Large Sample Estimators of the Stochastic Discount Factor*

Soohun Kim† Robert A. Korajczyk‡

April 16, 2018

Abstract

We propose several large sample estimators of the stochastic discount factor (SDF) for pricing risky assets. Our estimators can utilize not only a set of factors implied by a specific asset pricing model but also a set of latent factors estimated by multivariate statistical methods. We suggest a correction for the bias induced by having a finite time series and show how to use the correction in exploiting unbalanced panels of individual stock returns. The estimators perform well in simulations designed to mimic the the U.S. equity markets. A Lasso penalized version of the estimators does a good job of excluding systematic, but unpriced factors. When applied to large cross sections of equity returns, the estimators provide evidence about which factors command a risk premium.

JEL classification codes: G12

Keywords: asset pricing, factor structure, stochastic discount factor

*We thank Christian Brownlees, Gregory Connor, Ravi Jagannathan, Clark Liu, Andreas Nuehierl, Kuntara Pukthuanthong, Richard Roll, Olivier Scaillet, and participants at the Conference on Asia-Pacific Financial Markets and ERFIN workshop at the Warsaw School of Economics for helpful comments.

†Scheller College of Business, Georgia Institute of Technology, Atlanta, GA 30308. E-mail: soohun.kim@scheller.gatech.edu.

‡Kellogg School of Management, Northwestern University, Evanston, IL 60208. E-mail: r-korajczyk@kellogg.northwestern.edu.

1 Introduction

In an economy without arbitrage opportunities, there exists a valid stochastic discount factor (SDF) such that the price of any security is obtained as the expected value of the discounted (by the SDF) future payoff. The stochastic discount factor representation of asset pricing models has been widely used in the empirical finance literature. It is often the case that empirical studies using asset returns to estimate the SDF use a small number of portfolios.[1] This paper proposes several alternative estimators of the stochastic discount factor so that empirical researchers can fully exploit useful information on asset prices from a large panel of either balanced or unbalanced financial data.

The intuition behind our estimators is similar to the idea in Hansen and Jagannathan (1997). Within a set of candidate SDFs, we search for the one which minimizes the norm of pricing errors. It turns out that we can consistently estimate the true SDF with the mild assumption that individual assets have an approximate factor structure as in Chamberlain and Rothschild (1983). In an economy where the returns of a large number assets are driven by a finite number of common factors, the true SDF is expressed as a linear function of the pervasive factors which can be either traded or non-traded. Our estimator minimizes the (cross-sectional) sum of squared SDF pricing errors across a large number of assets. For the case of balanced panel data, the minimization problem is reduced to a linear regression problem and the solution can be easily obtained. To accommodate unbalanced panel data, we split the time series into multiple non-overlapping time blocks and mimic the solution of the balanced panel case with each short time block data. With a proper small-sample (in the time dimension) bias correction similar to Litzenberger and Ramaswamy (1979) and Kim and Skoulakis (2017), our estimators can consistently estimate the true SDF. Additionally, we find that our estimators are robust to the latency of factors, being partially agnostic in the spirit of Pukthuanthong and Roll (2017), and can adopt regularization tools such as lasso regressions as in Kozak, Nagel, and Santosh (2017).

This paper is not the first attempt to exploit the SDF representations of a large number of assets. Araujo and Issler (2012) show that the SDF can be summarized by a scaled inverse of the cross-sectional geometric average of returns. We allow multiple pervasive risks (possibly priced and non-priced factors) in an economy and let our

[1]To the best of our knowledge, Cochrane (1996) and Jagannathan and Wang (1996) are first papers to propose SDF specification for asset pricing tests using a long time-series data of a small number of assets.

estimators find the priced factors among those by observing price dynamics of the large panel. In particular, our paper is closely related to Pukthuanthong and Roll (2017). They also propose an SDF estimator which minimizes the squared pricing errors. Their approach has an advantage of being agnostic in the sense that no assumptions are required for the return generating process or the nature of systematic risk, except the SDF representation. However, in simulations, we find that the downside of the lack of structure of this SDF estimator is that it provides a very noisy estimate of the true SDF in economies constructed to have risk matching that of the U.S. equity market. We argue that being slightly less agnostic by imposing a more restrictive factor structure on the SDF, as in our estimators, leads to significant improvement in the performance of the estimated SDF with empirically relevant panel sizes. An alternative approach is proposed by Kozak, Nagel, and Santosh (2017) which estimates the SDF using a vast array of characteristics on individual stocks.

