mathbf{A}$. When $\mathbf{A}$ is symmetric it is useful to consider only its distinct elements by using the vech$(\mathbf{A})$ operator. This is relevant when $\mathbf{A}$ is a random matrix and one needs to evaluate the covariance matrix of its elements which would be singular if duplications are not removed. The two operators are linked by the duplication matrix $\mathbf{D}_N$ (see Magnus and Neudecker (2001, p.48)) yielding vec$(\mathbf{A}) = \mathbf{D}_N$vech$(\mathbf{A})$ for a $N \times N$ symmetric matrix $\mathbf{A}$.

Assumption 7 *(idiosyncratic covariance matrix)*

$$\begin{aligned}
&N^{-1}\mathbf{e}'\mathbf{D}_N\boldsymbol{\mu}_{ht} \to_p a_t, && (38)\\
&\boldsymbol{\mu}_{ht}'\mathbf{D}_N'\mathbf{D}_N\boldsymbol{\mu}_{ht} = o_p(N^2), && (39)\\
&\rho(\mathbf{D}_N\mathbf{\Omega}_{ht}\mathbf{D}_N') = o_p(1), && (40)\\
&E_t\left[\rho^2(\mathbf{H}_t)\right] = O_p(1), && (41)\\
&\boldsymbol{\mu}_{h^2,t}'\mathbf{D}_N'\mathbf{D}_N\boldsymbol{\mu}_{h^2,t} = o_p(N^2), && (42)\\
&\rho(\mathbf{D}_N\mathbf{\Omega}_{h^2,t}\mathbf{D}_N') = o_p(1), && (43)\\
&E_t\left[\rho^2(\mathbf{H}_t^2)\right] = O_p(1). && (44)
\end{aligned}$$

Conditions (38) and (39) place limits on the magnitude of the off-diagonal elements of $\mathbf{H}_t$ as $N \to \infty$. These conditions are satisfied, for instance, when $\mathbf{H}_t$ is diagonal and $E_t(\mathbf{H}_t)$ has bounded maximum eigenvalue. In turn this last condition is implied by (41) and Jensen's inequality since the maximum eigenvalue is a convex function. Condition (40) states that the cross-sectional variation of the elements of $\mathbf{H}_t$ is asymptotically negligible for large N. The same comments apply to $\mathbf{H}_t^2$ when conditions (42)-(43)-(44) hold. As indicated below, we make stronger assumptions on $\mathbf{H}_t$ in order to permit weaker conditions on the factor loadings, $\mathbf{B}$, in particular to allow the factor loadings to be (cross-sectionally) strongly dependent.

Define

$$\boldsymbol{\mu}_{vt} = E_t(\mathbf{v}_t), \ \boldsymbol{\Omega}_{vt} = E_t[(\mathbf{v}_t - \boldsymbol{\mu}_{vt})(\mathbf{v}_t - \boldsymbol{\mu}_{vt})'] = \mathbf{P}_{vt}\mathbf{P}_{vt}'.$$

In terms of the properties of the pricing error $\mathbf{v}_t$ we need to distinguish between the asymptotic no-arbitrage pricing and the unconstrained case.

Assumption 8 *(asymptotic no-arbitrage pricing)*

$$N^{-1/2}\mathbf{e}'E_t(\mathbf{H}_t\mathbf{v}_t) \to_p c_t, \tag{45}$$

$$E_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t) \to_p d_t, \tag{46}$$

$$\boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt} = O_p(1), \tag{47}$$

$$\rho(\mathbf{P}_{vt}) = o_p(N^{-\frac{1}{4}}), \tag{48}$$

$$\mathrm{tr}(\boldsymbol{\Omega}_{vt}) = O_p(1), \tag{49}$$

$$\rho(\boldsymbol{\Omega}_{vt}^{-1}) = O_p(N^{\frac{1}{2}}). \tag{50}$$

Assumption 9 *(unconstrained case)*

$$N^{-1}\mathbf{e}'E_t(\mathbf{H}_t\mathbf{v}_t) \to_p c_t, \tag{51}$$

$$N^{-1}E_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t) \to_p d_t, \tag{52}$$

$$\boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt} = O_p(N), \tag{53}$$

$$\rho(\mathbf{P}_{vt}) = o_p(N^{\frac{1}{4}}), \tag{54}$$

$$\mathrm{tr}(\boldsymbol{\Omega}_{vt}) = O_p(N), \tag{55}$$

$$\rho(\boldsymbol{\Omega}_{vt}^{-1}) = O_p(1). \tag{56}$$

The main differences between Assumption 8 and 9 are due to the fact that each element of the vector $\mathbf{v}_t$ will vanish at rate $O_p(N^{-\frac{1}{2}})$ in the former case whereas these will be $O_p(1)$ in the latter.

We have established that the high level conditions in Assumptions 2, 4 and 5 are implied by the primitive assumptions 6, 7, 8 and 9. See the Appendix, Part B, for a proof.

3 Asymptotic results for portfolio weights and portfolio returns

3.1 Trading strategies

We present the trading strategies of interest establishing a useful correspondence between their weights and associated portfolio returns. We consider two different trading strategies. The first is the mean-variance portfolio, defined by the solution to the following optimization problem[5]

$$\mathbf{w}_{t-1}^{mv} = \text{argmax}_{\mathbf{w}} \left(\mathbf{w}'\boldsymbol{\mu}_{t-1} + r_{0,t-1} - \frac{\kappa_{t-1}}{2} \mathbf{w}'\boldsymbol{\Sigma}_{t-1}\mathbf{w} \right), \tag{57}$$

where $\mathbf{w}_t^{mv} = (w_{1t}^{mv}, w_{2t}^{mv}, ..., w_{Nt}^{mv})'$, $0 < \kappa_{t-1} < \infty$ is the, possibly time-varying, coefficient of risk aversion. The solution is

$$\mathbf{w}_{t-1}^{mv} = \frac{1}{\kappa_{t-1}} \boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\mu}_{t-1}. \tag{58}$$

The associated portfolio return is

$$\rho_t^{mv} = (\mathbf{r}_t - r_{0,t-1}\mathbf{e})'\mathbf{w}_{t-1}^{mv} + r_{0,t-1}, \tag{59}$$

with conditional mean $\mu_{\rho,t-1}^{mv} = E(\rho_t^{mv} \mid \mathcal{A}_{(N),t-1}) = \kappa_{t-1}^{-1} \left(\boldsymbol{\mu}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\mu}_{t-1} \right)$, and the volatility $\sigma_{\rho,t-1}^{mv} = \sqrt{\text{var}(\rho_t^{mv} \mid \mathcal{A}_{(N),t-1})} = \kappa_{t-1}^{-1} \left(\boldsymbol{\mu}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\mu}_{t-1} \right)^{1/2}$.

Next, we consider a portfolio that is used to establish the APT, here called the arbitrage pricing (`ap`) portfolio. In particular using Ingersoll (1984, Theorem 1) formalization, the `ap` portfolio weights are given by

$$\mathbf{w}_{t-1}^{ap} = \frac{1}{\kappa_{t-1}} \mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1}. \tag{60}$$

with associated portfolio return

$$\rho_t^{ap} = (\mathbf{r}_t - r_{0,t-1}\mathbf{e})'\mathbf{w}_{t-1}^{ap} + r_{0,t-1}, \tag{61}$$

$\mu_{\rho,t-1}^{ap} = E(\rho_t^{ap} \mid \mathcal{A}_{(N),t-1}) = \kappa_{t-1}^{-1} \left(\boldsymbol{\mu}_{t-1}'\mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1} \right)$, and $\sigma_{\rho,t-1}^{ap} = \sqrt{\text{var}(\rho_t^{ap} \mid \mathcal{A}_{(N),t-1})} = \kappa_{t-1}^{-1} \left(\hat{\mathbf{v}}_{t-1}'\mathbf{H}_{t-1}\boldsymbol{\Sigma}_{t-1}\mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1} \right)^{1/2}$. The scaling factor, κ_{t-1}, does not affect our main results and without loss of generality will be set as $\kappa_{t-1} = 1$.[6]

The following theorem, which is valid for any finite $N > k$, clarifies the relationships that exist between `mv` and `ap` portfolios.

Theorem 1 *Suppose the vector of asset returns,* $\mathbf{r}_t$*, follow the factor model, (1), and that Assumptions 1-3 hold. Set*

$$\breve{\boldsymbol{\Sigma}}_t = \left[\mathbf{H}_t - \mathbf{H}_t\mathbf{B}(\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t \right]^{-1}. \tag{62}$$

[5] Recalling that $\boldsymbol{\mu}_{t-1}$ is the conditional mean of the excess return, then the `mv` portfolio return is obtained investing $\mathbf{w}_{t-1}^{mv}$ in the N risky assets and $1 - \mathbf{e}'\mathbf{w}_{t-1}^{mv}$ in the risk free asset.

[6] For instance, Ingersoll (1984) sets $\kappa_t = \hat{\mathbf{v}}_t'\mathbf{H}_t\hat{\mathbf{v}}_t$. However this choice makes the `ap` weights invalid under exact pricing, one of the cases we wish to explore, since an indeterminate form arises.

Then for any finite $N > k$, the `ap` *portfolio weights satisfy*

$$\mathbf{w}_{t-1}^{ap} = \breve{\mathbf{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} = \mathbf{\Sigma}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} = \breve{\mathbf{\Sigma}}_{t-1}^{-1}\mathbf{v}_{t-1} = \breve{\mathbf{\Sigma}}_{t-1}^{-1}\boldsymbol{\mu}_{t-1}. \tag{63}$$

The `ap` *portfolio return satisfies*

$$\rho_t^{ap} = r_{0,t-1} + (\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1})'\breve{\mathbf{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1}, \tag{64}$$

and the difference between the `mv` *and the* `ap` *portfolio returns satisfies*

$$\rho_t^{mv} - \rho_t^{ap} = \boldsymbol{\varepsilon}_t'\mathbf{\Sigma}_{t-1}^{-1}\mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1} + \hat{\boldsymbol{\lambda}}_{t-1}'\mathbf{B}'\mathbf{\Sigma}_{t-1}^{-1}\mathbf{B}(\hat{\boldsymbol{\lambda}}_{t-1} + \mathbf{z}_t), \tag{65}$$

where $\hat{\boldsymbol{\lambda}}_t$ and $\hat{\mathbf{v}}_t$ are defined in (11) and (12), respectively.

Remark 1(a) The third part of (63) is a function of the exact (inverse) covariance matrix, $\mathbf{\Sigma}_{t-1}^{-1}$. The fourth part is a function of the (true) pricing errors $\mathbf{v}_{t-1}$. The fifth part shows clearly the tight analogies that exist between `ap` and `mv` portfolios. Notably, it suggests that for estimation of the `ap` portfolio weights there is no need to identify the pricing errors nor the factor risk premia.
Remark 1(b) Expression (65) shows that two terms make the difference between the `mv` and `ap` portfolio returns. The first term can be shown to be asymptotically negligible for $N \to \infty$, in fact $O_p(N^{-\frac{1}{2}})$, whereas the second term, that involves factors and their risk premia, will be $O_p(1)$, irrespective of the assumed form of no-arbitrage pricing condition.
Remark 1(c) The `ap` portfolio return is functionally independent of the factors and their risk premia for any $N > k$.
Remark 1(d) Under exact no-arbitrage pricing, we have $\mathbf{w}_{t-1}^{ap} = \mathbf{0}$ and $\rho_t^{ap} = r_{0,t-1}$.

3.2 Main Results

We now present the limit behaviour of the `mv` and `ap` portfolio weights and their returns, distinguishing between the various form of no-arbitrage pricing conditions as spelled out by Assumption 5. Throughout this section it will be understood that Assumptions 1, 2, 3 and 4 (or equivalently Assumptions

1, 3, 6, and 7) hold, and that[7]

$$\mathbf{A}_t = \mathbf{D}_t - a_t\left(\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'\right) > 0, \ a.s., \tag{66}$$

$$b_t = 1 + a_t\left(\boldsymbol{\mu}_\beta'\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right) > 0, \ a.s., \tag{67}$$

$$e_t = d_t + (a_t d_t - c_t^2)\left(\boldsymbol{\mu}_\beta'\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right) > 0, \ a.s. \tag{68}$$

Recall that $\boldsymbol{\mu}_\beta$, $\mathbf{D}_t$ and a_t are defined by (5), (6), and (15), respectively, and d_t and c_t, are defined under Assumption 5, depending on the nature of the no-arbitrage pricing condition assumed.

When considering the `mv` and `ap` strategies different notions of diversifiability are needed. Given the factor model (1), we distinguish between diversification of the idiosyncratic part from the market neutrality condition that eliminates the effects of common factors. Regarding diversification of the idiosyncratic part we build upon Chamberlain (1983, Definiton 1), where hereafter $\parallel \mathbf{w} \parallel = (\mathbf{w}'\mathbf{w})^{\frac{1}{2}}$ denotes the Euclidean norm of a given vector of portfolio weights, $\mathbf{w} = (w_1, w_2, ..., w_N)'$.

Definition 1 *(well-diversification) The portfolio* $\mathbf{w}$ *is well diversified if*

$$\parallel \mathbf{w} \parallel \to_p 0 \ \text{as } N \to \infty.$$

Remark (a) Well-diversification of a given portfolio $\mathbf{w}$ implies that idiosyncratic risk, namely the contribution of $\boldsymbol{\varepsilon}_t$, to the portfolio excess return, $\rho_t - r_{0,t-1} = \mathbf{w}'(\mathbf{r}_t - \mathbf{e}r_{0,t-1})$, vanishes in mean square, that is (assume that $\mathbf{w} \in \mathcal{A}_{(N),t-1}$):

$$\text{var}(\mathbf{w}'\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{(N),t-1}) = \mathbf{w}'\mathbf{G}_{t-1}\mathbf{w} \to_p 0 \tag{69}$$

if the maximum eigenvalue of $\mathbf{G}_{t-1}$ does not grow too quickly since

$$\mathbf{w}'\mathbf{G}_{t-1}\mathbf{w} \leq \parallel \mathbf{w} \parallel \rho(\mathbf{G}_{t-1}) \ a.s.$$

For instance, if model (1) has an approximate factor structure in the sense of Chamberlain and Rothschild (1983, Definition 2), then (69) holds but, in fact, $\rho(\mathbf{G}_{t-1})$ is allowed to grow with N and yet (69) would still be satisfied.
Remark (b) We consider a slightly different definition from Chamberlain (1983). Here we wish to distinguish the properties of the portfolio weights from the properties of the limit portfolio return, which instead will depend

[7] Note that

$$\mathbf{A}_t^{-1} = \mathbf{D}_t^{-1} + a_t\left(\frac{\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'\mathbf{D}_t^{-1}}{1 - a_t\boldsymbol{\mu}_\beta'\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta}\right),$$

exists since by Assumptions 2 and 4 $a_t > 0$, $\mathbf{D}_t > 0$, and $a_t\boldsymbol{\mu}_\beta'\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta < 1$ a.s.

on the properties of the portfolio weights together with the characteristics of the data generating process.

Remark (c) According to Green and Hollifield (1992, Definition P1) portfolio $\mathbf{w}$ is well diversified if the sup-norm of $\mathbf{w}$ satisfies $\| \mathbf{w} \|_\infty = \max_{i=1,..,N} | w_i | \rightarrow_p 0$. This is implied by Definition 1 since, as indicated in Green and Hollifield (1992, eq.(45)),

$$\| \mathbf{w} \|_\infty \leq \| \mathbf{w} \| \leq N^{\frac{1}{2}} \| \mathbf{w} \|_\infty .$$

Definition 2 *(asymptotic market neutrality) The portfolio* $\mathbf{w}$ *is said to be asymptotically market (or beta) neutral if*

$$\| \mathbf{B}'\mathbf{w} \| \rightarrow_p 0 \text{ as } N \rightarrow \infty.$$

Remark (a) For the `ap` portfolio, Definition 2 applies for all finite $N > k$ as well, since

$$\mathbf{B}'\mathbf{w}_t^{ap} = \mathbf{0}, \text{ for any } N > k.$$

Remark (b) When a portfolio $\mathbf{w}$ satisfies Definition 2, the contribution to the portfolio return of both the common risk, $\mathbf{z}_t$, and the risk premia, $\boldsymbol{\lambda}_t$, vanish. This is due to the fact that any variable that is loaded by $\mathbf{B}$ is eliminated, and neither the (conditional) mean nor the variance of the portfolio return are affected by any variations in $\mathbf{z}_t$ and $\boldsymbol{\lambda}_t$.

Remark (c) A portfolio that satisfies Definition 2 need not be, and in general will not be, well diversified, as acknowledged for instance by Hubermann (1982, p.187).