We also contribute to a broader literature of using individual stocks for the empirical studies of asset pricing models. The arbitrage pricing theory of Ross (1976) and Chamberlain and Rothschild (1983) provides a framework to dichotomize a large cross section of returns into pervasive factors and diversifiable risks. A long literature derives methods to extract pervasive factors from a large cross sectional data (e.g., Connor and Korajczyk (1986, 1987, 1988), Stroyny (1992), Stock and Watson (1998, 2002), and Jones (2001)). We contribute to this literature by providing a simple tool to select a priced factor among the pervasive common factors extracted from a large panel of data. As pointed out in Merton (1973), Jagannathan and Wang (1996), Campbell and Vuolteenaho (2004), Kelly and Pruitt (2013) and Jagannathan and Marakani (2015), not all pervasive factors (i.e., those that explain common movements in asset returns) need be important for explaining the cross section of asset prices.

Alternatively, the pricing of a given pervasive factor can be examined with the beta pricing form. Exploiting a large cross section of assets in estimating beta pricing models, a series of papers have proposed risk premia estimators using large cross-sections (see Litzenberger and Ramaswamy (1979), Shanken (1992) and Jagannathan, Skoulakis, and Wang (2010)). The recent papers by Gagliardini, Ossola, and Scaillet (2016) and Kim and Skoulakis (2017, 2018) obtain the large panel asymptotic distribution of the risk premia estimator along with an estimator of its variance-covariance matrix. Our paper is differentiated in that we are using the SDF representation, not a beta pricing representation. This difference is particularly important when we use large panel data where the measurement errors in individual asset betas can severely bias estimated risk premia. Although the equivalence between SDF form and beta form is well known in the small-N/large-T setup (Jagannathan and Wang (2002)), more work is required to understand the differences between two approaches in the large panel data setting.

In Section 2, we describe our large cross-sectional economy and propose several large sample estimators of the stochastic discount factor (SDF) for pricing risky assets. In Section 3, we simulate an economy in which asset risks match those in the U.S. equity markets and examine the performance of our SDF estimators across various sample sizes. The estimators perform well and the imposition of a factor structure improves the estimators' performance relative to purely agnostic alternatives. The bias correction for unbalanced panels works. A Lasso version of the estimator is able to exclude unpriced factors. In Section 4, we apply our SDF estimators to a large cross section of assets and provide evidence on the size and significance of risk premia on candidate factors. We find that the value factor, HML, from Fama and French (1993, 2015) is subsumed by investment and profitability factors in for some sets of test assets. Section 5 concludes. All proofs are in the Appendix.

2 Economy

We assume that the gross return generating process of each individual security follows a K-factor model. In particular, the gross return of the i-th asset at time t is expressed as

$$R_{i,t} = \alpha_i + \boldsymbol{\beta}_i' \mathbf{f}_t + e_{i,t}, \text{ for } i = 1, \cdots, N \text{ and } t = 1, \cdots, T, \tag{2.1}$$

where $\boldsymbol{\beta}_i$ is the $(K \times 1)$ vector of factor loadings of the i-th asset on the $(K \times 1)$ vector of factor realizations, $\mathbf{f}_t$. As is standard, we assume $\mathbb{E}[e_{i,t}] = 0$ and $\mathbb{E}[\mathbf{f}_t e_{i,t}] = \mathbf{0}_K$, a $(K \times 1)$ vector of zeros. We allow the factor of $\mathbf{f}_t$ to be either traded excess returns, traded gross returns, latent, or nontraded factors.

With some mild assumptions on the cross-sectional dependency among residuals of $e_{i,t}$, Ross (1976) and Chamberlain and Rothschild (1983) show that in an economy without statistical arbitrage, there exists a scalar, λ_0, the gross return on the riskless asset, and a $(K \times 1)$ vector, $\boldsymbol{\lambda}_f$, such that

$$\mathbb{E}[R_{i,t}] \approx \lambda_0 + \boldsymbol{\beta}_i' \boldsymbol{\lambda}_f. \tag{2.2}$$

We assume that exact factor pricing holds, so that the equation (2.2) holds as an equality (as in Connor (1984)). Let the $(K \times 1)$ vector $\boldsymbol{\mu}_f$ be $\boldsymbol{\mu}_f = \mathbb{E}[\mathbf{f}_t]$. By combining the return generating process of (2.1) and the exact form of the pricing restriction of (2.2), we have

$$\mathbb{E}\left[R_{i,t}\right] = \alpha_i + \boldsymbol{\beta}_i'\boldsymbol{\mu}_f = \lambda_0 + \boldsymbol{\beta}_i'\boldsymbol{\lambda}_f,$$

implying that

$$\alpha_i = \lambda_0 + \boldsymbol{\beta}_i'\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f\right).$$

Then, plugging the above expression into the process of (2.1) yields

$$R_{i,t} = \lambda_0 + \boldsymbol{\beta}_i'\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f + \mathbf{f}_t\right) + e_{i,t}. \tag{2.3}$$

The equation (2.3) allows for many different specifications of the nature of the factor vector, $\mathbf{f}_t$. If $\mathbf{f}_t$ is an observed vector of portfolio excess returns (as in Fama and French (1993)) then $\boldsymbol{\mu}_{\boldsymbol{f}} = \boldsymbol{\lambda}_{\boldsymbol{f}}$. If $\mathbf{f}_t$ is an observed vector of portfolio gross returns, then $\boldsymbol{\mu}_f = \mathbf{1}_K\lambda_0 + \boldsymbol{\lambda}_f$ and spanning of the mean-variance frontier by the factors implies that (2.3) reduces to (see Huberman and Kandel (1987)):

$$R_{i,t} = \boldsymbol{\beta}_i'\mathbf{f}_t + e_{i,t}, \tag{2.4}$$

with the added constraint that $\boldsymbol{\beta}_i'\mathbf{1}_K = 1$. If $\mathbf{f}_t$ is an observed vector of pre-whitened macroeconomic variables, then $\boldsymbol{\mu}_f = \mathbf{0}_K$. In the literature, there are a number of papers which use a combination of traded excess returns and pre-whitened macroeconomic variables, such as Chen, Roll, and Ross (1986) or Shanken and Weinstein (2006). In this case the expected value of the factors is the factor risk premium for the excess return factors and zero for the pre-whitened variables. Finally, if $\mathbf{f}_t$ is an unobserved vector of latent portfolio excess returns (as in Connor and Korajczyk (1986)) then $\boldsymbol{\mu}_f = \boldsymbol{\lambda}_f$, but the procedure requires an consistent estimator of the excess returns on factor mimicking portfolios.

Next, we specify the stochastic discount factor (SDF) m_t in this economy such that

$$\mathbb{E}\left[R_{i,t}m_t\right] = 1 \text{ for } i = 1, \cdots, N.$$

The realized SDF is a linear function of the realization of the systematic factors:

$$m_t = \delta_0 + \mathbf{f}_t'\boldsymbol{\delta}_f, \tag{2.5}$$

which satisfies $\mathbb{E}\left[R_{i,t}m_t\right] = 1$ when the scalar δ_0 and the $(K \times 1)$ vector, $\boldsymbol{\delta}_f$, are given by

$$\delta_0 = \frac{1}{\lambda_0}\left(1 + \boldsymbol{\mu}_f' \Sigma_f^{-1} \boldsymbol{\lambda_f}\right) \tag{2.6}$$

$$\boldsymbol{\delta}_f = -\frac{1}{\lambda_0}\left(\Sigma_f^{-1} \boldsymbol{\lambda_f}\right), \tag{2.7}$$

where

$$\Sigma_f = \mathbb{E}\left[\left(\mathbf{f}_t - \boldsymbol{\mu}_f\right)\left(\mathbf{f}_t - \boldsymbol{\mu}_f\right)'\right].$$

The expected value of m_t is λ_0^{-1}.

So far, we describe an economy with N assets and specify the form of stochastic discount factor as a linear function of systematic factors, which prices the gross returns of the N assets. In many cases, when a risk free asset exists, empirical research studies the returns of the N assets in excess of the risk free return. If there exists a risk free asset, then the expression of (2.3) implies that the gross return of the risk free asset is λ_0 since it has neither any exposure to the factor ($\boldsymbol{\beta}_i = \mathbf{0}_K$) nor residual risk ($e_{i,t} = 0$). Hence, from (2.3), the excess return of the i-th asset at time t can be written as

$$R_{i,t}^e = R_{i,t} - \lambda_0 = \boldsymbol{\beta}_i'\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f + \mathbf{f}_t\right) + e_{i,t}. \tag{2.8}$$

Now, we characterize a stochastic discount factor m_t^e which prices the excess returns of the N assets, i.e.,

$$\mathbb{E}\left[R_{i,t}^e m_t^e\right] = 0 \text{ for } i = 1, \cdots, N.$$

It can be shown that we can construct a stochastic discount factor

$$m_t^e = 1 + \mathbf{f}_t' \boldsymbol{\delta}^e, \tag{2.9}$$

satisfying $\mathbb{E}\left[R_{i,t}^e m_t^e\right] = 0$, with the $(K \times 1)$ vector of $\boldsymbol{\delta}^e$, given by

$$\boldsymbol{\delta}^e = -\left(\boldsymbol{\Sigma}_f + \boldsymbol{\lambda}_f \boldsymbol{\mu}_f'\right)^{-1} \boldsymbol{\lambda}_f. \tag{2.10}$$

We obtain an extra degree of freedom when pricing excess returns, rather than gross returns, since we do not require the SDF to pin down the mean of m or, equivalently, the riskless rate of return, thus it can be off in pricing gross returns by a constant which cancels when excess returns are analyzed (see Cochrane (2005, section 6.3)).

Neither of the stochastic discount factors, m_t (for gross returns) nor m_t^e (for excess returns), are observable since the parameters, $\boldsymbol{\lambda}_f$, $\boldsymbol{\mu}_f$,$\boldsymbol{\Sigma}_f$, λ_0, and possibly the factors themselves, are unobservable. We propose several alternative estimators of the SDFs which are based on using large cross-sections of individual assets or portfolios. We start with an estimator assuming a balanced panel of asset returns and then extend to a number of alternative estimators.