Remark (d) Definition 2 will be a relevant property so long as the portfolio weights do not decay to zero too quickly, for otherwise asymptotic market neutrality will be trivially achieved. For instance, if $\max_{i=1,...,n} |w_i| = o_p(N^{-1})$, the weights will satisfy Definition 2 but they would not be of interest as they also eliminate the contribution of all other components of asset excess returns' distribution. Instead, we consider the existence of portfolio weights that are asymptotically market neutral in the sense of Definition 2, whilst at the same time satisfy the condition $0 \leq c < N(\mathbf{w}'\mathbf{w})$ (in probability) for some $c > 0$. This can occur if the individual weights are granular, namely $w_i = O_p(N^{-1})$, *and* the market neutrality condition holds. The granularity of portfolio weights on its own does not imply market neutrality. For example, consider the equally weighted portfolio discussed in Section 4.3 below.

We now derive the limiting properties of the `mv` and `ap` portfolios under various types of no-arbitrage conditions.

Theorem 2 *(exact no-arbitrage pricing)*
Suppose the vector of asset returns, $\mathbf{r}_t$*, follow the factor model, (1), and that Assumptions 1-4 hold. Then under the exact pricing condition (17):*
(i) For any i

$$Nw_{it}^{mv} - \mathbf{e}_i'\mathbf{H}_t\mathbf{B}\mathbf{D}_t^{-1}\boldsymbol{\Omega}_t^{-1}\boldsymbol{\lambda}_t \rightarrow_p 0 \tag{70}$$

and

$$w_{it}^{ap} = 0. \tag{71}$$

(ii) If it is further assumed that

$$N^{-1/2}\mathbf{e}'\mathbf{H}_{t-1}\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{t-1} \rightarrow_d N(0, a_{t-1}), \;\; N^{-1/2}\mathbf{B}'\mathbf{H}_{t-1}\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{t-1} \rightarrow_d N(\mathbf{0}, \mathbf{D}_{t-1}), \tag{72}$$

then

$$\rho_t^{mv} - r_{0,t-1} \;\rightarrow_p\; \boldsymbol{\lambda}_{t-1}'\boldsymbol{\Omega}_{t-1}^{-1}(\boldsymbol{\lambda}_{t-1} + \mathbf{z}_t), \tag{73}$$

$$\left(\frac{\mu_{\rho,t-1}^{mv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{mv}}\right) \;\rightarrow_p\; (\boldsymbol{\lambda}_{t-1}'\boldsymbol{\Omega}_{t-1}^{-1}\boldsymbol{\lambda}_{t-1})^{\frac{1}{2}}, \tag{74}$$

and

$$\rho_t^{ap} = r_{0,t-1}.$$

To summarize, under exact no-arbitrage pricing:
Remark 2(a)w^{mv} is well-diversified (Definition 1), but it is not market neutral (Definition 2). w^{ap} satisfies Definitions 1 for any $N > k$.
Remark 2(b) The limit `mv` portfolio excess return, and its (*ex ante*) Sharpe ratio, are only a function of the factors characteristics. The `ap` portfolio excess return is identically zero and its Sharpe ratio is not defined.
Remark 2(c) The `mv` portfolio weights are $O_p(N^{-1})$, and a function of the factors' characteristics.
Remark 2(d) Primitive conditions for (72) can be readily established.

Theorem 3 *(asymptotic no-arbitrage pricing)*
Suppose the vector of asset returns, $\mathbf{r}_t$*, follow the factor model, (1), and that Assumptions 1-4 hold. Then under (18), (19) and (21) we have:*
(i) For any i

$$N^{\frac{1}{2}}w_{it}^{mv} - \mathring{w}_{it} \rightarrow_p 0 \tag{75}$$

and

$$N^{\frac{1}{2}}w_{it}^{ap} - \mathring{w}_{it} \rightarrow_p 0, \tag{76}$$

where

$$\mathring{w}_{it} = N^{\frac{1}{2}}\mathbf{e}_i'\mathbf{H}_t\mathbf{v}_t - (c_t/b_t)\mathbf{e}_i'\mathbf{H}_t\mathbf{B}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta. \tag{77}$$

(ii) If it is further assumed that (19), (20), (26), and (72) hold, and

$$\mathbf{v}'_{t-1}\mathbf{H}_{t-1}\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{t-1} \rightarrow_d N(0, d_{t-1}), \tag{78}$$

then

$$\rho_t^{mv} - r_{0,t-1} \mid \mathcal{A}_{t-1} \rightarrow_d \frac{e_{t-1}}{b_{t-1}} + x_t + \boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}_{t-1}^{-1}\boldsymbol{\lambda}_{t-1} + \boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}_{t-1}^{-1}\mathbf{z}_t, \tag{79}$$

$$\left(\frac{\mu_{\rho,t-1}^{mv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{mv}}\right) \rightarrow_p \left(\frac{e_{t-1}}{b_{t-1}} + \boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}_{t-1}^{-1}\boldsymbol{\lambda}_{t-1}\right)^{\frac{1}{2}}, \tag{80}$$

and

$$\rho_t^{ap} - r_{0,t-1} \mid \mathcal{A}_{t-1} \rightarrow_d \left(\frac{e_{t-1}}{b_{t-1}}\right) + x_t, \tag{81}$$

$$\left(\frac{\mu_{\rho,t-1}^{ap} - r_{0,t-1}}{\sigma_{\rho,t-1}^{ap}}\right) \rightarrow_p \left(\frac{e_{t-1}}{b_{t-1}}\right)^{\frac{1}{2}}, \tag{82}$$

where $x_t \sim N(0, e_{t-1}/b_{t-1})$.

To summarize, under asymptotic no-arbitrage pricing:
Remark 3(a) w_t^{mv} is neither well-diversified in the sense of Definition 1, nor market neutral in the sense of Definition 2. w_t^{ap} is market neutral but not well-diversified.
Remark 3(b) The limit `mv` portfolio excess return, and the associated *ex ante* Sharpe ratio, are function of both factors and asset-specific characteristics. The *ex ante* Sharpe ratio is positive and bounded. The same features apply to the limit `ap` portfolio excess return with the notable difference of being functionally independent of the common factors. In general, the limit Sharpe ratio for the `ap` portfolio is smaller than that of the `mv` portfolio.
Remark 3(c) The `mv` and `ap` portfolio weights are both $O_p(N^{-\frac{1}{2}})$. Their limit approximation, $\mathring{w}_{it}$, is the same and does not depend on the distribution of the common factors, $\mathbf{z}_t$. It is only a function of asset specific characteristics.
Remark 3(d) Primitive conditions for (78) can be easily established.

Theorem 4 *(unconstrained pricing)*
Suppose the vector of asset returns, $\mathbf{r}_t$*, follow the factor model, (1), and that Assumptions 1-4 hold. Then under (22), (23) and (25), we have:*

(i) For any i

$$w_{it}^{mv} - \breve{w}_{it} \rightarrow_p 0, \tag{83}$$

$$w_{it}^{ap} - \breve{w}_{it} \rightarrow_p 0, \tag{84}$$

where

$$\breve{w}_{it} = \mathbf{e}_i'\mathbf{H}_t\mathbf{v}_t - (c_t/b_t)\mathbf{e}_i'\mathbf{H}_t\mathbf{B}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta. \tag{85}$$

(ii) If it is further assumed that (23), (24) and (26) hold, then

$$N^{-1}(\rho_t^{mv} - r_{0,t-1}) \to_p \left(\frac{e_{t-1}}{b_{t-1}}\right), \tag{86}$$

$$N^{-\frac{1}{2}}\left(\frac{\mu_{\rho,t-1}^{mv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{mv}}\right) \to_p \left(\frac{e_{t-1}}{b_{t-1}}\right)^{\frac{1}{2}}, \tag{87}$$

and

$$N^{-1}(\rho_t^{ap} - r_{0,t-1}) \to_p \left(\frac{e_{t-1}}{b_{t-1}}\right), \tag{88}$$

$$N^{-\frac{1}{2}}\left(\frac{\mu_{\rho,t-1}^{ap} - r_{0,t-1}}{\sigma_{\rho,t-1}^{ap}}\right) \to_p \left(\frac{e_{t-1}}{b_{t-1}}\right)^{\frac{1}{2}}. \tag{89}$$

To summarize, in the unconstrained pricing case:

Remark 4(a) w_t^{mv} is neither well-diversified nor market neutral in the sense of Definitions 1 and 2. w_t^{ap} is market neutral by construction but is not well-diversified in the sense of Definition 1.

Remark 4(b) The `mv` and `ap` portfolio weights are both $O_p(1)$ with the same limit approximation, $\breve{w}_{it}$, which is only a function of asset specific characteristics.

Remark 4(c) The limit `mv` and `ap` portfolio returns, and their associated *ex ante* Sharpe ratios, coincide and are only functions of asset specific characteristics.

Remark 4(d) Normalization by N^{-1} for the return and by $N^{-\frac{1}{2}}$ for the Sharpe ratio is necessary and shows that without the no-arbitrage pricing restrictions, and if $\mathbf{B}$, $\mathbf{H}_{t-1}$ and $\boldsymbol{\mu}_{t-1}$ are known, then the Sharpe ratio of both the `mv` and `ap` portfolios will be unbounded in N, which is clearly not compatible with any form of no-arbitrage[8]. It would be interesting to find out if the Sharpe ratio will continue to be unbounded in N in this *unconstrained* case if estimation uncertainty is taken into account.

[8] MacKinley (1995) pointed out the lack of no-arbitrage implications in terms of Sharpe ratio. See also Cochrane and Saa-Requejo (2001).

4 Discussion of results

4.1 Market neutrality and portfolio diversification

The above analysis makes it clear that a well diversified portfolio need not be market neutral, whilst a market neutral portfolio need not be well diversified. A simple example of the former is the equal weighted portfolio, and an example of the latter is the `ap` portfolio (which is market neutrality for any $N > k$ by construction). Our results also show that although the `mv` strategy is not market neutral *per se* , it can be adjusted to achieve asymptotic market neutrality. The adjustment needed depends on the type of no-arbitrage condition assumed, but does not require knowledge of any of the parameters. In particular, we show that `mv` portfolio can not be made market neutral under exact pricing, but can be adjusted to become asymptotically market neutral under asymptotic no-arbitrage or unconstrained pricing cases, once the `mv` portfolio weights are adjusted by a factor of $N^{-\frac{1}{2}}$ and N^{-1}, respectively.

Whether asymptotic market neutrality holds depends on the limiting behaviour of $\boldsymbol{\Sigma}_t^{-1}\mathbf{B}$, and the type of no-arbitrage restrictions under consideration. In particular, Lemma B in the Appendix shows that, under suitable regularity conditions,

$$\parallel \boldsymbol{\Sigma}_t^{-1}\mathbf{B} \parallel^2 = O_p(N^{-1}).$$

This in turn implies that, for any $N \times p$ (fixed p) matrix $\mathbf{X}_t$ which satisfies $\mathbf{X}_t'\mathbf{H}_t\mathbf{B} = O_p(N)$,

$$\mathbf{X}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{B} = O_p(1). \tag{90}$$

This includes the case where $\mathbf{X}_t$ is set to $\mathbf{B}$ when (6) holds (see (113)). It also implies

$$\mathbf{X}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{B} = O_p(N^{-\frac{1}{2}}) \tag{91}$$

when $\mathbf{X}_t$ is a random zero mean matrix such that $\mathbf{X}_t'\mathbf{H}_t\mathbf{B} = O_p(N^{\frac{1}{2}})$ (see for instance (114) with $\mathbf{X}_t$ set to $\boldsymbol{\varepsilon}_t$).

Bearing in mind (90) and (91), the `mv` portfolio weights can be written as (using (58) and setting $\kappa_t = 1$ for simplicity)

$$\mathbf{w}_t^{mv} = \boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t = \boldsymbol{\Sigma}_t^{-1}\mathbf{v}_t + \boldsymbol{\Sigma}_t^{-1}\mathbf{B}\boldsymbol{\lambda}_t.$$

This, together with (13), yields the following decomposition of the `mv` portfolio excess return, $\rho_t^{mv} - r_{0,t-1} = \mathbf{w}_t^{mv\prime}(\mathbf{r}_t - \mathbf{e}r_{0,t-1})$,

$$\begin{aligned}
\rho_t^{mv} - r_{0,t-1} &= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_t + \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\boldsymbol{\lambda}_{t-1} + (\boldsymbol{\lambda}_{t-1} + \mathbf{z}_t)'\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_t && (92)\\
&+\boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} + \boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\boldsymbol{\lambda}_{t-1} && (93)\\
&+(\boldsymbol{\lambda}_{t-1} + \mathbf{z}_t)'\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\boldsymbol{\lambda}_{t-1}. && (94)
\end{aligned}$$

In the exact pricing case $\mathbf{v}_t = \mathbf{0}$ and the portfolio excess return simplifies to

$$\rho_t^{mv} - r_{0,t-1} = (\boldsymbol{\lambda}_{t-1} + \mathbf{z}_t)'\mathbf{B}'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\boldsymbol{\lambda}_t + \boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\boldsymbol{\lambda}_t.$$

By (90), with $\mathbf{X}_t$ set to $\mathbf{B}$, the first term on the right hand side is $O_p(1)$ and by (91), with $\mathbf{X}_t$ set to $\boldsymbol{\varepsilon}_t$, the second term is $O_p(N^{-\frac{1}{2}})$. Therefore, market neutrality does not apply and, indeed, by Theorem 2, in this case we have $\rho_t^{mv} - r_{0,t-1} \rightarrow_p \quad \boldsymbol{\lambda}_{t-1}'\boldsymbol{\Omega}_{t-1}^{-1}(\boldsymbol{\lambda}_{t-1} + \mathbf{z}_t)$.

Consider now the unconstrained case. Then $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_t^{-1}\mathbf{B} = O_p(1)$ by (90), with $\mathbf{X}_t$ set to $\mathbf{v}_t$, and thus the second and third term on the right hand side of (92) are $O_p(1)$. Of the two terms in the right hand side of (93), which involve $\boldsymbol{\varepsilon}_t$, the first one is $O_p(N^{\frac{1}{2}})$, since we assume that a central limit theorem applies, and the second term is $O_p(N^{-\frac{1}{2}})$ as outlined previously. Likewise, as before, term (94) is $O_p(1)$. The dominating term will then be $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_t$ which is $O_p(N)$ under mild assumptions. It easily follows that if we normalize the mv portfolio weights by N^{-1}, implying that $N^{-1}w_{it}^{mv} = O_p(N^{-1})$, all terms but $N^{-1}\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_t$ will vanish, including all terms involving risk premia $\boldsymbol{\lambda}_{t-1}$ and common factor risk, $\mathbf{z}_t$. Therefore, Definition 2 applies to the adjusted portfolio weights $N^{-1}w_t^{mv}$ in the unconstrained pricing case. The same argument also applies when the asymptotic no-arbitrage pricing holds and one then considers the `mv` portfolio weights normalized by $N^{-\frac{1}{2}}$ to which Definition 2 applies. Note that, once again, $N^{-\frac{1}{2}}w_{it}^{mv} = O_p(N^{-1})$.

The `ap` portfolio is market neutral by construction, for any finite $N > k$ and for any form of no-arbitrage assumptions since $\mathbf{w}_t^{ap} = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\boldsymbol{\mu}_{t-1}$ by Theorem 1, and recalling that $\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{B} = \mathbf{0}$. One would therefore expect a close relationship between the `ap` and `mv` strategies in the non-exact pricing cases, since as we have seen the `mv` strategy can be made market neutral (as $N \rightarrow \infty$) in such cases. This is in line with our earlier result showing that in the unconstrained case the `ap` and `mv` strategies are in fact asymptotically equivalent in terms of portfolio weights as well as the limit portfolio returns (here once both normalized by the same power of N). This equivalence is due to $\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}$ and $\boldsymbol{\Sigma}_{t-1}^{-1}$ being asymptotically equivalent since $\| \breve{\boldsymbol{\Sigma}}_{t-1}^{-1} - \boldsymbol{\Sigma}_{t-1}^{-1} \|^2 = O_p(N^{-1})$. Such equivalence plays no role in exact pricing since in that case the `ap` weights are identically null, unlike the `mv` weights which depend on the factor characteristics. As a result it is not surprising that the limit behaviour of the two strategies differ markedly in the exact pricing case.

Consider now the issue of portfolio diversification. The degree of portfolio diversification achieved crucially depends on the assumed nature of the arbitrage condition. Under the exact pricing case, the `mv` portfolio weights can be written in vector form as

$$\mathbf{w}_t^{mv} = N^{-1}\mathbf{H}_t\mathbf{B}\mathbf{D}_t^{-1}\boldsymbol{\Omega}_t^{-1}\boldsymbol{\lambda}_t + O_p(N^{-3/2}),$$

and

$$\mathbf{w}_t^{mv\prime}\mathbf{w}_t^{mv} = N^{-1}\boldsymbol{\lambda}_t'\boldsymbol{\Omega}_t^{-1}\mathbf{D}_t^{-1}\left(\frac{\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B}}{N}\right)\mathbf{D}_t^{-1}\boldsymbol{\Omega}_t^{-1}\boldsymbol{\lambda}_t + O_p(N^{-5/2})$$

yielding $\mathbf{w}_t^{mv\prime}\mathbf{w}_t^{mv} \to_p 0$ as $N \to \infty$, and the `mv` portfolio in this case is well-diversified in the sense of Definition 1. It is also easily seen that $|w_{it}^{mv}| \,/\, \|\mathbf{w}_t^{mv}\| \to_p 0$ as $N \to \infty$ and the portfolio weights are granular in the sense of Chudik, Pesaran, and Tosetti (2009).