2.1 Balanced Panel Estimator

In this section, we assume that we observe the gross returns of $R_{i,t}$ or the excess returns of $R_{i,t}^e$ for assets $i = 1, \cdots, N$ over the time period $t = 1, \cdots, T$. It is convenient to represent the gross return generating process of (2.3) and the excess return generating process of (2.8) in matrix form:

$$\mathbf{R} = \lambda_0 \mathbf{1}_N \mathbf{1}_T' + \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E}, \tag{2.11}$$

and

$$\mathbf{R}^e = \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E}, \tag{2.12}$$

where the (i,t) element of the $(N \times T)$ matrices of $\mathbf{R}$ and $\mathbf{R}^e$ are $R_{i,t}$ and $R_{i,t}^e$, respectively, $\mathbf{1}_N$ is the $(N \times 1)$ vector of ones, the i-th row of the $(N \times K)$ matrix of $\mathbf{B}$ is $\boldsymbol{\beta}_i'$, the t-th row of the $(T \times K)$ matrix of $\mathbf{F}$ is $\mathbf{f}_t'$, and the (i,t) element of the $(N \times T)$ matrix of $\mathbf{E}$ is $e_{i,t}$.

We make standard assumptions on the systematic factors and factor loadings.

Assumption 1. *As $N \to \infty$, $\frac{1}{N}\mathbf{B}'\mathbf{1}_N \to \boldsymbol{\mu}_\beta$ and $\frac{1}{N}\mathbf{B}'\mathbf{B} \to \mathbf{V}_\beta = \Sigma_\beta + \boldsymbol{\mu}_\beta\boldsymbol{\mu}_\beta'$, where Σ_β is a positive definite matrix. Also, as $T \to \infty$, $\frac{1}{T}\mathbf{F}'\mathbf{1}_T \xrightarrow{p} \boldsymbol{\mu}_f$ and $\frac{1}{T}\mathbf{F}'\mathbf{F} \xrightarrow{p} \mathbf{V}_f = \Sigma_f + \boldsymbol{\mu}_f\boldsymbol{\mu}_f'$, where Σ_f is a positive definite matrix.*

Assumption 1 specifies that loadings on each factor are pervasive across a large number of assets and that each factor is neither redundant nor non-stationary over time, which are reasonably acceptable for the return generating process. The assumption does not imply that all pervasive factors are priced, so that it allows factors that explain common variation but are not deemed important by investors.

Next, we make assumptions on the distributional properties of the residual terms of $e_{i,t}$. We use $\mathbf{0}_{m\times n}$ to denote the $(m \times n)$ matrix of zeros.

Assumption 2. *As $N, T \to \infty$, $\frac{\mathbf{1}_N'\mathbf{E}\mathbf{1}_T}{NT} \xrightarrow{p} 0$, $\frac{\mathbf{F}'\mathbf{E}'\mathbf{1}_N}{NT}, \frac{\mathbf{B}'\mathbf{E}\mathbf{1}_T}{NT} \xrightarrow{p} \mathbf{0}_K$ and $\frac{\mathbf{B}'\mathbf{E}\mathbf{F}}{NT} \xrightarrow{p} \mathbf{0}_{K\times K}$. Also, as $N, T \to \infty$, $\frac{1}{N}\left(\frac{\mathbf{E}\mathbf{1}_T}{T}\right)'\left(\frac{\mathbf{E}\mathbf{1}_T}{T}\right) \xrightarrow{p} 0$, $\frac{1}{N}\left(\frac{\mathbf{E}\mathbf{F}}{T}\right)'\left(\frac{\mathbf{E}\mathbf{1}_T}{T}\right) \xrightarrow{p} \mathbf{0}_K$, and $\frac{1}{N}\left(\frac{\mathbf{E}\mathbf{F}}{T}\right)'\left(\frac{\mathbf{E}\mathbf{F}}{T}\right) \xrightarrow{p} \mathbf{0}_{K\times K}$.*

The first set of conditions in Assumption 2 states that the average residual terms over the $(N \times T)$ panel data converges to zero even when the average is weighted by factor realizations (in time-series dimension) or factor loadings (in cross-sectional dimension). The second set of conditions in Assumption 2 imposes that the time-series averages of residuals and the product of residuals and factors are sufficiently close to zero so that the squared cross-sectional averages converge to zeros.

The following theorem establishes that we can recover the stochastic discount factor as a linear function of factors from a large panel data.

Theorem 2.1. *With Assumptions 1 and 2, as $N, T \to \infty$, $\widetilde{m}_t = \widetilde{\delta}_0 + \mathbf{f}_t' \widetilde{\boldsymbol{\delta}}_f$ and $\widetilde{m}_t^e = 1 + \mathbf{f}_t' \widetilde{\boldsymbol{\delta}}^e$ converge to m_t and m_t^e given in (2.5) and (2.9), respectively, when the $((K+1) \times 1)$ vector of $\widetilde{\boldsymbol{\delta}} = \left[\widetilde{\delta}_0 \ \widetilde{\boldsymbol{\delta}}_f'\right]'$ and the $(K \times 1)$ vector of $\widetilde{\boldsymbol{\delta}}^e$ are constructed by*

$$\widetilde{\boldsymbol{\delta}} = \left(\frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{R} \mathbf{F}_\triangle}{NT^2}\right)^{-1} \left(\frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{1}_N}{NT}\right) \tag{2.13}$$

$$\widetilde{\boldsymbol{\delta}}^e = -\left(\frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{F}}{NT^2}\right)^{-1} \left(\frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{1}_T}{NT^2}\right), \tag{2.14}$$

where $\mathbf{F}_\triangle = [\mathbf{1}_T \ \mathbf{F}]$.

The estimator proposed in Theorem 2.1 can be intuitively understood as follows. By specifying the $(T \times 1)$ vector of the realized SDF, $[m_1 \cdots m_T]'$, as $\mathbf{m} = \mathbf{F}_\triangle \boldsymbol{\delta}$, the realized mispricing of the N assets' gross returns can be formulated by

$$\mathbf{1}_N - \frac{\mathbf{R}\mathbf{m}}{T} = \mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_\triangle}{T}\boldsymbol{\delta},$$

and the estimator $\widetilde{\boldsymbol{\delta}}$ in (2.13) can be obtained as the solution of the minimizing the squared pricing error:

$$\widetilde{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}} \left(\mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_\triangle}{T}\boldsymbol{\delta}\right)' \left(\mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_\triangle}{T}\boldsymbol{\delta}\right). \tag{2.15}$$

Similarly, given the $(T \times 1)$ vector of the realized SDF, $[m_1^e \cdots m_T^e]'$, denoted by $\mathbf{m}^e = \mathbf{1}_T + \mathbf{F}\boldsymbol{\delta}^e$, the estimator $\widetilde{\boldsymbol{\delta}}^e$ in (2.14) can be interpreted as the solution of the minimizing the squared pricing error:

$$\widetilde{\boldsymbol{\delta}}^e = \arg\min_{\boldsymbol{\delta}^e} \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T}\boldsymbol{\delta}^e\right)' \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T}\boldsymbol{\delta}^e\right). \tag{2.16}$$

The formation of $\widetilde{\boldsymbol{\delta}}$ and $\widetilde{\boldsymbol{\delta}}^e$ as in (2.15) and (2.16) implies that they are regression coefficients (regressing $\mathbf{1}_N$ on $\frac{\mathbf{RF}_\triangle}{T}$ for gross returns and regressing $\frac{\mathbf{R}^e\mathbf{1}_N}{T}$ on $\frac{\mathbf{R}^e\mathbf{F}}{T}$ for excess returns). This facilitates the adoption of popular regularization/selection tools such as Ridge regression or Lasso. These tools are extremely useful in that we can generalize the asset pricing model by incorporating a relatively large number of factors and let the statistical tools select factors which are useful in minimizing pricing errors. For example, we can apply the estimator to K pervasive factors to choose $L < K$ priced factors:

$$\widetilde{\boldsymbol{\delta}}_{lasso} = \arg\min_{\delta_0,\mathbf{d}} \left(\mathbf{1}_N - \frac{\mathbf{RF}_\triangle}{T}\boldsymbol{\delta}\right)' \left(\mathbf{1}_N - \frac{\mathbf{RF}_\triangle}{T}\boldsymbol{\delta}\right) + \gamma\|\mathbf{d}\|_1, \tag{2.17}$$

where $\gamma > 0$, $\mathbf{d} = [d_1 \cdots d_K]'$, $\|\mathbf{d}\|_1 = |d_1| + \cdots + |d_K|$, $\mathcal{F}$ is the $(T \times K)$ matrix such that $\mathbf{F} = \mathcal{F}\mathcal{S}_f$, $\boldsymbol{\delta}_f = \mathcal{S}_f'\mathbf{d}$, $\boldsymbol{\delta} = [\delta_0 \ \boldsymbol{\delta}_f]'$, and the $(K \times L)$ factor selection matrix of $\mathcal{S}_f$ is defined as follows. The (k, l) element of $\mathcal{S}_f$ is 1 if d_k is the l-th non zero element among $d_1, \cdots, d_k$ and 0 otherwise. Similarly, the extension of $\widetilde{\boldsymbol{\delta}}^e$ with the lasso penalty and the extended factor $\mathcal{F}$ can be formulated by

$$\widetilde{\boldsymbol{\delta}}^e_{lasso} = \arg\min_{\mathbf{d}} \left(\frac{\mathbf{R}^e\mathbf{1}_N}{T} - \frac{\mathbf{R}^e\mathbf{F}}{T}\boldsymbol{\delta}^e\right)' \left(\frac{\mathbf{R}^e\mathbf{1}_N}{T} - \frac{\mathbf{R}^e\mathbf{F}}{T}\boldsymbol{\delta}^e\right) + \gamma\|\mathbf{d}\|_1, \tag{2.18}$$

where $\boldsymbol{\delta}^e = \mathcal{S}_f'\mathbf{d}$.