The above result does not carry over to the other two cases. In the asymptotic no-arbitrage case

$$\mathring{\mathbf{w}}_t = N^{\frac{1}{2}}\mathbf{H}_t\mathbf{v}_t - (c_t/b_t)\mathbf{H}_t\mathbf{B}\mathbf{A}_\mathbf{t}^{-1}\boldsymbol{\mu}_\beta,$$

which yields

$$\begin{aligned}\mathbf{w}_t^{mv\prime}\mathbf{w}_t^{mv} \;&=\; \mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{v}_t - 2(c_t/b_t)\left(\frac{\mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{B}}{N}\right)\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta \\ &\quad +(c_t/b_t)^2\boldsymbol{\mu}_\beta'\mathbf{A}_t^{-1}\left(\frac{\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B}}{N}\right)\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta + O_p(N^{-1/2}),\end{aligned}$$

which is $O_p(1)$. In this case the weights are not granular and the possibility that one or more assets in the portfolio will be given sizeable weights can not be ruled out even if $N \to \infty$. The situation is even more extreme if the unconstrained pricing case is considered. In this case

$$\begin{aligned}N^{-1}\breve{\mathbf{w}}_t'\breve{\mathbf{w}}_t \;&=\; \left(\frac{\mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{v}_t}{N}\right) + (c_t/b_t)^2\boldsymbol{\mu}_\beta'\mathbf{A}_t^{-1}\left(\frac{\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B}}{N}\right)\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta \\ &\quad -2(c_t/b_t)\left(\frac{\mathbf{v}_t'\mathbf{H}_t\mathbf{H}_t\mathbf{B}}{N}\right)\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta,\end{aligned}$$

and $\mathbf{w}_t^{mv\prime}\mathbf{w}_t^{mv} = O_p(N)$.

4.2 Contribution of factors to portfolio return

The `mv` and `ap` strategies are closely related. Their associated portfolio returns coincide, as N gets arbitrarily large, in the unconstrained case and are positively correlated, albeit not perfectly so, in the asymptotic no-arbitrage pricing case. To see this consider (65). It is easily seen that the difference between ρ_t^{mv} and ρ_t^{ap} is always at most $O_p(1)$ irrespective of the type of no-arbitrage assumption postulated. In fact, of the two terms that makes *exactly* the difference between the portfolio returns, $\boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1}$ is $O_p(N^{-\frac{1}{2}})$

whereas $\hat{\boldsymbol{\lambda}}'_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}(\hat{\boldsymbol{\lambda}}_{t-1}+\mathbf{z}_t)$ is $O_p(1)$, as illustrated in the previous section. Both terms have limits that are independent of $\mathbf{v}_t$, and thus do not depend on whether some form of no-arbitrage holds or not. The latter matters, however, for the limit behaviour of the portfolio return themselves (rather than for their difference) as evident from (64) which we use to establish the limit behaviour of the `ap` portfolio return. Now, in the exact pricing case $\mathbf{v}_t = \mathbf{0}$ yielding $\rho^{ap}_t = r_{0,t-1}$ and thus one finds zero (conditional) correlation between the two portfolio returns. In the asymptotic no-arbitrage pricing case, $\rho^{ap}_t - r_{0,t-1} = O_p(1)$ since, by (63), $\boldsymbol{\varepsilon}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\hat{\mathbf{v}}_{t-1} = \boldsymbol{\varepsilon}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1}$ and $\hat{\mathbf{v}}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\hat{\mathbf{v}}_{t-1} = \mathbf{v}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1}$ where both terms are $O_p(1)$. In this case, the `mv` and `ap` portfolio returns are positively, but not perfectly, correlated. Lack of perfect correlation is due to the fact the `ap` portfolio return is, by construction, independent of the common factors whereas the latter, as indicated by (65), contains $O_p(1)$ terms involving the common factors. Finally, in the unconstrained case, the `mv` and `ap` portfolio returns are perfectly correlated since they both involve the term $\hat{\mathbf{v}}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\hat{\mathbf{v}}_{t-1} = \mathbf{v}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1}$ which is $O_p(N)$ and dominates all the other terms which are at most $O_p(N^{\frac{1}{2}})$, as $\boldsymbol{\varepsilon}_t\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\hat{\mathbf{v}}_{t-1}$, or even $O_p(1)$, as the terms in (65). As a consequence, the unconstrained case requires the normalization by N^{-1} to establish well-defined limits for the `mv` and `ap` portfolio returns.

Extension to the minimum variance (minimizing the variance subject to a target expected return) and maximum expected return (maximizing the expected return subject to a target portfolio variance) portfolios is straightforward. However, our focus on `mv` trading strategies is less restrictive than it might appear at first since our results equally apply to other trading strategies so long as they can be written as $\boldsymbol{\Sigma}^{-1}_{t-1}\boldsymbol{\delta}_{t-1}$ for some $N \times 1$ vector $\boldsymbol{\delta}_{t-1} = \boldsymbol{\delta}(\mathcal{A}_{(N),t-1})$, a general function of $\mathcal{A}_{(N),t-1}$, the exact form of which could depend on the type of trading strategy under consideration. For instance, our results hold for certain dynamic trading strategies where the portfolio weights can be written as the sum of the `mv` component and of an inter-temporal hedging component, both of which employ the inverse of the covariance matrix in the suitable way. See, for example, Campbell and Viceira (2001) and Campbell, Chan, and Viceira (2003).

The different forms of no-arbitrage imply different behaviour of the *ex ante* Sharpe ratio. In the exact and asymptotic no-arbitrage pricing cases, the Sharpe ratio turns out to be bounded even for large N. It is well known that this is a fundamental property which is key to establishing the APT (see, for instance, Ingersoll (1984, Theroem 1)). Instead, the Sharpe ratio of `mv` and `ap` portfolios diverge to plus infinity at rate $N^{\frac{1}{2}}$ in the unconstrained case. This result generalizes MacKinley (1995) who provides an example where such divergence occurs, in particular when there is a non-priced missing factor from the postulated factor structure.[9]

4.3 Comparisons with some sub-optimal portfolios

It is of interest also to compare the limit properties of optimal `mv` portfolios to sub-optimal but relatively easy to implement portfolios such as the global-minimum-variance (`gmv`) portfolio and the equal weighted (`ew`) portfolio. The `gmv` portfolio weights, $\mathbf{w}_t^{gmv} = (w_{1t}^{gmv}, w_{2t}^{gmv}, ..., w_{Nt}^{gmv})'$, are the solution to the problem:

$$\mathbf{w}_t^{gmv} = \text{argmin}_{\mathbf{w}} \left(\mathbf{w}'\boldsymbol{\Sigma}_t\mathbf{w}\right), \text{ such that } \mathbf{w}'\mathbf{e} = 1, \tag{95}$$

yielding

$$\mathbf{w}_t^{gmv} = \frac{\boldsymbol{\Sigma}_t^{-1}\mathbf{e}}{\mathbf{e}'\boldsymbol{\Sigma}_t^{-1}\mathbf{e}}. \tag{96}$$

It is well known that this portfolio does not belong to the efficient frontier, except when the conditional expected returns $\mu_{i,t-1}$ are the same across i. Nevertheless, this portfolio is still of interest since its implementation does not require the estimation of expected returns. Jagannathan and Ma (2003) show that, in terms of asset allocation, its out-of-sample performance is comparable with the performance of portfolios that do belong to the tangency portfolios.

Let

$$\mathring{w}_{it}^{gmv} = N^{-1}\left[\left(\frac{b_t}{a_t}\right)\mathbf{e}_i'\mathbf{H}_t\mathbf{e} - \mathbf{e}_i'\mathbf{H}_t\mathbf{B}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right]. \tag{97}$$

Then, under any of the no-arbitrage conditions listed in Assumption 5 we have:

$$N(w_{it}^{gmv} - \mathring{w}_{it}^{gmv}) \rightarrow_p 0. \tag{98}$$

For portfolio returns and the *ex ante* Sharpe ratio we have under exact no-arbitrage, for a random variable $g_t \sim N(0, a_{t-1}/b_{t-1})$,

$$N^{\frac{1}{2}}(\rho_t^{gmv} - r_{0,t-1}) \mid \mathcal{A}_{t-1} \rightarrow_d g_t,$$

$$N^{\frac{1}{2}}\left(\frac{\mu_{\rho,t-1}^{gmv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{gmv}}\right) \rightarrow_p \sqrt{a}_{t-1}\boldsymbol{\mu}'_\beta\mathbf{A}_{t-1}^{-1}\boldsymbol{\Omega}_{t-1}^{-1}\boldsymbol{\lambda}_{t-1}.$$

Under asymptotic no-arbitrage

$$N^{\frac{1}{2}}(\rho_t^{gmv} - r_{0,t-1}) \mid \mathcal{A}_{t-1} \rightarrow_d g_t + \frac{c_{t-1}}{a_{t-1}},$$

[9] In case of a priced missing factor there is a correspondence between the pricing error and the innovation variance that ensures boundedness of the Sharpe ratio. See MacKinley (1995, eq. (17)).

$$\left(\frac{\mu_{\rho,t-1}^{gmv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{gmv}}\right) \to_p \frac{c_{t-1}}{\sqrt{a_{t-1}b_{t-1}}}.$$

Finally, under the unrestricted case

$$(\rho_t^{gmv} - r_{0,t-1}) \to_p \frac{c_{t-1}}{a_{t-1}},$$

$$N^{-\frac{1}{2}}\left(\frac{\mu_{\rho,t-1}^{gmv} - r_{0,t-1}}{\sigma_{\rho,t-1}^{gmv}}\right) \to_p \frac{c_{t-1}}{\sqrt{a_{t-1}b_{t-1}}}.$$

Therefore, under all the three cases, the `gmv` portfolio is well-diversified and asymptotically market neutral in the sense of Definitions 1 and 2. In the exact and asymptotic no-arbitrage pricing cases, the portfolio return converges to the risk free rate, but not so in the unconstrained case where a weighted average of the pricing errors $\mathbf{v}_{t-1}$ makes the limit excess return. Moreover, in this last case its *ex ante* Sharpe ratio diverges at rate $O_p(N^{\frac{1}{2}})$ (as in the case of `mv` and `ap` portfolios), but importantly it is not guaranteed to diverge to infinity (c_t need not be positive). This largely reflects the sub-optimal nature of the `gmv` strategy as it does not make use of the expected return predictions.

Consider now the `ew` portfolio defined by $w_t^{ew} = N^{-1}\mathbf{e}$. Since $\text{var}(N^{-1}\mathbf{e}'\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{(N),t-1}) = N^{-2}\mathbf{e}'\mathbf{G}_{t-1}\mathbf{e} \leq N^{-1}\rho\left(\mathbf{G}_{t-1}\right)$ then for $\rho\left(\mathbf{G}_{t-1}\right) = o_p(N)$ it readily follows that

$$\rho_t^{ew} - r_{0,t-1} = N^{-1}(\mathbf{r}_t - r_{0,t-1}\mathbf{e})'\mathbf{e} \to_p \boldsymbol{\mu}_{v,t-1} + \boldsymbol{\mu}_\beta'\boldsymbol{\lambda}_{t-1},$$

where $N^{-1}\mathbf{e}'\mathbf{v}_t \to_p \boldsymbol{\mu}_{vt}$, and

$$\left(\frac{\mu_{\rho,t-1}^{ew} - r_{0,t-1}}{\sigma_{\rho,t-1}^{ew}}\right) \to_p \frac{\boldsymbol{\mu}_{v,t-1} + \boldsymbol{\mu}_\beta'\ \boldsymbol{\lambda}_{t-1}}{\boldsymbol{\mu}_\beta'\boldsymbol{\Omega}_{t-1}\boldsymbol{\mu}_\beta}.$$

Hence, the equal weighted portfolio is well-diversified, but is not market neutral. Well-diversification occurs when $\rho\left(\mathbf{G}_{t-1}\right) = o_p(N)$. The *ex ante* Sharpe ratio of the equal weighted portfolio is bounded in N, but need not be positive. This is in contrast to the `mv` and `ap` portfolios that ensure a non-negative limit for the *ex ante* Sharpe ratio. The Sharpe ratio of the equal weighted portfolio is always a function of the factors conditional moments. The only consequence of either the exact or approximate no-arbitrage pricing case is that $\boldsymbol{\mu}_{vt} = \mathbf{0}$, without changing much in terms of the behaviour of the Sharpe ratio of the `ew` portfolio. Therefore, the relatively favourable evidence provided in the empirical literature for the `ew` portfolio (see DeMiguel, Garlappi, and Uppal (2009)) most likely is due to the negative impact of the estimation uncertainty on the performance of the `mv` or `ap` portfolios.

4.4 Contribution of factors to portfolio weights

The conditional distribution of the common factors is irrelevant, as far as the form of the limiting mv and ap portfolio weights is concerned, both in the asymptotic no-arbitrage pricing and unconstrained cases. For the (sub-optimal) global minimum variance portfolio weights (97) this holds for the exact pricing case as well.

An immediate implication is that when these portfolio weights are set empirically, one can avoid specifying, let alone estimating, the conditional mean and the conditional covariance matrix of the common factors. Obviously, one is not exempted from estimation of the factors themselves (when latent) in so far as they help in estimation of $\mathbf{v}_{t-1}$ and $\mathbf{H}_{t-1}$. For a finite N, this estimation strategy clearly would involve an approximation error since the finite-N expression of the portfolio weights will necessarily be a function of $\boldsymbol{\Omega}_{t-1}$ and $\boldsymbol{\lambda}_{t-1}$. However, such approximation error decreases to zero as N increases and, at the same time, using either the limit portfolio formulae (77), (85), and (97), is likely to be robust to the consequences of incorrectly specifying, or poorly estimating $\boldsymbol{\Omega}_{t-1}$ and $\boldsymbol{\lambda}_{t-1}$.

Part (i) of Theorems 2, 3, and 4 can be interpreted as a cross sectional consistency result, showing the form of the limit approximations, as $N \to \infty$, of the portfolio weights. These results can be strengthened, by showing that the asymptotic distribution of the portfolio weights, centered around the limit portfolio weights, is distributed independently of the conditional moments of $\mathbf{z}_t$. In other words, the contribution of these moments to the (finite-N) portfolio weights vanishes at a suitably fast rate, faster than the rate required to obtain the asymptotic distribution of the portfolio weights. Moreover, the stated results of the above theorems hold not only point-wise for each $i = 1, 2, ..., N$, but also hold jointly for the entire vector of portfolio weights. For instance, considering the mv strategy, it can be shown that $\left\|\mathbf{w}_t^{mv} - N^{-1}\mathbf{H}_t\mathbf{B}\mathbf{D}_t^{-1}\boldsymbol{\Omega}_t^{-1}\boldsymbol{\lambda}_t\right\| = o_p(N^{-1})$ in the exact pricing case, $\left\|\mathbf{w}_t^{mv} - N^{-\frac{1}{2}}\mathring{\mathbf{w}}_t^{mv}\right\| = o_p(N^{-\frac{1}{2}})$ in the asymptotic no-arbitrage pricing case and $\|\mathbf{w}_t^{mv} - \breve{\mathbf{w}}_t^{mv}\| = o_p(1)$ in the unconstrained pricing case. The last two bounds also apply when substituting $\mathbf{w}_t^{mv}$ with $\mathbf{w}_t^{ap}$.

Another important consequence of part (i) of these theorems is that, with the exception of the mv portfolio weights in the exact pricing case, the limiting portfolio weights will not be time-varying unless $\mathbf{H}_t$ is, that is only if the idiosyncratic component $\boldsymbol{\varepsilon}_t$ features dynamic conditional heteroskedasticity. If we relax our assumptions, say allowing $\mathbf{B}$ to be time-varying $\mathbf{B}_t$, then portfolio weights become time-varying even if $\mathbf{H}_t = \mathbf{H}$. But for this case to be genuinely interesting, $\mathbf{B}_t$ needs to be independent from the factors $\mathbf{z}_t$.

This rules out the case $\mathbf{B}_t = \mathbf{B}\mathbf{\Omega}_t^{\frac{1}{2}}$, which, as far as the dynamics of $\mathbf{r}_t$ is concerned, is observationally equivalent to (1). If instead one alternatively assumes the parameter-free form $\mathbf{\Omega}_t = \mathbf{I}_k$, our result continues to apply since the limit portfolios continue to be functionally independent of any parametric aspect of $\mathbf{\Omega}_t$.

4.5 Short-selling and factor dominance

Green and Hollifield (1992) argue that the possibility of short-selling, in the sense of a repeated finding of negative optimal portfolio weights, is related to the presence of one dominant factor, challenging the common view that large negative weights are simply the consequences of estimation uncertainty. Jagannathan and Ma (2003) provide some simulation results which suggests that this effect depends on the cross-sectional dispersion of the factor loadings. Our result provide some analytical insights on this issue.