Theorem 2.1 assumes that we have the true, but possibly mean-deficient, factors. An alternative approach is to treat $\mathbf{F}$ as latent factors which are estimated through multivariate statistical techniques. For this case we do not directly observe factor realizations, $\mathbf{F}$, but estimate those with $\mathbf{F}^*$ such that $\mathbf{F}^* = \mathbf{F}\mathcal{O} + o_p(1)$, where $\mathcal{O}$ is a $K \times K$ linear transformation. The following corollary shows that the consistent estimation of the stochastic discount factor is still feasible. In practice, we obtain estimates of the latent factors by applying principal components analysis (PCA) to large cross sectional data as in Connor and Korajczyk (1986) or Stock and Watson (1998).

Corollary 2.1. *With Assumptions 1 and 2, given a consistent factor estimator of* $\mathbf{F}^* = \mathbf{F}\mathcal{O} + o_p(1)$ *for a some rotation matrix of* $\mathcal{O}$*, as* $N, T \to \infty$*,* $\widetilde{m}_t^* = \widetilde{\delta}_0^* + \mathbf{f}_t^{*\prime}\widetilde{\boldsymbol{\delta}}_f^*$ *and* $\widetilde{m}_t^{*e} = 1 + \mathbf{f}_t^{*\prime}\widetilde{\boldsymbol{\delta}}^{*e}$ *converge to* m_t *and* m_t^e *given in* (2.5) *and* (2.9)*, respectively, when the*

$((K+1)\times 1)$ *vector of* $\widetilde{\boldsymbol{\delta}}^{*} = \left[\widetilde{\delta}_0^{*}\ \widetilde{\boldsymbol{\delta}}_f^{*\prime}\right]'$ *and the* $(K\times 1)$ *vector of* $\widetilde{\boldsymbol{\delta}}^{*e}$ *are constructed by*

$$\widetilde{\boldsymbol{\delta}}^{*} = \left(\frac{\mathbf{F}_{\triangle}^{*\prime}\mathbf{R}'\mathbf{R}\mathbf{F}_{\triangle}^{*}}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle}^{*\prime}\mathbf{R}'\mathbf{1}_N}{NT}\right) \tag{2.19}$$

$$\widetilde{\boldsymbol{\delta}}^{*e} = -\left(\frac{\mathbf{F}^{*\prime}\mathbf{R}^{e\prime}\mathbf{R}^{e}\mathbf{F}^{*}}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}\mathbf{R}^{e\prime}\mathbf{R}^{e}\mathbf{1}_T}{NT^2}\right), \tag{2.20}$$

where $\mathbf{F}_{\triangle}^{*} = [\mathbf{1}_T\ \mathbf{F}^{*}]$ *and* $\mathbf{F}^{*} = [\mathbf{f}_1^{*}\ \cdots\ \mathbf{f}_T^{*}]'$.

Furthermore, even without factor estimates, we can consistently estimate the stochastic discount factor for the gross returns with some restrictions on the residual variances and the sequential asymptotics of $N\to\infty$ and then $T\to\infty$. It is worth to emphasize that the SDF estimator of this case is identical to that by Pukthuathong and Roll (2017).

Proposition 2.1. *With Assumptions 1 and 2, the homoskedasticity condition of* $\frac{1}{N}\mathbf{E}'\mathbf{E}\xrightarrow{p} s\mathbf{I}_T$, *as* $N\to\infty$ *and then* $T\to\infty$, $\breve{m}_t$ *converges to* m_t *in* (2.5), *when* $\breve{m}_t$ *is given by*

$$\breve{m}_t = \iota_t'\left(\frac{\mathbf{R}'\mathbf{R}}{NT^2}\right)^{-1}\left(\frac{\mathbf{R}'\mathbf{1}_N}{NT}\right), \tag{2.21}$$

where ι_t *is the* $(T\times 1)$ *vector of zeros except the* t*-th element of one.*

The Pukthuathong and Roll (2017) estimator is totally agnostic as to the number of pervasive factors. Our estimator is equivalent to that of Pukthuathong and Roll (2017) when we let $K = T$.[2]

2.2 Unbalanced Panel Estimator

Since the estimators proposed in the previous subsection utilize the full panel data of large N and T, they are appropriate for the case using a large number of portfolios over a long horizon. However, if empirical researchers wish to use individual stocks to construct the SDF, the estimators are problematic due to the survivorship biases induced by requiring a balanced panel of individual assets. In this section, we propose estimators