To simplify the exposition we shall consider the case where $\mathbf{H}_t$ is diagonal, the factor loadings, $\boldsymbol{\beta}_i$, are *i.i.d.* with finite covariance matrix $\mathbf{\Sigma}_\beta$, and focus on the `mv` portfolio under the unconstrained pricing case. Using (85), and noting that $\mathbf{e}_i'\mathbf{H}_t\mathbf{v}_t = v_{it}h_{ii,t}$ and $\mathbf{B}'\mathbf{H}_t\mathbf{e}_i = h_{ii,t}\boldsymbol{\beta}_i$, $\mathbf{A}_t = a_t\mathbf{\Sigma}_\beta$, and after some algebra we have

$$w_{it}^{mv} \rightarrow_p h_{ii,t}\left[v_{it} - \frac{\mu_{vt}^h\left(\boldsymbol{\mu}_\beta'\mathbf{\Sigma}_\beta^{-1}\boldsymbol{\beta}_i\right)}{1+\boldsymbol{\mu}_\beta'\mathbf{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta}\right],$$

where for simplicity we set $\mu_{vt}^h = c_t/a_t$, namely the probability limit of $(\Sigma_{j=1}^N v_{jt}h_{jj,t}/\Sigma_{j=1}^N h_{jj,t})$, and as before $\mu_{\beta j}$ and β_{ij} are the j^{th} element of $\boldsymbol{\mu}_\beta = (\mu_{\beta 1}, \mu_{\beta 2}, ..., \mu_{\beta k})'$ and $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, ..., \beta_{ik})'$, respectively. Since the factor loadings are assumed to be independently distributed, when $\mu_{\beta j} \neq 0$ for $j = 1, 2, ..., k$, the above result can be written equivalently[10] as

$$w_{it}^{mv} \rightarrow_p \frac{h_{ii,t}}{1+\boldsymbol{\mu}_\beta'\mathbf{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta}\left[v_{it} + \left(\frac{\mu_{\beta 1}}{\sigma_{\beta 1}}\right)^2\left(v_{it} - \mu_{vt}^h\frac{\beta_{i1}}{\mu_{\beta 1}}\right) + ... + \left(\frac{\mu_{\beta k}}{\sigma_{\beta k}}\right)^2\left(v_{it} - \mu_{vt}^h\frac{\beta_{ik}}{\mu_{\beta k}}\right)\right]. \tag{99}$$

Note that μ_{vt}^h can be viewed as a weighted average of the pricing errors, v_{jt}, with the weights given by the relative inverse volatilities of the idiosyncratic shocks, $h_{jj,t}/\Sigma_{i=1}^N h_{ii,t}$, for $j = 1, 2, ..., N$. The terms $v_{it} - \mu_{vt}^h\left(\beta_{ij}/\mu_{\beta j}\right)$ then measures the extent to which the pricing error of the ith asset deviates from the mean pricing errors across all assets, corrected by $\beta_{ij}/\mu_{\beta j}$.

[10] The requirement that $\mu_{\beta j} \neq 0$ is not restrictive since the term associated with the factor with a zero mean factor loading can be dropped from the expression.

Green and Hollifield (1992) argue that the possibility of short-selling, in the sense of a repeated finding of negative optimal portfolio weights, is related to the presence of one dominant factor. Our result sheds some light on this. One can see from (99) that the limit portfolio weights only depend on factor loadings if the mean of these loadings is non-zero (i.e. if $\mu_{\beta j} \neq 0$). Such factors are regarded as *dominant* by Jagannathan and Ma (2003).[11]

Therefore, a negative weight is more likely for the asset for which $v_{it} < 0$. Moreover, when $v_{it} - \mu_{vt}^{h}\left(\beta_{ij}/\mu_{\beta j}\right) > 0$ for some $1 \leq j \leq k$, a negative weight is more likely to arise whenever the jth factor loading assumes a value larger than their cross-sectional average. This effect is magnified, the larger is the *Sharpe ratio* of the factor loading, defined by $\mu_{\beta j}/\sigma_{\beta j}$. A large beta dispersion implies a smaller chance of finding negative weights, corroborating the findings based on simulations reported by Jagannathan and Ma (2003). On the other hand, note also that the larger the number of dominant factors under consideration (in the sense of Jagannathan and Ma (2003)), the less likely it is that a negative weight would be encountered. Similar outcomes obtain for non-diagonal $\mathbf{H}_t$. This is in line with Green and Hollifield (1992)'s conjecture about the presence of a single dominant factor whenever large negative weights are observed.

5 Final remarks

In this paper we have investigated the limit properties of mean-variance (`mv`) and arbitrage pricing (`ap`) trading strategies, as the number of assets diverge to infinity. Specifically, we have focussed on the issue of portfolio diversification and the extent to which different portfolio strategies are able to diversify the idiosyncratic versus the market risks. We have extended the results obtained in the literature for the exact pricing case to two other cases of asymptotic no-arbitrage pricing and the unconstrained pricing scenarios. Under the non-exact pricing cases it is established that the two portfolios (`mv` and `ap`) are closely related and are in fact asymptotically equivalent in the unconstrained case. The results are related to a number of issues of interest in the literature on asset pricing such as the prevalence of short-selling, the number of dominant factors and the granularity property of the portfolio weights. We also consider diversifiability and optimality of the global minimum variance and the equally weighted portfolios, two sub-optimal yet popular strategies. For the first time, the paper also provides a set of primi-

[11] Note that a dominant factor is necessarily strong in the sense discussed earlier. But a strong factor need not be dominant since it is possible for the factor loadings to be divergent, in the sense of $N^{-1}\Sigma_{i=1}^{N}\left|\beta_{ij}\right| \rightarrow_p K > 0$ with $N^{-1}\Sigma_{i=1}^{N}\beta_{ij} \rightarrow_p \mu_{\beta j} = 0$.

tive conditions on the asset return distributions that ensure the more familiar higher level assumptions needed for the validity of the limit results provided in the asset pricing literature as well as in this paper.

Our results have important practical implications. It is well know that under the assumption of correct model specification, factor model-based optimal portfolios weights leads to more efficient estimates of the corresponding portfolio variance, as compared to the familiar sample moment plug-in estimates (see the empirical results of Chan, Karceski, and Lakonishok (1999) and the theoretical results of Fan, Fan, and Lv (2007)). However, the asymptotic independence of `mv` portfolio weights from the common factors' conditional distribution, established in this paper, suggests that in the case of large portfolios it might be prudent to side-step the tasks of specification and estimation of the conditional distribution of the factors and instead use the approximate formulae for the portfolio weights advanced in this paper. In this way it might be possible to avoid the adverse effects of model and parameter uncertainties that surround the specification of the unobserved common factor models. But before this issue can be examined one also needs to consider the extent to which the properties of the limit portfolios are still valid when the remaining unknown parameters are replaced by their estimates. This is particularly important with respect to the market neutrality properties of the `mv` and `ap` strategies. This papers clarifies that an essential ingredient for building market neutral trading strategies is to recover the factor loadings corresponding to the *strong factors*. Instead, knowledge of the factors conditional distribution is not relevant here. Therefore, in so far as market neutrality is concerned, the development of efficient testing and estimation procedures to carry inference on the loading of the (dominant) factors appear of primary importance. Double asymptotic results will need to be established, where both the cross-section and the time series dimensions diverge to infinity, unlike this paper whose results hold at each point in time. The examination of estimation uncertainty and its effects on the properties of the limit portfolios is beyond the scope of the present paper and will be the subject of future research.

Appendix

Part A: Lemmas

The following lemmas derive the means and variances of linear, bilinear and quadratic forms in random vectors and matrices.[12]

Lemma A.1 Let

$$Q_{1,N} = \boldsymbol{\alpha}'\mathbf{A}\boldsymbol{\alpha},$$

where the $N \times N$ random matrix $\mathbf{A}$ is symmetric, $\mathbf{A} > 0$ *a.s.* such that the $N(N+1)/2 \times 1$ random vector $\mathbf{a} = \text{vech}(\mathbf{A})$ has mean $\boldsymbol{\mu}_a$ and covariance matrix $\boldsymbol{\Omega}_a = \mathbf{P}_a\mathbf{P}'_a$. The random $N \times 1$ vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_N)'$ has mean $E(\boldsymbol{\alpha}) = \boldsymbol{\mu}_\alpha$, covariance matrix $\boldsymbol{\Omega}_\alpha = \mathbf{P}_\alpha\mathbf{P}'_\alpha > 0$ and finite fourth moment satisfying $E(v_i^3) = \gamma_1$ and $E(v_i^4) = \gamma_2 + 3$ for $\mathbf{v} = (v_1, v_2, ..., v_N)' = \mathbf{P}_\alpha^{-1}(\boldsymbol{\alpha} - \boldsymbol{\mu}_\alpha)$. When:

$\boldsymbol{\alpha}$ and $\mathbf{A}$ are, element by element, mutually independent

then

$$\begin{aligned}\text{var}(Q_{1,N}) &= 2\text{tr}[E(\mathbf{C}^2)] + 4\boldsymbol{\mu}'_\alpha E(\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A})\boldsymbol{\mu}_\alpha + \gamma_2\text{tr}[E(\mathbf{C}\odot\mathbf{C})] \\ &\quad +4\gamma_1\boldsymbol{\mu}'_\alpha E[\mathbf{A}\mathbf{P}_\alpha(\mathbf{I}_N\odot\mathbf{C})]\mathbf{e} + \text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha)'\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}'_N\text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha)\end{aligned}$$

where $\mathbf{C} = \mathbf{P}'_\alpha\mathbf{A}\mathbf{P}_\alpha$, $\text{tr}(\mathbf{A})$, $\rho(\mathbf{A})$ and $\text{vech}(\mathbf{A})$ denote the trace, the spectral norm and the vech (the vector with all the distinct elements) of a matrix $\mathbf{A}$, $\odot$ and $\otimes$ are the Hadamard and Kronecker product operators, respectively, and $\mathbf{D}_N$ is the $N^2 \times (N(N+1)/2)$ duplication matrix (see Magnus and Neudecker (2001, p.48)).

An alternative representation is

$$\begin{aligned}\text{var}(Q_{1,N}) &= 2\text{tr}\left[\mathbf{D}'_N\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\Omega}_\alpha\right)\mathbf{D}_N\boldsymbol{\Omega}_a\right] + 4\text{tr}\left[\mathbf{D}'_N\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha\right)\mathbf{D}_N\boldsymbol{\Omega}_a\right] \\ &\quad +2\boldsymbol{\mu}'_a\mathbf{D}'_N\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\Omega}_\alpha\right)\mathbf{D}_N\boldsymbol{\mu}_a + 4\boldsymbol{\mu}'_a\mathbf{D}'_N\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha\right)\mathbf{D}_N\boldsymbol{\mu}_a \\ &\quad +\gamma_2\text{tr}\left[E\left(\mathbf{C}\odot\mathbf{C}\right)\right] + 4\gamma_1\boldsymbol{\mu}'_\alpha\mathbf{P}'^{-1}_\alpha E\left[\mathbf{C}(\mathbf{I}_N\odot\mathbf{C})\right]\mathbf{e} \\ &\quad +\text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha)'\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}'_N\text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha).\end{aligned}$$

Proof: since

$$Q_{1,N} = \mathbf{v}'\mathbf{P}'_\alpha\mathbf{A}\mathbf{P}_\alpha\mathbf{v} + 2\mathbf{v}'\mathbf{P}'_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha + \boldsymbol{\mu}'_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha$$

[12] Moments of quadratic forms when the weight matrix $\mathbf{A}$ is non-stochastic are available in the statistics literature. See, for example, Ullah (2004). But to our knowledge such moments have not been derived for the case where $\mathbf{A}$ is a matrix with stochastic elements.

it follows that

$$\text{var}(Q_{1,N}) = E\left[\text{var}(Q_{1,N}\,|\mathbf{A})\right] + \text{var}(E(Q_{1,N}\,|\mathbf{A})). \tag{100}$$

Since $\boldsymbol{\alpha}$ and $\mathbf{A}$ are independently distributed $E(Q_{1,N}|\mathbf{A}) = \text{tr}(\mathbf{C}) + \boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\mu}_\alpha$ and

$$\begin{aligned}
&E(Q_{1,N}^2\,|\mathbf{A}) = \\
&E\left[(\mathbf{v}'\mathbf{C}\mathbf{v})^2\,|\mathbf{A}\right] + 4\boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha + (\boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\mu}_\alpha)^2 \\
&+4\boldsymbol{\mu}_\alpha'\mathbf{A}'\mathbf{P}_\alpha E\left(\mathbf{v}\left(\mathbf{v}'\mathbf{C}\mathbf{v}\right)\right) + 2\left(\boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\mu}_\alpha\right)\text{tr}(\mathbf{C}).
\end{aligned}$$

Using results in Ullah (2004) we have,

$$\begin{aligned}
E\left[(\mathbf{v}'\mathbf{C}\mathbf{v})^2\,|\mathbf{A}\right] &= [\text{tr}(\mathbf{C})]^2 + 2\text{tr}(\mathbf{C}^2) + \gamma_2\text{tr}(\mathbf{C}\odot\mathbf{C}) \\
E\left[\mathbf{v}\left(\mathbf{v}'\mathbf{C}\mathbf{v}\right)|\mathbf{A}\right] &= \gamma_1(\mathbf{I}_n\odot\mathbf{C})\mathbf{e}.
\end{aligned}$$

Hence, after some algebra we obtain

$$\text{var}(Q_{1,N}\,|\mathbf{A}) = 2\text{tr}(\mathbf{C}^2) + 4\boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha + \gamma_2\text{tr}(\mathbf{C}\odot\mathbf{C}) + 4\gamma_1\boldsymbol{\mu}_\alpha'\mathbf{A}\mathbf{P}_\alpha(\mathbf{I}_N\odot\mathbf{C})\mathbf{e}.$$

Then, regarding $\text{var}[E(Q_{1,N}\,|\mathbf{A})]$, let $\mathbf{F} = \boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha'$ and note that $E(Q_{1,N}\,|\mathbf{A}) = \text{vec}(\mathbf{F})'\text{vec}(\mathbf{A})$. Since $\mathbf{A} = \mathbf{A}'$, using the duplication matrix $\mathbf{D}_N$, we have $\text{vec}(\mathbf{A}) = \mathbf{D}_N\text{vech}(\mathbf{A}) = \mathbf{D}_N\mathbf{a}$ where $\mathbf{a}$ is a $N(N+1)/2\times 1$ vector, composed of the distinct elements of $\mathbf{A}$. Similarly $\text{vec}(\mathbf{F}) = \mathbf{D}_N\text{vech}(\mathbf{F}) = \mathbf{D}_N\mathbf{f}$. Setting

$$\mathbf{d} = \mathbf{D}_N'\mathbf{D}_N\mathbf{f}$$

one gets

$$\text{var}[E(Q_{1,N}\,|\mathbf{A})] = \text{var}(\mathbf{d}'\mathbf{a}) = \text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha')'\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}_N'\text{vec}(\boldsymbol{\Omega}_\alpha + \boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha').$$

Adding the above expression to the expectation of $E(\text{var}(Q_{1,N}\,|\mathbf{A}))$ yields the first representation of $\text{var}(Q_{1,N})$.

Now we establish the second representation. Since

$$\begin{aligned}
\text{tr}(\mathbf{C}^2) &= \text{tr}\left(\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A}\right) = \text{vec}(\mathbf{A})'\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\Omega}_\alpha\right)\text{vec}(\mathbf{A}) \\
\boldsymbol{\mu}_\alpha'\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha &= \text{tr}(\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha'\mathbf{A}) = \text{vec}(\mathbf{A})'\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha'\right)\text{vec}(\mathbf{A})
\end{aligned}$$

then

$$\begin{aligned}
\text{var}(Q_{1,N}\,|\mathbf{A}) &= 2\text{vec}(\mathbf{A})'\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\Omega}_\alpha\right)\text{vec}(\mathbf{A}) + 4\text{vec}(\mathbf{A})'\left(\boldsymbol{\Omega}_\alpha\otimes\boldsymbol{\mu}_\alpha\boldsymbol{\mu}_\alpha'\right)\text{vec}(\mathbf{A}) \\
&\quad+\gamma_2\text{tr}(\mathbf{C}\odot\mathbf{C}) + 4\gamma_1\boldsymbol{\mu}_\alpha'\mathbf{P}_\alpha'^{-1}\mathbf{C}(\mathbf{I}_N\odot\mathbf{C})\mathbf{e}.
\end{aligned}$$

For $\mathbf{d}$ as defined above and

$$\begin{aligned} \mathbf{G}_1 &= \mathbf{D}_N' \left(\boldsymbol{\Omega}_\alpha \otimes \boldsymbol{\Omega}_\alpha \right) \mathbf{D}_N, \\ \mathbf{G}_2 &= \mathbf{D}_N' \left(\boldsymbol{\Omega}_\alpha \otimes \boldsymbol{\mu}_\alpha \boldsymbol{\mu}_\alpha' \right) \mathbf{D}_N, \end{aligned}$$

one gets

$$\begin{aligned} \mathrm{var}(Q_{1,N}) &= 2E\left(\mathbf{a}'\mathbf{G}_1\mathbf{a}\right) + 4E\left(\mathbf{a}'\mathbf{G}_2\mathbf{a}\right) + \gamma_2 \mathrm{tr}\left[E\left(\mathbf{C} \odot \mathbf{C}\right)\right] \\ &+ 4\gamma_1 \boldsymbol{\mu}_\alpha' \mathbf{P}_\alpha'^{-1} E\left[\mathbf{C}(\mathbf{I}_N \odot \mathbf{C})\right]\mathbf{e} + \mathrm{var}\left(\mathbf{d}'\mathbf{a}\right) \end{aligned}$$

and noting that $\mathbf{G}_1, \mathbf{G}_2$ and $\mathbf{d}$ are non-stochastic, we have

$$\begin{aligned} \mathrm{var}(Q_{1,N}) &= 2\mathrm{tr}\left[\mathbf{G}_1\left(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}_a'\right)\right] + 4\mathrm{tr}\left[\mathbf{G}_2\left(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}_a'\right)\right] + \mathbf{d}'\boldsymbol{\Omega}_a\mathbf{d} \\ &\quad + \gamma_2 \mathrm{tr}\left[E\left(\mathbf{C} \odot \mathbf{C}\right)\right] + 4\gamma_1 \boldsymbol{\mu}_\alpha' \mathbf{P}_\alpha'^{-1} E\left[\mathbf{C}(\mathbf{I}_N \odot \mathbf{C})\right]\mathbf{e} \\ &= 2\mathrm{tr}\left(\mathbf{G}_1\boldsymbol{\Omega}_a\right) + 4\mathrm{tr}\left(\mathbf{G}_2\boldsymbol{\Omega}_a\right) + 2\boldsymbol{\mu}_a'\mathbf{G}_1\boldsymbol{\mu}_a + 4\boldsymbol{\mu}_a'\mathbf{G}_2\boldsymbol{\mu}_a + \mathbf{d}'\boldsymbol{\Omega}_a\mathbf{d} \\ &\quad + \gamma_2 \mathrm{tr}\left[E\left(\mathbf{C} \odot \mathbf{C}\right)\right] + 4\gamma_1 \boldsymbol{\mu}_\alpha' \mathbf{P}_\alpha'^{-1} E\left[\mathbf{C}(\mathbf{I}_N \odot \mathbf{C})\right]\mathbf{e}. \ \square \end{aligned}$$

Lemma A.2 Let

$$Q_{2,N} = \boldsymbol{\gamma}'\mathbf{A}\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha}$ and $\mathbf{A}$ satisfy the same assumptions of Lemma A.1 and $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_N)'$ has finite fourth moment with $\boldsymbol{\Sigma}_{\gamma\alpha} = E(\boldsymbol{\gamma}\boldsymbol{\alpha}')$, $\boldsymbol{\sigma}_{\gamma\alpha} = \mathrm{vec}(\boldsymbol{\Sigma}_{\gamma\alpha})$ and $\boldsymbol{\Sigma} = E\left(\boldsymbol{\alpha}\boldsymbol{\alpha}' \otimes \boldsymbol{\gamma}\boldsymbol{\gamma}'\right)$. When

$(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ and $\mathbf{A}$ are mutually independent, element by element

then:

$$\mathrm{var}(Q_{2,N}) = \mathrm{tr}\left(\boldsymbol{\Omega}_a\mathbf{D}_N'\boldsymbol{\Sigma}\mathbf{D}_N\right) + \boldsymbol{\mu}_a'\mathbf{D}_N'\boldsymbol{\Sigma}\mathbf{D}_N\boldsymbol{\mu}_a - \left(\boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\boldsymbol{\mu}_a\right)^2.$$

Proof: the result follows substituting into (100)

$$\begin{aligned} E(Q_{2,N}\,|\mathbf{A}) &= \mathrm{tr}\left(\mathbf{A}\boldsymbol{\gamma}\boldsymbol{\alpha}'\right) = \mathrm{tr}\left(\boldsymbol{\Sigma}_{\gamma\alpha}\mathbf{A}\right) = \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\mathbf{a}. \\ \mathrm{var}\left[E(Q_{2,N}\,|\mathbf{A})\right] &= \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}_N'\boldsymbol{\sigma}_{\gamma\alpha} \end{aligned}$$

and

$$\begin{aligned} \mathrm{var}(Q_{2,N}\,|\mathbf{A}) &= E(\boldsymbol{\alpha}'\mathbf{A}\boldsymbol{\gamma}\boldsymbol{\gamma}'\mathbf{A}\boldsymbol{\alpha}\,|\mathbf{A}) - \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\mathbf{a}\mathbf{a}'\mathbf{D}_N'\boldsymbol{\sigma}_{\gamma\alpha} \\ &= \mathrm{tr}\left[E\left(\boldsymbol{\alpha}\boldsymbol{\alpha}'\mathbf{A}\boldsymbol{\gamma}\boldsymbol{\gamma}'\mathbf{A}\,|\mathbf{A}\right)\right] - \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\mathbf{a}\mathbf{a}'\mathbf{D}_N'\boldsymbol{\sigma}_{\gamma\alpha} \\ &= \mathbf{a}'\mathbf{D}_N'\left[E\left(\boldsymbol{\alpha}\boldsymbol{\alpha}' \otimes \boldsymbol{\gamma}\boldsymbol{\gamma}'\right)\right]\mathbf{D}_N\mathbf{a} - \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\mathbf{a}\mathbf{a}'\mathbf{D}_N'\boldsymbol{\sigma}_{\gamma\alpha} \\ E\left[\mathrm{var}(Q_{2,N}\,|\mathbf{A})\right] &= \mathrm{tr}\left[\left(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}_a'\right)\mathbf{D}_N'\boldsymbol{\Sigma}\mathbf{D}_N\right] - \boldsymbol{\sigma}_{\gamma\alpha}'\mathbf{D}_N\left(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}_a'\right)\mathbf{D}_N'\boldsymbol{\sigma}_{\gamma\alpha}. \ \square \end{aligned}$$

Lemma A.3 Let
$$Q_{3,N} = \boldsymbol{\delta}'\mathbf{A}\boldsymbol{\alpha},$$
where $\boldsymbol{\alpha}$ and $\mathbf{A}$ satisfy the same assumptions of Lemma A.1 and $\boldsymbol{\delta} = (\delta_1, ..., \delta_N)'$ is a $N \times 1$ non random vector. Then

$$\text{var}(Q_{3,N}) = \delta' E(\mathbf{A}\boldsymbol{\Omega}_\alpha \mathbf{A})\boldsymbol{\delta} + (\delta' \otimes \boldsymbol{\mu}'_\alpha)\mathbf{D}_N \boldsymbol{\Omega}_a \mathbf{D}'_N(\boldsymbol{\delta} \otimes \boldsymbol{\mu}_\alpha).$$

An alternative representation is

$$\text{var}(Q_{3,N}) = \text{tr}\left[(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\Omega}_\alpha)\mathbf{D}_N(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}'_a)\mathbf{D}'_N\right] + \text{tr}\left[(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha)\,\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}'_N\right].$$

Proof: it is clear that

$$\begin{aligned}
E(Q_{3,N}\,|\mathbf{A}\,) &= \boldsymbol{\mu}'_\alpha \mathbf{A}\boldsymbol{\delta} = (\boldsymbol{\delta}' \otimes \boldsymbol{\mu}'_\alpha)\text{vec}(\mathbf{A}) = (\boldsymbol{\delta}' \otimes \boldsymbol{\mu}'_\alpha)\mathbf{D}_N\mathbf{a}, \\
\text{var}(Q_{3,N}\,|\mathbf{A}\,) &= \delta'\mathbf{A}\boldsymbol{\Omega}_\alpha\mathbf{A}\boldsymbol{\delta} \\
&= \text{tr}\,(\boldsymbol{\Omega}_\alpha \mathbf{A}\boldsymbol{\delta}\boldsymbol{\delta}'\mathbf{A}) \\
&= \text{vec}(\mathbf{A})'(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\Omega}_\alpha)\text{vec}(\mathbf{A}) \\
&= \mathbf{a}'\mathbf{D}'_N(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\Omega}_\alpha)\mathbf{D}_N\mathbf{a}.
\end{aligned}$$

The first representation uses simply the expectation of the first term on the right hand side of $\text{var}(Q_{3,N}\,|\mathbf{A}\,)$ above together with

$$\begin{aligned}
\text{var}\left[E(Q_{3,N}\,|\mathbf{A}\,)\right] &= (\delta' \otimes \boldsymbol{\mu}'_\alpha)\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}'_N(\boldsymbol{\delta} \otimes \boldsymbol{\mu}_\alpha) \\
&= \text{tr}\left[(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\mu}_\alpha\boldsymbol{\mu}'_\alpha)\,\mathbf{D}_N\boldsymbol{\Omega}_a\mathbf{D}'_N\right].
\end{aligned}$$

The second representation relies on

$$\begin{aligned}
E\left[\text{var}(Q_{3,N}\,|\mathbf{A}\,)\right] &= E\left[\mathbf{a}'\mathbf{D}'_N(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\Omega}_\alpha)\mathbf{D}_N\mathbf{a}\right] \\
&= \text{tr}\left[(\boldsymbol{\delta}\boldsymbol{\delta}' \otimes \boldsymbol{\Omega}_\alpha)\mathbf{D}_N(\boldsymbol{\Omega}_a + \boldsymbol{\mu}_a\boldsymbol{\mu}'_a)\mathbf{D}'_N\right].
\end{aligned}$$

The results follows substituting suitably the above expressions into (100). □

Next, we introduce a lemma where we show that for a given t and as $N \to \infty$, $\boldsymbol{\Sigma}_{t-1}^{-1}$ and $\mathbf{B}$ are asymptotically orthogonal. This result turns out to be critical for characterizing the behavior of optimal portfolios as N gets large.

Lemma B *Let (6) holds. Recalling that $\mathbf{e}_i$ is the i^{th} column of the identity matrix $\mathbf{I}_N$, then for any i, t*

$$\mathbf{e}'_i\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\beta}^j \to_p 0 \text{ as } N \to \infty, \;\; 1 \leq j \leq k, \tag{101}$$

where $\boldsymbol{\beta}^j$ denotes the j-th column of $\mathbf{B}$*.*

Let (6) and (14) hold. Then for any t

$$\| \boldsymbol{\Sigma}_t^{-1}\boldsymbol{\beta}^j \|^2 = O_p(N^{-1}), \ \ 1 \leq j \leq k, \quad as\ N \to \infty. \tag{102}$$

Proof: the results follow from the identity

$$\boldsymbol{\Sigma}_t^{-1} = \mathbf{H}_t - \mathbf{H}_t\mathbf{B}(N^{-1}\boldsymbol{\Omega}_t^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-1}\mathbf{B}'\mathbf{H}_t. \tag{103}$$

Pre-multiplying both sides by $\mathbf{e}_i'$ and post-multiplying both sides by $\boldsymbol{\beta}^j$ yields (101).

We deal with (102) more explicitly. First note that, setting $\mathbf{e}_j$ to be the j^{th} column of the identity $\mathbf{I}_k$ matrix,

$$\begin{aligned}
& (N^{-1}\boldsymbol{\Omega}_t^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-1}\mathbf{B}'\mathbf{H}_t\boldsymbol{\beta}^j - \mathbf{e}_j \\
= \ & (N^{-1}\boldsymbol{\Omega}_t^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-1}\mathbf{B}'\mathbf{H}_t\boldsymbol{\beta}^j - (N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-1}\mathbf{B}'\mathbf{H}_t\boldsymbol{\beta}^j \\
= \ & N^{-1}\left[-(N^{-1}\boldsymbol{\Omega}_t^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\boldsymbol{\Omega}_t^{-1}(N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-1}\mathbf{B}'\mathbf{H}_t\boldsymbol{\beta}^j\right] \\
\equiv \ & N^{-1}\mathbf{g}_t^{(j)},
\end{aligned}$$

where notice that $\mathbf{g}_t^{(j)}$ is a $k \times 1$ vector with a finite norm.

Therefore, substituting the latter expression into (103) and recalling that $\mathbf{B}\mathbf{e}_j = \boldsymbol{\beta}^j$ it follows that

$$\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\beta}^j = \mathbf{H}_t\boldsymbol{\beta}^j - \mathbf{H}_t\mathbf{B}(\mathbf{e}_j + N^{-1}\mathbf{g}_t^{(j)}),$$

yielding

$$\begin{aligned}
\| \ \ & \boldsymbol{\Sigma}_t^{-1}\boldsymbol{\beta}^j \|^2 = \boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j + N^{-1}\mathbf{g}_t^{(j)\prime}\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j - \boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j \\
& -N^{-1}\mathbf{g}_t^{(j)\prime}\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j - N^{-1}\boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\mathbf{B}\mathbf{g}_t^{(j)} - N^{-2}\mathbf{g}_t^{(j)\prime}\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B}\mathbf{g}_t^{(j)} \\
& -\boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j + \boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\boldsymbol{\beta}^j + N^{-1}\boldsymbol{\beta}^{j\prime}\mathbf{H}_t\mathbf{H}_t\mathbf{B}\mathbf{g}_t^{(j)} \\
= \ \ & N^{-1}\mathbf{g}_t^{(j)\prime}\left(N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{H}_t\mathbf{B}\right)\mathbf{g}_t^{(j)} = O_p(N^{-1}\mathbf{g}_t^{(j)\prime}\mathbf{F}_t\mathbf{g}_t^{(j)}). \ \square
\end{aligned}$$

Part B: Primitive assumptions

Here we show formally how Assumptions 2, 4 and 5 are implied by the primitive Assumptions 6, 7 and 8. Using the conditional *cross-sectional* mean and variance operators $E_t(\cdot)$ and $\mathrm{var}_t(\cdot)$, defined in Section 2.1, we establish mean-square convergence conditional on the sigma-algebra of cross sectionally-invariant events, $\mathcal{A}_t$. For non-time varying arguments we use the standard operators $E(\cdot)$ and var$(\cdot)$.

Consider first Assumption 2. Then (5) follows by (29) and Lemma A.3 with $\boldsymbol{\delta} = \mathbf{e}$, $\boldsymbol{\alpha} = \boldsymbol{\beta}^j$, $\mathbf{A} = \mathbf{I}_N$ yielding $\mathbf{C} = \mathbf{P}'_{\beta j}\mathbf{P}_{\beta j}$ with c_{rs} as its $(r,s)^{th}$ element, $\mu_A = \text{vech}(\mathbf{I}_N)$, $\boldsymbol{\Omega}_A = \mathbf{0}$. Hence, by (33),

$$\text{var}(\boldsymbol{\beta}^{j\prime}\mathbf{e}/N) = \mathbf{e}'\boldsymbol{\Omega}_{\beta j}\mathbf{e}/N^2 \leq \rho(\boldsymbol{\Omega}_{\beta j})/N \leq \rho^2(\mathbf{P}_{\beta j})/N = O_p(N^{-\frac{1}{2}}),$$

yielding convergence in quadratic mean of $\boldsymbol{\beta}^{j\prime}\mathbf{e}/N$ to $\mu_{\beta j}$. (6) follows by convergence in quadratic mean since (30) holds and

$$\text{var}_t(\boldsymbol{\beta}^{j\prime}\mathbf{H}_t\boldsymbol{\beta}^l) = o_p(N^2),$$

using Lemmas A.1, with $\boldsymbol{\alpha}$ set to $\boldsymbol{\beta}^j$, $\mathbf{A}$ to $\mathbf{H}_t$ and $\mathbf{C}$ to $\mathbf{P}'_{\beta j}\mathbf{H}_t\mathbf{P}_{\beta j}$.