[2] Let $\widetilde{\mathbf{m}} = [\widetilde{m}_1\ \cdots\ \widetilde{m}_T]$ and $\breve{\mathbf{m}} = [\breve{m}_1\ \cdots\ \breve{m}_T]$. Then, when $K=T$, it holds that

$$\widetilde{\mathbf{m}} = \mathbf{F}_{\triangle}\widetilde{\boldsymbol{\delta}} = \mathbf{F}_{\triangle}\left(\frac{\mathbf{F}_{\triangle}'\mathbf{R}'\mathbf{R}\mathbf{F}_{\triangle}}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle}'\mathbf{R}'\mathbf{1}_N}{NT}\right) = \mathbf{F}_{\triangle}\mathbf{F}_{\triangle}^{-1}\left(\frac{\mathbf{R}'\mathbf{R}}{NT^2}\right)^{-1}(\mathbf{F}_{\triangle}')^{-1}\mathbf{F}_{\triangle}'\left(\frac{\mathbf{R}'\mathbf{1}_N}{NT}\right) = \breve{\mathbf{m}}.$$

which deal with unbalanced panel data by estimating the SDF over non-overlapping time periods of length τ with the number of blocks increasing as T approaches infinity. The estimators proposed above (and the estimator of Pukthuathong and Roll (2017)) are biased for finite values of T since the errors in the regression of $\mathbf{1}_N$ (for gross returns) and $\frac{\mathbf{R}^e \mathbf{1}_N}{T}$ (for excess returns) are correlated with the regressor. This requires us to propose a bias correction for the SDF estimators.

We split the time period of length T into B non-overlapping time blocks of length τ such that $T = B\tau$. We fix τ. Hence, as T increases, B increases. We use $b = 1, \cdots, B$ as an index of time blocks. For example, the first block of $b = 1$ covers the time period $t = 1, \cdots, \tau$ and the second block of $b = 2$ covers the time period $t = \tau + 1, \cdots, 2\tau$. Pick a specific b. Then, collect all individual stocks with full return data over the b-th time block, $t = (b-1)\tau + 1, \cdots, b\tau$. Although this restriction can be relaxed by assuming missing-at-random within a block (as in Connor and Korajczyk (1987) and Stock and Watson (1998)), we require full returns in a single time block for simplicity. We relabel stocks in block b with the index of $i_{[b]} = 1, \cdots, N_{[b]}$, where $N_{[b]}$ is the number of individual stocks with full returns over the b-th time block. Note that $i_{[b]}$ does not have to be identical to the original index of i and that $i_{[b]}$ will, in general, be different from $i_{[b']}$ when $b \neq b'$.

Next, we express the observed return generating process in the b-th time block similarly to the original full panel representation of (2.11) and (2.12):

$$\mathbf{R}_{[b]} = \lambda_0 \mathbf{1}_{N_{[b]}} \mathbf{1}'_\tau + \mathbf{B}_{[b]} \left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f \right) \mathbf{1}'_\tau + \mathbf{B}_{[b]} \mathbf{F}'_{[b]} + \mathbf{E}_{[b]}, \tag{2.22}$$

and2

$$\mathbf{R}^e_{[b]} = \mathbf{B}_{[b]} \left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f \right) \mathbf{1}'_\tau + \mathbf{B}_{[b]} \mathbf{F}'_{[b]} + \mathbf{E}_{[b]}, \tag{2.23}$$

where the $\left(i_{[b]}, s\right)$ element of the $\left(N_{[b]} \times \tau\right)$ matrices of $\mathbf{R}_{[b]}$ and $\mathbf{R}^e_{[b]}$ are $R_{i_{[b]},(b-1)\tau+s}$ and $R^e_{i_{[b]},(b-1)\tau+s}$, respectively, $\mathbf{1}_m$ is the $(m \times 1)$ vector of ones, the $i_{[b]}$-th row of the $\left(N_{[b]} \times K\right)$ matrix of $\mathbf{B}_{[b]}$ is $\boldsymbol{\beta}'_{i_{[b]}}$, the s-th row of the $(\tau \times K)$ matrix of $\mathbf{F}_{[b]}$ is $\mathbf{f}'_{(b-1)\tau+s}$, and the $\left(i_{[b]}, s\right)$ element of the $\left(N_{[b]} \times \tau\right)$ matrix of $\mathbf{E}_{[b]}$ is $e_{i_{[b]},(b-1)\tau+s}$.

We need the assumptions of the availability of large cross-sectional data in each time block and the time-invariant first two moments in the cross-sectional distribution of factor loadings.