Using repetitively the bound $\rho(\boldsymbol{\Omega}_{\beta j}) \leq \rho^2(\mathbf{P}_{\beta j})$, and the result $\text{tr}(\mathbf{DAD}') \leq \rho(\mathbf{A})tr(\mathbf{DD}')$ for any matrix $\mathbf{D}$ and positive semi-definite $\mathbf{A}$ (see Lutkepohl (1996, p.44)), we have

$$\begin{aligned} E_t\left[\text{tr}(\mathbf{C}_t^2)\right] = E_t\left[\text{tr}(\boldsymbol{\Omega}_{\beta j}\mathbf{H}_t)^2\right] &\leq \text{tr}(\boldsymbol{\Omega}_{\beta j}^2)E_t\left[\rho^2(\mathbf{H}_t)\right] \\ &\leq \rho(\boldsymbol{\Omega}_{\beta j})\text{tr}(\boldsymbol{\Omega}_{\beta j})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{\frac{3}{2}}), \\ E_t\left(\boldsymbol{\mu}_\beta^{j\prime}\mathbf{H}_t\boldsymbol{\Omega}_{\beta j}\mathbf{H}_t\boldsymbol{\mu}_\beta^j\right) &\leq \boldsymbol{\mu}_\beta^{j\prime}\boldsymbol{\mu}_\beta^j\rho(\boldsymbol{\Omega}_{\beta j})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{\frac{3}{2}}), \end{aligned}$$

and

$$\begin{aligned} E_t\left[\text{tr}\left(\mathbf{C}_t \odot \mathbf{C}_t\right)\right] &= E_t\left\{\text{tr}[(\mathbf{I}_N \odot \mathbf{C}_t)\mathbf{C}_t]\right\} = \sum_{i=1}^N E_t(c_{ii,t}^2) \leq N E_t\left[\rho^2(\mathbf{C}_t)\right] \\ &= N\rho^2(\boldsymbol{\Omega}_{\beta j})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^2), \end{aligned}$$

where $\mathbf{C}_t = \mathbf{P}'_{\beta j}\mathbf{H}_t\mathbf{P}_{\beta j}$, $(\mathbf{I}_N \odot \mathbf{C}_t)$ is a diagonal matrix with $c_{ii,t}$ on the $(i,i)^{th}$ position, and $\max_{i=1,\ldots,N} \mid c_{ii,t} \mid \leq \rho(\mathbf{C}_t)$, a.s. Also

$$\begin{aligned} \left|\boldsymbol{\mu}_\beta^{j\prime}\mathbf{P}_{\beta j}^{\prime -1}E_t\left(\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right)\mathbf{e}\right| &\leq \left(\boldsymbol{\mu}_\beta^{j\prime}\boldsymbol{\Omega}_{\beta j}^{-1}\boldsymbol{\mu}_\beta^j\right)^{\frac{1}{2}}\left\{\mathbf{e}'E_t\left[(\mathbf{I}_N \odot \mathbf{C}_t)\mathbf{C}_t\right]E_t\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\mathbf{e}\right\}^{\frac{1}{2}} \\ &\leq \left(\boldsymbol{\mu}_\beta^{j\prime}\boldsymbol{\mu}_\beta^j\right)^{\frac{1}{2}}\rho^{\frac{1}{2}}(\boldsymbol{\Omega}_{\beta j}^{-1})N^{\frac{1}{2}}\rho\left\{E_t[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)]\right\} = o_p(N^2). \end{aligned}$$

But, since $\text{tr}(\mathbf{C}_t \odot \mathbf{C}_t) = \text{tr}\left[(\mathbf{I}_N \odot \mathbf{C}_t)\mathbf{C}_t\right]$ by Magnus and Neudecker (2001, Theorem 7, p.46) and $\rho(\cdot)$ is a convex function (Magnus and Neudecker 2001, Theorem 5, p.205), then $\rho\left\{E_t\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\right\} \leq E_t\left\{\rho\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\right\} \leq E_t\left[\rho^2(\mathbf{C}_t)\right] \leq \rho^4(\mathbf{P}_{\beta j})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N)$. The expression (35) is redundant when $\gamma_1 = 0$.

For the off-diagonal terms of $\mathbf{B}'\mathbf{H}_t\mathbf{B}$ use Lemmas A.2 with $\boldsymbol{\alpha}$ set to $\boldsymbol{\beta}^j, \boldsymbol{\gamma}$ to $\boldsymbol{\beta}^l$, and $\mathbf{A}$ to $\mathbf{H}_t$, and note that

$$\begin{aligned} \boldsymbol{\mu}'_{ht}\mathbf{D}'_N\boldsymbol{\Omega}_{\beta jl}\mathbf{D}_N\boldsymbol{\mu}_{ht} &\leq \boldsymbol{\mu}'_{ht}\mathbf{D}'_N\mathbf{D}_N\boldsymbol{\mu}_{ht}\rho(\boldsymbol{\Omega}_{\beta jh}) = o_p(N^2), \\ \text{tr}\left(\boldsymbol{\Omega}_{ht}\mathbf{D}'_N\boldsymbol{\Sigma}_{\beta jh}\mathbf{D}_N\right) &\leq \rho(\mathbf{D}_N\boldsymbol{\Omega}_{ht}\mathbf{D}'_N)\text{tr}(\boldsymbol{\Sigma}_{\beta jh}) = o_p(N^2), \end{aligned}$$

where recall that $\mathbf{\Omega}_{\beta jl} = \mathbf{\Sigma}_{\beta jl} - E\left[\mathrm{vec}(\boldsymbol{\beta}^l \boldsymbol{\beta}^{j'})\right] E\left[\mathrm{vec}'(\boldsymbol{\beta}^l \boldsymbol{\beta}^{j'})\right]$, with $\mathrm{vec}(\boldsymbol{\beta}^l \boldsymbol{\beta}^{j'}) = (\boldsymbol{\beta}^j \otimes \boldsymbol{\beta}^l)$.

Now consider Assumption 4. Then (14) easily follows much in the same way as done for (6), although we need the slightly stronger conditions (31), (42), (43), and (44). Next, (15) follows by (38) and Lemma A.3 with $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ set to $\mathbf{e}$, and $\mathbf{A}$ set to $\mathbf{H}_t$, since

$$\mathrm{var}_t\left(\mathbf{e}'\mathbf{H}_t\mathbf{e}\right) = (\mathbf{e}' \otimes \mathbf{e}')(\mathbf{D}_N \mathbf{\Omega}_{h,t} \mathbf{D}_N')(\mathbf{e} \otimes \mathbf{e}) \leq N^2 \rho(\mathbf{D}_N \mathbf{\Omega}_{h,t} \mathbf{D}_N').$$

Finally, consider Assumption 5. By Lemma A.3 with $\boldsymbol{\alpha}$ set to $\mathbf{v}_t$, $\boldsymbol{\delta}$ to $\mathbf{e}$, $\mathbf{A}$ to $\mathbf{H}_t$ and $\mathbf{C}$ to $\mathbf{P}_{vt}'\mathbf{H}_t\mathbf{P}_{vt}$, (19) follows by convergence in quadratic mean since

$$\mathrm{var}_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{e}) = o_p(N)$$

and (45) holds. In fact,

$$\begin{aligned}
&(\mathbf{e}' \otimes \boldsymbol{\mu}_{vt}')\mathbf{D}_N \mathbf{\Omega}_{ht} \mathbf{D}_N'(\mathbf{e} \otimes \boldsymbol{\mu}_{vt}) \leq N \boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt} \rho(\mathbf{D}_N \mathbf{\Omega}_{ht} \mathbf{D}_N') = o_p(N), \\
&E_t\left(\mathbf{e}'\mathbf{H}_t \mathbf{\Omega}_{vt} \mathbf{H}_t \mathbf{e}\right) \leq N \rho(\mathbf{\Omega}_{vt}) E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{\frac{1}{2}}).
\end{aligned}$$

To establish (20) we use (46) and Lemma A.1 with $\boldsymbol{\alpha}$ set to $\mathbf{v}_t$, $\mathbf{A}$ to $\mathbf{H}_t$. In fact, following the same steps used above to get the variance of $\boldsymbol{\beta}^{j'}\mathbf{H}_t\boldsymbol{\beta}^j$,

$$E_t\left[\mathrm{tr}(\mathbf{C}_t^2)\right] \leq \mathrm{tr}(\mathbf{\Omega}_{vt}^2) E_t\left[\rho^2(\mathbf{H}_t)\right] \ \leq \rho(\mathbf{\Omega}_{vt}) \mathrm{tr}(\mathbf{\Omega}_{vt}) E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{-\frac{1}{2}}),$$

$$E_t\left(\boldsymbol{\mu}_{vt}\mathbf{H}_t \mathbf{\Omega}_{vt} \mathbf{H}_t \boldsymbol{\mu}_{vt}\right) \leq \boldsymbol{\mu}_{v,t}'\boldsymbol{\mu}_{v,t} \rho(\mathbf{\Omega}_{v,t}) E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{-\frac{1}{2}}),$$

$$E_t\left[\mathrm{tr}\left(\mathbf{C}_t \odot \mathbf{C}_t\right)\right] \leq N E_t\left[\rho^2(\mathbf{C}_t)\right] = N \rho^4(\mathbf{P}_{v,t}) E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(1),$$

and

$$\left|\boldsymbol{\mu}_{vt}'\mathbf{P}_{vt}'^{-1} E_t\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\mathbf{e}\right| \leq N^{\frac{1}{2}} \left(\boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt}\right)^{\frac{1}{2}} \rho^{\frac{1}{2}}(\mathbf{\Omega}_{vt}^{-1}) \rho\left\{E_t[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)]\right\} = o_p(N^{-\frac{1}{4}}),$$

since, as already seen, $\rho\left\{E_t[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)]\right\} \leq \rho^4(\mathbf{P}_{v,t}) E_t\left[\rho^2(\mathbf{H}_t)\right]$ which is $o_p(N^{-1})$. Therefore $\mathrm{var}_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t) = o_p(1)$. Condition (50) is redundant when $\mathbf{v}_t$ has a symmetric distribution around the mean.

Following the same steps, one establishes (23). In fact, (51) holds and by Lemma A.3 with $\boldsymbol{\alpha}$ set to $\mathbf{v}_t$, $\boldsymbol{\delta}$ to $\mathbf{e}$, and $\mathbf{A}$ to $\mathbf{H}_t$, then $\mathrm{var}_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{e}) = o_p(N^2)$. Furthermore

$$\begin{aligned}
&(\mathbf{e}' \otimes \boldsymbol{\mu}_{vt}')\mathbf{D}_N \mathbf{\Omega}_{ht} \mathbf{D}_N'(\mathbf{e} \otimes \boldsymbol{\mu}_{vt}) \leq N \boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt} \rho(\mathbf{D}_N \mathbf{\Omega}_{ht} \mathbf{D}_N') = o_p(N^2), \\
&\mathbf{e}'E_t(\mathbf{H}_t \mathbf{\Omega}_{vt} \mathbf{H}_t)\mathbf{e} \leq N \rho(\mathbf{\Omega}_{vt}) E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^2).
\end{aligned}$$

To establish (24) we use (52) and rely on Lemma A.1 setting $\boldsymbol{\alpha}$ to $\mathbf{v}_t$, and $\mathbf{A}$ to $\mathbf{H}_t$ yielding $\text{var}_t(\mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t) = o_p(N^2)$ since

$$\begin{aligned}
E_t\left[\text{tr}(\mathbf{C}_t^2)\right] &\leq \text{tr}(\boldsymbol{\Omega}_{vt}^2)E_t\left[\rho^2(\mathbf{H}_t)\right] \leq \rho(\boldsymbol{\Omega}_{vt})\text{tr}(\boldsymbol{\Omega}_{vt})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N^{\frac{3}{2}}), \\
E_t\left(\boldsymbol{\mu}_{vt}\mathbf{H}_t\boldsymbol{\Omega}_{vt}\mathbf{H}_t\boldsymbol{\mu}_{vt}\right) &\leq \boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt}\rho(\boldsymbol{\Omega}_{vt})E_t[\rho^2(\mathbf{H}_t)] = o_p(N^{\frac{3}{2}}), \\
E_t\left[\text{tr}\left(\mathbf{C}_t \odot \mathbf{C}_t\right)\right] &\leq NE_t[\rho^2(\mathbf{C}_t)] = N\rho^4(\mathbf{P}_{vt})E_t[\rho^2(\mathbf{H}_t)] = o_p(N^2), \\
\left|\boldsymbol{\mu}_{vt}'\mathbf{P}_{vt}'^{-1}E_t\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\mathbf{e}\right| &\leq N^{\frac{1}{2}}\left(\boldsymbol{\mu}_{vt}'\boldsymbol{\mu}_{vt}\right)^{\frac{1}{2}}\rho^{\frac{1}{2}}(\boldsymbol{\Omega}_{vt}^{-1})\rho\left\{E_t\left[\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t)\right]\right\} = o_p(N^2),
\end{aligned}$$

where all the arguments used earlier still apply and noting in particular that $\rho\left[E_t(\mathbf{C}_t(\mathbf{I}_N \odot \mathbf{C}_t))\right] \leq \rho^4(\mathbf{P}_{vt})E_t\left[\rho^2(\mathbf{H}_t)\right] = o_p(N)$. Condition (56) is redundant when $\mathbf{v}_t$ has a symmetric distribution around the mean.

In so far as (21) and (25) are concerned, one follows the same steps used for (20) and (24) but, in addition, we need to strengthen (39), (40), and (41) by (42), (43), and (44), respectively.

Part C: Proofs of Theorems

Proof of Theorem 1
By simple steps

$$\begin{aligned}
& \left[\mathbf{I}_N - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\right]\boldsymbol{\Sigma}_{t-1}^{-1} \\
= & \left[\mathbf{I}_N - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\right]\left[\mathbf{H}_{t-1} - \mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\right] \\
= & \mathbf{H}_{t-1} - \mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1} - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1} \\
& + \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1} \\
= & \mathbf{H}_{t-1} - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1} \\
= & \left[\mathbf{I}_N - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\right]\left[\mathbf{H}_{t-1} - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\right] \\
= & \left[\mathbf{I}_N - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\right]\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}
\end{aligned}$$

and

$$\left[\mathbf{I}_N - \mathbf{H}_{t-1}\mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\right]\breve{\boldsymbol{\Sigma}}_{t-1}^{-1} = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}. \tag{104}$$

Recalling that $\boldsymbol{\mu}_{t-1} = \mathbf{B}\boldsymbol{\lambda}_{t-1} + \mathbf{v}_{t-1} = \mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1} + \hat{\mathbf{v}}_{t-1}$ where

$$\hat{\mathbf{v}}_{t-1} = [\mathbf{I}_N - \mathbf{B}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}]\mathbf{v}_{t-1},$$

mow yields

$$\mathbf{H}_{t-1}\hat{\mathbf{v}}_{t-1} = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{v}_{t-1} = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}(\mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1}) = \breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\boldsymbol{\mu}_{t-1}, \tag{105}$$

and $\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} = \boldsymbol{\Sigma}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1}$, thus establishing (63). Since $\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{B} = \mathbf{0}$, it follows that

$$\hat{\mathbf{v}}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B} = \mathbf{0}. \tag{106}$$

Now (64) follows (with $\kappa_t = 1$) recalling that

$$\rho_t^{ap} - r_{0,t-1} = (\mathbf{z}_t + \hat{\boldsymbol{\lambda}}_{t-1})'\mathbf{B}'\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} + (\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1})'\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} = (\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1})'\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1}.$$

Finally (65) follows since

$$\begin{aligned} & \rho_t^{mv} - r_{0,t-1} = (\mathbf{r}_t - r_{0,t-1}\mathbf{e})'\mathbf{w}_{t-1}^{mv} = [\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1} + \mathbf{B}(\mathbf{z}_t + \hat{\boldsymbol{\lambda}}_{t-1})]'\boldsymbol{\Sigma}_{t-1}^{-1}(\hat{\mathbf{v}}_{t-1} + \mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1}) \\ = \; & (\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1})'\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\hat{\mathbf{v}}_{t-1} + \boldsymbol{\varepsilon}_t'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\hat{\lambda}_{t-1} + [\mathbf{B}(\mathbf{z}_t + \hat{\boldsymbol{\lambda}}_{t-1})]'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}\hat{\boldsymbol{\lambda}}_{t-1}. \end{aligned}$$

and recognizing that the first term after the third equality sign is $\rho_t^{ap} - r_{0,t-1}$.
□

All the limits below are based on $N \to \infty$, and the normalization $\kappa_{t-1} = 1$ is used throughout.

Proof of Theorem 2
(i) Since in general

$$w_{it}^{mv} = \mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\mu}_t = \mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\ \mathbf{v}_t + \mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\ \boldsymbol{\lambda}_t, \tag{107}$$

then for the exact pricing case $\mathbf{v}_t = \mathbf{0}$, and using

$$\boldsymbol{\Sigma}_t^{-1}\mathbf{B} = \mathbf{H}_t\mathbf{B}(\boldsymbol{\Omega}_t^{-1} + \mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\boldsymbol{\Omega}_t^{-1}. \tag{108}$$

under the assumptions we have

$$w_{it}^{mv} = \mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\ \boldsymbol{\lambda}_t = N^{-1}\mathbf{e}_i'\ \mathbf{H}_t\mathbf{B}(N^{-1}\boldsymbol{\Omega}_t^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\boldsymbol{\Omega}_t^{-1}\ \boldsymbol{\lambda}_t = O_p(N^{-1}). \tag{109}$$

Similarly, for the `ap` strategy, recalling that $\breve{\boldsymbol{\Sigma}}_t^{-1}\mathbf{B} = \mathbf{0}$,

$$w_{it}^{ap} = \mathbf{e}_i'\breve{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}_t = \mathbf{e}_i'\breve{\boldsymbol{\Sigma}}_t^{-1}\ \mathbf{v}_t,$$

and the result (71) follows since $\mathbf{v}_t = \mathbf{0}$.
(ii) In general

$$\begin{aligned} \boldsymbol{\rho}_t^{mv} - r_{0,t-1} &= \mathbf{w}_{t-1}^{mv\prime}(\mathbf{r}_t - r_{0,t-1}\mathbf{e}) \\ &= (\ \mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}(\ \boldsymbol{\lambda}_{t-1} + \mathbf{z}_t) && (110) \\ &\quad +(\ \mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})'\boldsymbol{\Sigma}_{t-1}^{-1}\ \mathbf{v}_{t-1} && (111) \\ &\quad +(\ \mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t. && (112) \end{aligned}$$

Using (108)

$$\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B} \to_p \boldsymbol{\Omega}_{t-1}^{-1} \tag{113}$$

implying that (110) converges in probability to $\boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}^{-1}_{t-1}(\ \boldsymbol{\lambda}_{t-1}+\mathbf{z}_t)$ since $\mathbf{v}_{t-1}=\mathbf{0}$. When the second part of (72) holds, using (103), by Billingsley (1986, Lemma 1, p.193)

$$N^{\frac{1}{2}}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{t-1} \rightarrow_d N\left(0, \boldsymbol{\Omega}^{-1}_{t-1}(\mathbf{A}^{-1}_{t-1}+a_{t-1}\boldsymbol{\mu}_\beta\boldsymbol{\mu}'^{-1}_\beta\boldsymbol{\Omega}^{-1}_{t-1})\right), \qquad (114)$$

implying that term (112) involving $\boldsymbol{\varepsilon}_t$ is $O_p(N^{-\frac{1}{2}})$. Since under exact pricing all terms involving $\mathbf{v}_{t-1}$ are null, one gets immediately (73). For (74) direct calculation yields

$$\mu^{mv}_{\rho,t-1}-r_{0,t-1}=\boldsymbol{\lambda}'_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}\ \boldsymbol{\lambda}_{t-1},$$
$$\sigma^{mv2}_{\rho,t-1}=\boldsymbol{\lambda}'_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}\boldsymbol{\Omega}_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}\ \boldsymbol{\lambda}_{t-1}+\ \boldsymbol{\lambda}'_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{G}_{t-1}\ \boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}\ \boldsymbol{\lambda}_{t-1}.$$

Then, by (113),

$$\mu^{mv}_{\rho,t-1}-r_{0,t-1}\rightarrow_p \boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}^{-1}_{t-1}\ \boldsymbol{\lambda}_{t-1},\ \ \sigma^{mv2}_{\rho,t-1}\rightarrow_p \boldsymbol{\lambda}'_{t-1}\boldsymbol{\Omega}^{-1}_{t-1}\ \boldsymbol{\lambda}_{t-1},$$

since for the second term on the right hand side of $\sigma^{mv2}_{\rho,t-1}$

$$\boldsymbol{\lambda}'_{t-1}\mathbf{B}'\boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{G}_{t-1}\ \boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}\ \boldsymbol{\lambda}_{t-1}=\ \boldsymbol{\lambda}'_{t-1}\mathbf{B}'\ \boldsymbol{\Sigma}^{-1}_{t-1}\mathbf{B}(\boldsymbol{\Omega}^{-1}_{t-1}+\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\boldsymbol{\Omega}^{-1}_{t-1}\ \boldsymbol{\lambda}_{t-1}=O_p(N^{-1}). \qquad (115)$$

For the `ap` strategy, using (64) and (105), we have

$$\rho^{ap}_t=r_{0,t-1}+(\boldsymbol{\varepsilon}_t+\hat{\mathbf{v}}_{t-1})'\breve{\boldsymbol{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1}. \qquad (116)$$

Hence, under $\mathbf{v}_{t-1}=0$, $\rho^{ap}_t=r_{0,t-1}$ as required. □

Proof of Theorem 3

(i) Consider (107). By identity (103), and recalling that here we are considering the asymptotic no-arbitrage pricing case, we have

$$\begin{aligned}\mathbf{e}'_i\boldsymbol{\Sigma}^{-1}_t\mathbf{v}_t &= \mathbf{e}'_i\left[\mathbf{H}_t\mathbf{v}_t-\mathbf{H}_t\mathbf{B}(\boldsymbol{\Omega}^{-1}_t+\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t\right]\\ &= \mathbf{e}'_i\mathbf{H}_t\mathbf{v}_t-N^{-\frac{1}{2}}\mathbf{e}'_i\mathbf{H}_t\mathbf{B}(N^{-1}\boldsymbol{\Omega}^{-1}_t+N^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}N^{-\frac{1}{2}}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t\\ &= \left[\mathbf{e}'_i\mathbf{H}_t\mathbf{v}_t-N^{-\frac{1}{2}}\mathbf{e}'_i\mathbf{H}_t\mathbf{B}(\mathbf{A}_t+a_t\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta)^{-1}\boldsymbol{\mu}_\beta c_t\right](1+o_p(1)),\end{aligned}$$

where

$$N^{-\frac{1}{2}}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t\rightarrow_p \boldsymbol{\mu}_\beta c_t, \qquad (117)$$

by the row-wise independence between $\mathbf{B}$ and $\mathbf{H}_t\mathbf{v}_t$. By the identity

$$\mathbf{D}^{-1}_t=(\mathbf{A}_t+a_t\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta)^{-1}=\mathbf{A}^{-1}_t-\frac{a_t}{(1+a_t\boldsymbol{\mu}'_\beta\mathbf{A}^{-1}_t\boldsymbol{\mu}_\beta)}\mathbf{A}^{-1}_t\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta\mathbf{A}^{-1}_t, \qquad (118)$$

it follows that

$$\mathbf{D}_t^{-1}\boldsymbol{\mu}_\beta c_t = c_t b_t^{-1}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta,$$

yielding

$$\begin{aligned}
\mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\mathbf{v}_t &= \frac{1}{b_t}\mathbf{e}_i'\mathbf{H}_t\left[\mathbf{v}_t + (a_t\,\mathbf{v}_t\boldsymbol{\mu}_\beta' - N^{-\frac{1}{2}}\mathbf{B}c_t)\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right](1+o_p(1)) \\
&= \left[\mathbf{e}_i'\mathbf{H}_t\mathbf{v}_t - N^{-\frac{1}{2}}(c_t/b_t)\mathbf{e}_i'\mathbf{H}_t\mathbf{B}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right](1+o_p(1)) \\
&= \mathring{w}_{it}(1+o_p(1)) = O_p(N^{-\frac{1}{2}}). \qquad (119)
\end{aligned}$$

Finally, $\mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\ \boldsymbol{\lambda}_t = O_p(N^{-1})$ by (109) which holds irrespective of the form of no-arbitrage since this term does not depend on $\mathbf{v}_{t-1}$.

Concerning the ap strategy, by Theorem 1, $\mathbf{w}_t^{ap} = \breve{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}_t = \breve{\boldsymbol{\Sigma}}_t^{-1}\ \mathbf{v}_t$. However

$$\mathbf{e}_i'\breve{\boldsymbol{\Sigma}}_t^{-1}\mathbf{v}_t - \mathbf{e}_i'\boldsymbol{\Sigma}_t^{-1}\mathbf{v}_t = \mathbf{e}_i'\mathbf{H}_t\mathbf{B}(\boldsymbol{\Omega}_t^{-1}+\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\boldsymbol{\Omega}_t^{-1}(\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t = O_p(N^{-\frac{3}{2}}),$$

and, therefore, $w_{it}^{ap} = \mathbf{e}_i'\breve{\boldsymbol{\Sigma}}_t^{-1}\ \mathbf{v}_t$ has the same limiting behaviour as (119).
(ii) Recalling (108) and (117), then $\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} = O_p(N^{-1})$ since

$$N^{\frac{1}{2}}\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B} = N^{-\frac{1}{2}}\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{B}(N^{-1}\boldsymbol{\Omega}_{t-1}^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\boldsymbol{\Omega}_{t-1}^{-1} \to_p b_{t-1}^{-1}c_{t-1}\boldsymbol{\mu}_\beta'\mathbf{A}_{t-1}^{-1}\boldsymbol{\Omega}_{t-1}^{-1}. \qquad (120)$$

Also, by (114), $\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t = O_p(N^{-\frac{1}{2}})$, implying that the dominant terms in (110), (111), and (112) are, respectively, $\boldsymbol{\lambda}_{t-1}'\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}(\boldsymbol{\lambda}_{t-1}+\mathbf{z}_t)$, $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\ \mathbf{v}_{t-1}$, and $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t$. The limit of the first term follows by (113). For the second term we establish

$$\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} \to_p e_{t-1}/b_{t-1}. \qquad (121)$$

In fact, by (103) and recalling (117),

$$\begin{aligned}
\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} &= \mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{v}_{t-1} - \mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1} \\
&= \mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{v}_{t-1} - \frac{\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{B}}{\sqrt{N}}\left(N^{-1}\boldsymbol{\Omega}_{t-1}^{-1} + N^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}\right)^{-1}\frac{\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}}{\sqrt{N}} \\
&\to_p d_{t-1} - c_{t-1}^2\boldsymbol{\mu}_\beta'(\mathbf{A}_{t-1} + a_{t-1}\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta')^{-1}\boldsymbol{\mu}_\beta.
\end{aligned}$$

By identity (118)

$$\boldsymbol{\mu}_\beta'(\mathbf{A}_{t-1} + a_{t-1}\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta')^{-1}\boldsymbol{\mu}_\beta = \boldsymbol{\mu}_\beta'\mathbf{A}_{t-1}^{-1}\boldsymbol{\mu}_\beta/b_{t-1},$$

and hence

$$\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} \to_p d_{t-1} - \frac{c_{t-1}^2}{b_{t-1}}\left(\boldsymbol{\mu}_\beta'\mathbf{A}_{t-1}^{-1}\boldsymbol{\mu}_\beta\right).$$

In view of (67) and (68), and after some re-arranging (121) follows, noting that $\boldsymbol{\mu}'_\beta \mathbf{A}_t^{-1} \boldsymbol{\mu}_\beta = (b_t - 1)/a_t$, and $e_t = d_t + (a_t d_t - c_t^2)(b_t - 1)/a_t$. For the third term, using the second part of (72), (78) and (117),

$$\mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \boldsymbol{\varepsilon}_t \mid \mathcal{A}_{t-1} \to_d x_t \sim N(0, e_{t-1}/b_{t-1}) \tag{122}$$

by a simple application of the continuous mapping theorem (see Billingsley (1986, Theorem 29.2)) applied to

$$\begin{aligned} \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \boldsymbol{\varepsilon}_t &= \mathbf{v}'_{t-1} \mathbf{H}_{t-1} \boldsymbol{\varepsilon}_t - \mathbf{v}'_{t-1} \mathbf{H}_{t-1} \ \mathbf{B} (\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}' \mathbf{H}_{t-1} \mathbf{B})^{-1} \mathbf{B}' \mathbf{H}_{t-1} \boldsymbol{\varepsilon}_t \\ &= \mathbf{v}'_{t-1} \mathbf{H}_{t-1} \boldsymbol{\varepsilon}_t - N^{-\frac{1}{2}} \mathbf{v}'_{t-1} \mathbf{H}_{t-1} \ \mathbf{B} (N^{-1} \boldsymbol{\Omega}_{t-1}^{-1} + N^{-1} \mathbf{B}' \mathbf{H}_{t-1} \mathbf{B})^{-1} N^{-\frac{1}{2}} \mathbf{B}' \mathbf{H}_{t-1} \boldsymbol{\varepsilon}_t. \end{aligned}$$

It is easy to see that $e_t > 0$, almost surely. In fact e_t is given by the sum of d_t and $(a_t d_t - c_t^2) \boldsymbol{\mu}'_\beta \mathbf{A}_t^{-1} \boldsymbol{\mu}_\beta$. The latter term is positive since $a_t d_t > c_t^2$ by our assumption and $\mathbf{A}_t$ is positive definite. The first term d_t is certainly non-negative since it equals the probability limit of the quadratic form $\mathbf{v}'_t \mathbf{H}_t \mathbf{v}_t$.

Regarding (80) consider

$$\mu_{\rho,t-1}^{mv} - r_{0,t-1} = (\mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})' \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \ \boldsymbol{\lambda}_{t-1} + (\mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})' \boldsymbol{\Sigma}_{t-1}^{-1} \ \mathbf{v}_{t-1}, \tag{123}$$

$$\begin{aligned} \sigma_{\rho,t-1}^{mv\,2} &= (\mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})' \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \boldsymbol{\Omega}_{t-1} \mathbf{B}' \boldsymbol{\Sigma}_{t-1}^{-1} (\ \mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1}) \\ &+ (\ \mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1})' \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{G}_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} (\mathbf{v}_{t-1} + \mathbf{B}\boldsymbol{\lambda}_{t-1}). \end{aligned} \tag{124}$$

By (113), (115), (120) and (121)

$$\begin{aligned} &\mu_{\rho,t-1}^{mv} - r_{0,t-1} = \boldsymbol{\lambda}'_{t-1} \mathbf{B} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \ \boldsymbol{\lambda}_{t-1} + \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \ \mathbf{v}_{t-1} + O_p(N^{-\frac{1}{2}}), \\ &\sigma_{\rho,t-1}^{mv\,2} = \boldsymbol{\lambda}'_{t-1} \mathbf{B}' \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \boldsymbol{\Omega}_{t-1} \mathbf{B}' \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \boldsymbol{\lambda}_{t-1} + \ \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{G}_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{v}_{t-1} + O_p(N^{-\frac{1}{2}}), \end{aligned}$$

where for $\sigma_{\rho,t-1}^{mv\,2}$ one reckons, by (120),

$$\begin{aligned} \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{G}_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{v}_{t-1} &= \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} (\boldsymbol{\Sigma}_{t-1} - \mathbf{B} \boldsymbol{\Omega}_{t-1} \mathbf{B}') \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{v}_{t-1} \\ &= \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{v}_{t-1} + O_p(N^{-1}). \end{aligned}$$

and

$$\mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{G}_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} \ \boldsymbol{\lambda}_{t-1} = \mathbf{v}'_{t-1} \boldsymbol{\Sigma}_{t-1}^{-1} \mathbf{B} (\boldsymbol{\Omega}_{t-1}^{-1} + \mathbf{B}' \mathbf{H}_{t-1} \mathbf{B})^{-1} \boldsymbol{\Omega}_{t-1}^{-1} \boldsymbol{\lambda}_{t-1} = O_p(N^{-\frac{3}{2}}).$$

Thus $\mu_{\rho,t-1}^{mv} - r_{0,t-1}$ and $\sigma_{\rho,t-1}^{mv\,2}$ both converge in probability to $e_{t-1}/b_{t-1} + \boldsymbol{\lambda}'_{t-1} \boldsymbol{\Omega}_{t-1}^{-1} \boldsymbol{\lambda}_{t-1}$ and (80) follows.

For the limit ap portfolio return, by repeated use of Theorem 1,

$$\begin{aligned} \rho_t^{ap} &= r_{0,t-1} + (\boldsymbol{\varepsilon}_t + \hat{\mathbf{v}}_{t-1})' \breve{\boldsymbol{\Sigma}}_{t-1}^{-1} \mathbf{v}_{t-1} = r_{0,t-1} + (\boldsymbol{\varepsilon}_t + \mathbf{v}_{t-1})' \breve{\boldsymbol{\Sigma}}_{t-1}^{-1} \mathbf{v}_{t-1} \\ &= r_{0,t-1} + \boldsymbol{\varepsilon}'_t \breve{\boldsymbol{\Sigma}}_{t-1}^{-1} \mathbf{v}_{t-1} + \mathbf{v}'_{t-1} \breve{\boldsymbol{\Sigma}}_{t-1}^{-1} \mathbf{v}_{t-1}. \end{aligned}$$

Then (81) follows by (121) and (122) and

$$\begin{aligned}\mathbf{v}'_{t-1}\left(\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}-\boldsymbol{\Sigma}_{t-1}^{-1}\right)\mathbf{v}_{t-1} &= \mathbf{v}'_{t-1}\mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1}+\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\boldsymbol{\Omega}_{t-1}^{-1}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}\\ &= O_p(N^{-1}), \qquad (125)\\ \boldsymbol{\varepsilon}'_t\left(\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}-\boldsymbol{\Sigma}_{t-1}^{-1}\right)\mathbf{v}_{t-1} &= \boldsymbol{\varepsilon}'_t\,\mathbf{H}_{t-1}\mathbf{B}(\boldsymbol{\Omega}_{t-1}^{-1}+\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\boldsymbol{\Omega}_{t-1}^{-1}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}\\ &= O_p(N^{-1}).\end{aligned}$$

which, in turn, follows by the second part of (72) and (117). Finally, regarding (82), by (125),

$$\begin{aligned}\mu_{\rho,t-1}^{ap}-r_{0,t-1} &= \mathbf{v}'_{t-1}\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{v}_{t-1}=\mathbf{v}'_{t-1}\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1}+O_p(N^{-1}),\\ \sigma_{\rho,t-1}^{ap\,2} &= \mathbf{v}'_{t-1}\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{G}_{t-1}\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{v}_{t-1}=\mathbf{v}'_{t-1}\breve{\boldsymbol{\Sigma}}_{t-1}^{-1}\mathbf{v}_{t-1}=\mathbf{v}'_{t-1}\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1}+O_p(N^{-1})\end{aligned}$$

where for the second equality on the left hand side of $\sigma_{\rho,t-1}^{ap\,2}$ we used (104). Then by (121) the result follows. □

Proof of Theorem 4 We follow the proofs of both part (i) and part (ii) of Theorem 3.
(i) Consider (107). As before $\mathbf{e}'_i\boldsymbol{\Sigma}_t^{-1}\mathbf{B}\ \boldsymbol{\lambda}_t=O_p(N^{-1})$ by (109) yielding $w_{it}^{mv}=\mathbf{e}'_i\boldsymbol{\Sigma}_t^{-1}\ \mathbf{v}_t+O_p(N^{-1})$. Then, in the unconstrained pricing case,

$$\begin{aligned}\mathbf{e}'_i\boldsymbol{\Sigma}_t^{-1}\ \mathbf{v}_t &= \mathbf{e}'_i\left[\mathbf{H}_t\ \mathbf{v}_t-\mathbf{H}_t\mathbf{B}(\boldsymbol{\Omega}_t^{-1}+\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t\right]\\ &= \left[\mathbf{e}'_i\mathbf{H}_t\ \mathbf{v}_t-\mathbf{e}'_i\mathbf{H}_t\mathbf{B}(\mathbf{A}_t+a_t\boldsymbol{\mu}_\beta\boldsymbol{\mu}'_\beta)^{-1}\boldsymbol{\mu}_\beta c_t\right](1+o_p(1)),\end{aligned}$$

where now

$$\frac{\mathbf{B}'\mathbf{H}_t\mathbf{v}_t}{N}\to_p\boldsymbol{\mu}_\beta c_t. \qquad (126)$$

Then, by identity (118) and re-arranging terms,

$$\begin{aligned}\mathbf{e}'_i\boldsymbol{\Sigma}_t^{-1}\ \mathbf{v}_t &= \left[\mathbf{e}'_i\mathbf{H}_t\mathbf{v}_t-(c_t/b_t)\mathbf{e}'_i\mathbf{H}_t\mathbf{B}\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta\right](1+o_p(1))\\ &= \breve{w}_{it}^{mv}(1+o_p(1))=O_p(1). \qquad (127)\end{aligned}$$

Concerning the `ap` strategy, by Theorem 1, $w_{it}^{ap}=\mathbf{e}'_i\breve{\boldsymbol{\Sigma}}_t^{-1}\boldsymbol{\mu}_t=\mathbf{e}'_i\breve{\boldsymbol{\Sigma}}_t^{-1}\ \mathbf{v}_t$ which behaves as (127) since

$$\mathbf{e}'_i\left(\breve{\boldsymbol{\Sigma}}_t^{-1}-\boldsymbol{\Sigma}_t^{-1}\right)\ \mathbf{v}_t=\mathbf{e}'_i\mathbf{H}_t\mathbf{B}(\boldsymbol{\Omega}_t^{-1}+\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\boldsymbol{\Omega}_t^{-1}(\mathbf{B}'\mathbf{H}_t\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_t\mathbf{v}_t=O_p(N^{-1}).$$

(ii) By (108) and (126) then $\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1}=O_p(1)$ since

$$\begin{aligned}\mathbf{v}'_{t-1}\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B} &= N^{-1}\mathbf{v}'_{t-1}\mathbf{H}_{t-1}\mathbf{B}(N^{-1}\boldsymbol{\Omega}_{t-1}^{-1}+N^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\boldsymbol{\Omega}_{t-1}^{-1}\\ &\to{}_p\, b_{t-1}^{-1}c_{t-1}\boldsymbol{\mu}'_\beta\mathbf{A}_{t-1}^{-1}\boldsymbol{\Omega}_{t-1}^{-1}, \qquad (128)\end{aligned}$$

We need to find the dominant term in (110)-(111)-(112). Since $\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t = O_p(N^{-\frac{1}{2}})$ by (114), $\boldsymbol{\lambda}_{t-1}'\mathbf{B}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{B}(\boldsymbol{\lambda}_{t-1}+\mathbf{z}_t) = O_p(1)$ by (113), we are left to evaluate $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t$ and $\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\ \mathbf{v}_{t-1}$. It turns out that these are $O_p(N^{\frac{1}{2}})$ and $O_p(N)$, respectively, and thus the latter is the dominant term. In fact, by (128),

$$\begin{aligned}
\operatorname{var}\left(\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t \mid \mathcal{A}_{(N),t-1}\right) &= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{G}_{t-1}\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} \\
&= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}(\boldsymbol{\Sigma}_{t-1}-\mathbf{B}\boldsymbol{\Omega}_{t-1}\mathbf{B}')\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} \\
&= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} + O_p(1)
\end{aligned}$$

yielding

$$\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\boldsymbol{\varepsilon}_t = O_p(N^{\frac{1}{2}}). \tag{129}$$

The last term satisfies

$$\frac{\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1}}{N} \to_p e_{t-1}/b_{t-1}. \tag{130}$$

In fact, by (103) and recalling (126),

$$\begin{aligned}
\frac{\mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1}}{N} &= \frac{\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}}{N} - \left(\frac{\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{B}}{N}\right)\left(\boldsymbol{\Omega}_{t-1}^{-1}+\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1} \\
&= \frac{\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}}{N} - \left(\frac{\mathbf{v}_{t-1}'\mathbf{H}_{t-1}\mathbf{B}}{N}\right)\left(\frac{\boldsymbol{\Omega}_{t-1}^{-1}}{N}+\frac{\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}}{N}\right)^{-1}\left(\frac{\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}}{N}\right) \\
&\to_p d_{t-1} - c_{t-1}^2\boldsymbol{\mu}_\beta'\left(\mathbf{A}_{t-1}+a_{t-1}\boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'\right)^{-1}\boldsymbol{\mu}_\beta,
\end{aligned}$$

from which (130) follows simple manipulations, in view of identity (118). Collecting terms, the limit of the normalized (by N^{-1}) `mv` portfolio excess return will be (130) establishing (86). It is easy to see that $e_t > 0$, almost surely. In fact it is given by the sum of d_t and $(a_t d_t - c_t^2)\boldsymbol{\mu}_\beta'\mathbf{A}_t^{-1}\boldsymbol{\mu}_\beta$. The latter term is positive since $a_t d_t > c_t^2$ by our assumption and $\mathbf{A}_t$ is positive definite. The first term d_t is certainly non-negative since it equals the probability limit of the quadratic form $N^{-1}\ \mathbf{v}_t'\mathbf{H}_t\mathbf{v}_t$.

Regarding (87), consider again (123) and (124). By (113), (115), (128) and (130)

$$\begin{aligned}
\mu_{\rho,t-1}^{mv} - r_{0,t-1} &= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} + O_p(1), \\
\sigma_{\rho,t-1}^{mv\,2} &= \mathbf{v}_{t-1}'\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{G}_{t-1}\boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} + O_p(1),
\end{aligned}$$

where for $\sigma_{\rho,t-1}^{mv\,2}$ we use

$$\mathbf{v}_{t-1}'\ \boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{G}_{t-1}\ \boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} = \mathbf{v}_{t-1}'\ \boldsymbol{\Sigma}_{t-1}^{-1}(\boldsymbol{\Sigma}_{t-1}-\mathbf{B}\boldsymbol{\Omega}_{t-1}\mathbf{B}')\ \boldsymbol{\Sigma}_{t-1}^{-1}\mathbf{v}_{t-1} = \ \mathbf{v}_{t-1}'\ \boldsymbol{\Sigma}_{t-1}^{-1}\ \mathbf{v}_{t-1}+O_p(1).$$

and

$$\mathbf{v}'_{t-1}\mathbf{\Sigma}^{-1}_{t-1}\mathbf{G}_{t-1}\ \mathbf{\Sigma}^{-1}_{t-1}\mathbf{B}\ \boldsymbol{\lambda}_{t-1} =\ \mathbf{v}'_{t-1}\ \mathbf{\Sigma}^{-1}_{t-1}\mathbf{B}(\mathbf{\Omega}^{-1}_{t-1}+\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{\Omega}^{-1}_{t-1}\ \boldsymbol{\lambda}_{t-1} = O_p(N^{-1}).$$

Therefore $N^{-1}(\mu^{mv}_{\rho,t-1} - r_{0,t-1})$ and $N^{-1}\sigma^{mv\,2}_{\rho,t-1}$ both converge in probability to e_{t-1}/b_{t-1} and (87) follows since the Sharpe ratio requires the square root of $\sigma^{mv\,2}_{\rho,t-1}$ in the denominator.

For the limit ap portfolio return, by repeated use of Theorem 1,

$$\rho^{ap}_t = r_{0,t-1} + \boldsymbol{\varepsilon}'_t\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1} + \mathbf{v}'_{t-1}\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1}.$$

Then (88) follows normalizing $\rho^{ap}_t - r_{0,t-1}$ by N^{-1} and applying (129) and (130) since

$$\begin{aligned}\mathbf{v}'_{t-1}\left(\breve{\mathbf{\Sigma}}^{-1}_{t-1} - \mathbf{\Sigma}^{-1}_{t-1}\right)\mathbf{v}_{t-1} &= \mathbf{v}'_{t-1}\mathbf{H}_{t-1}\mathbf{B}(\mathbf{\Omega}^{-1}_{t-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{\Omega}^{-1}_{t-1}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}\\ &= O_p(1), \qquad (131)\\ \boldsymbol{\varepsilon}'_t\left(\breve{\mathbf{\Sigma}}^{-1}_{t-1} - \mathbf{\Sigma}^{-1}_{t-1}\right)\mathbf{v}_{t-1} &= \boldsymbol{\varepsilon}'_t\mathbf{H}_{t-1}\mathbf{B}(\mathbf{\Omega}^{-1}_{t-1} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{\Omega}^{-1}_{t-1}(\mathbf{B}'\mathbf{H}_{t-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{H}_{t-1}\mathbf{v}_{t-1}\\ &= O_p(N^{-\frac{1}{2}}).\end{aligned}$$

which follows by the second part of (72) and (126). Finally, regarding (89), by (131),

$$\begin{aligned}\mu^{ap}_{\rho,t-1} - r_{0,t-1} &= \mathbf{v}'_{t-1}\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1} = \mathbf{v}'_{t-1}\mathbf{\Sigma}^{-1}_{t-1}\mathbf{v}_{t-1} + O_p(1),\\ \sigma^{ap\,2}_{\rho,t-1} &= \mathbf{v}'_{t-1}\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{G}_{t-1}\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1} = \mathbf{v}'_{t-1}\breve{\mathbf{\Sigma}}^{-1}_{t-1}\mathbf{v}_{t-1} = \mathbf{v}'_{t-1}\mathbf{\Sigma}^{-1}_{t-1}\mathbf{v}_{t-1} + O_p(1),\end{aligned}$$

where for the second equality on the left hand side of $\sigma^{ap\,2}_{\rho,t-1}$ we used (104). Therefore, by (130), $N^{-1}(\mu^{ap}_{\rho,t-1} - r_{0,t-1})$ and $N^{-1}\sigma^{ap\,2}_{\rho,t-1}$ both converge in probability to e_{t-1}/b_{t-1} and (89) follows. □

References

Aguilar, O., and M. West (2000): "Bayesian dynamic factor models and portfolio allocation," *Journal of Business and Economic Statistics*, 18, 338–357.

Billingsley, P. (1986): *Probability and measure.* New York: Wiley, second edn.

Brandt, M. (2004): "Portfolio choice problem," in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia,and L. Hansen. Amsterdam: Elsevier Science.

CAMPBELL, J., Y. CHAN, AND L. VICEIRA (2003): "A multivariate model of strategic asset allocation," *Journal of Financial Economics*, 67, 41–80.

CAMPBELL, J., AND L. VICEIRA (2001): "Who should buy long-term bonds?," *American Economic Review*, 91, 99–127.

CHAMBERLAIN, G. (1983): "Funds, factors and diversification in arbitrage pricing models," *Econometrica*, 51, 1305–1324.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1983): "Arbitrage, factor structure and mean-variance analysis on large asset markets," *Econometrica*, 51, 1281–1304.

CHAN, L., J. KARCESKI, AND J. LAKONISHOK (1999): "On portfolio optimization: forecasting covariances and choosing the risk model," *Review of Financial Studies*, 12, 937–974.

CHIB, S., F. NARDARI, AND N. SHEPHARD (2002): "Analysis of high dimensional multivariate stochastic volatility models," Washington University in Saint Louis, Preprint.

CHUDIK, A., M. PESARAN, AND E. TOSETTI (2009): "Weak and strong cross section dependence and estimation of large panels," Cambridge Working Papers in Economics 0924, Faculty of Economics, University of Cambridge.

COCHRANE, J., AND J. SAA-REQUEJO (2001): "Beyond arbitrage: good-deal asset price bounds in incomplete markets," *Journal of Political Economy*, 108, 79–119.

CONNOR, G. (1984): "A unified beta pricing theory," *Journal of Economic Theory*, 34, 13–31.

CONNOR, G., R. A. KORAJCZYK, AND O. LINTON (2006): "The common and specific components of dynamic volatility," *Journal of Econometrics*, 132, 231–255.

DEMIGUEL, V., L. GARLAPPI, AND R. UPPAL (2009): "Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy?," *Review of Financial Studies*, 22, 1915–1953.

DIEBOLD, F., AND M. NERLOVE (1989): "The dynamics of exchange rate volatility: a multivariate latent factor ARCH model," *Journal of Applied Econometrics*, 4, 1–21.

DOZ, C., AND E. RENAULT (2006): "Factor Stochastic Volatility in Mean Models: A GMM Approach," *Econometric Reviews*, 25, 275–309.

FAN, J., Y. FAN, AND J. LV (2007): "High-dimensional covariance matrix estimation using a factor model," *Journal of Econometrics*, 147, 186–197.

FIORENTINI, G., E. SENTANA, AND N. SHEPHARD (2004): "Likelihood-based estimation of latent generalized ARCH structures," *Econometrica*, 72, 1481–1517.

FLEMING, J., C. KIRBY, AND B. OSTDIEK (2001): "The economic value of volatility timing," *Journal of Finance*, 56, 329–351.

GREEN, R., AND B. HOLLIFIELD (1992): "When will mean-variance efficient portfolios be well diversified?," *Journal of Finance*, 47, 1785–1809.

GRINBLATT, M., AND S. TITMAN (1987): "The relation between mean-variance efficiency and arbitrage pricing," *Journal of Business*, 60, 97–112.

HAN, Y. (2006): "Asset allocation with a high-dimensional latent factor stochastic volatility model," *The Review of Financial Studies*, 19, 237–271.

HANSEN, L., AND S. RICHARD (1987): "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," *Econometrica*, 55, 587–613.

HUBERMANN, G. (1982): "A simple approach to the arbitrage pricing theory," *Journal of Economic Theory*, 28, 183–191.

INGERSOLL, J. (1984): "Some results in the theory of arbitrage pricing," *Journal of Finance*, 39, 1021–1039.

JAGANNATHAN, R., AND T. MA (2003): "Risk reduction in large portfolios: why imposing the wrong constraints helps," *Journal of Finance*, 58, 1651–1684.

JOHANNES, M., N. POLSON, AND J. STROUD (2002): "Sequential optimal portfolio performance: market and volatility timing," Graduate School of Business, University of Chicago, Preprint.

KANDEL, S., AND R. STAMBAUGH (1996): "On the predictability of stock returns: an asset allocation perspective," *Journal of Finance*, 51, 385–424.

King, M., E. Sentana, and S. Wadhwani (1994): "Volatility and links between national stock markets," *Econometrica*, 62, 901–933.

Lintner, J. (1965): "The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets," *Review of Economics and Stastistics*, 47, 13–37.

Lutkepohl, H. (1996): *Handbook of Matrices*. John Wiley & Sons, New York.

MacKinley, A. (1995): "Multifactor models do not explain deviations from the CAPM," *Journal of Financial Economics*, 38, 3–28.

Magnus, J., and H. Neudecker (2001): *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Chichester (UK): John Wiley & Sons Ltd, second edn.

Markowitz, H. (1952): "Portfolio selection," *Journal of Finance*, 7, 77–91.

Merton, R. (1973): "An intertemporal capital asset pricing model," *Econometrica*, 41, 867–887.

Pesaran, M., and A. Timmermann (1995): "Predictability of stock returns: robustness and economic significance," *Journal of Finance*, 50, 1201–1228.

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

Sentana, E. (2004): "Factor representing portfolios in large asset markets," *Journal of Econometrics*, 119, 257–289.

Sharpe, W. (1964): "Capital asset prices: a theory of market equilibrium under conditions of risk," *Journal of Finance*, 1919, 425–442.

Ullah, A. (2004): *Finite Sample Econometrics*. Oxford: Oxford University Press.