Assumption 3. *As $N \to \infty$, $\min_{b \in 1,\cdots,B} \left(N_{[b]}\right) \to \infty$. Also, as $N_{[b]} \to \infty$, $\frac{1}{N_{[b]}} \mathbf{B}'_{[b]} \mathbf{1}_{N_{[b]}} \to \boldsymbol{\mu}_\beta$, $\frac{1}{N_{[b]}} \mathbf{B}'_{[b]} \mathbf{B}_{[b]} \to \mathbf{V}_\beta = \Sigma_\beta + \boldsymbol{\mu}_\beta \boldsymbol{\mu}'_\beta$, and $\frac{\mathbf{E}'_{[b]} \mathbf{E}_{[b]}}{N} \xrightarrow{p} \mathbf{V}_{e,[b]}$, where $\mathbf{V}_{e,[b]}$ is a $(\tau \times \tau)$ diagonal matrix.*

It is worth to highlight that this assumption allows the beta at the individual stock level to vary over time. We require that only the first two moments of the cross-sectional distribution factor loadings are stable over time. Also, note that from the last limit in the assumption, the variance of residuals can vary within a block as well as across blocks, similar to Jones (2001).

Lastly, we need the following regularity condition in our short-time block structure.

Assumption 4. *Consider a continuous function of* $\mathbf{F}_{[b]}$*;* $f \quad : \mathbb{R}^{\tau \times K} \to \mathbb{R}^m$*. Then, there exists a finite number* $M > 0$ *such that*

$$\lim_{T\to\infty} \frac{1}{B} \sum_{b=1}^{B} \sqrt{f\left(\mathbf{F}_{[b]}\right)' f\left(\mathbf{F}_{[b]}\right)} < M.$$

Assumption 4 simply restricts the behavior of factor realizations to be stationary enough so that the block-wise averages do not explode as T increases.

So far, we specified all assumptions that we need to construct a stochastic discount factor utilizing unbalanced panel data. Before we present our main theorem, we need to introduce an estimator of $\mathbf{V}_{e,[b]}$ which will be an essential element of our small-τ bias correction. A bias-correction in a short time series has been addressed in other papers (Litzenberger and Ramaswamy (1979), Shanken (1992)) and the relation of our correction to those papers will be discussed later. We utilize the estimator of $\mathbf{V}_{e,[b]}$ proposed by Kim and Skoulakis (2017).

Lemma 2.1. *Let Assumptions 1, 3, and 4 be in effect. Define* $\widehat{\mathbf{V}}_{e,[b]}$ *by*

$$\widehat{\mathbf{V}}_{e,[b]} = diag\left(\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1} \mathcal{S}' vec\left(\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right)\right), \tag{2.24}$$

where

$$\begin{aligned} \mathbf{H}_{[b]} &= \mathbf{J}_\tau - \mathbf{J}_\tau \mathbf{F}_{[b]} \left(\mathbf{F}'_{[b]}\mathbf{J}_\tau \mathbf{F}_{[b]}\right)^{-1} \mathbf{F}'_{[b]}\mathbf{J}_\tau, \\ \mathbf{J}_\tau &= \mathbf{I}_\tau - \frac{1}{\tau}\mathbf{1}_{\tau\times\tau}, \end{aligned} \tag{2.25}$$

the operator $\odot$ *denotes the Hadamard product (entry-wise product), the* (i,j)*-th element of the* $(\tau^2 \times \tau)$ *selection matrix of* $\mathcal{S}$ *is given by* $\mathbf{1}\left(i = (j-1)\,\tau + j\right)$ *and the* $(N \times \tau)$ *matrix of* $\widehat{\mathbf{E}}_{[b]}$ *is defined, for the case of using the gross returns, by* $\widehat{\mathbf{E}}_{[b]} = \mathbf{R}_{[b]}\mathbf{H}_{[b]}$ *and, for the case of using excess returns, by* $\widehat{\mathbf{E}}_{[b]} = \mathbf{R}^e_{[b]}\mathbf{H}_{[b]}$*. Then, it holds that as* $N, T \to \infty$*,* $\widehat{\mathbf{V}}_{e,[b]} \xrightarrow{p} \mathbf{V}_{e,[b]}$ *for each* $b = 1, \cdots, B$*.*

The intuition of the estimator defined by (2.24) follows. Given the expressions of $\mathbf{R}_{[b]}$ and $\mathbf{H}_{[b]}$ in (2.22) and (2.25), respectively, it holds that $\widehat{\mathbf{E}}_{[b]} = \mathbf{R}_{[b]}\mathbf{H}_{[b]} = \mathbf{E}_{[b]}\mathbf{H}_{[b]}$. Hence, Assumption 3 implies that $\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}} = \mathbf{H}_{[b]}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}\right)\mathbf{H}_{[b]} \xrightarrow{p} \mathbf{H}_{[b]}\mathbf{V}_{e,[b]}\mathbf{H}_{[b]}$. To extract the diagonal matrix of $\mathbf{V}_{e,[b]}$ in the probability limit of $\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}$, we manipulate the matrix of $\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}$ as in (2.24).

Lastly, we state our main theorem for unbalanced panels which shows that a consistent estimator of the SDF can be constructed with unbalanced panel data.

Theorem 2.2. *With Assumptions 1, 3, and 4, as* $N, T \to \infty$, $\widehat{m}_t = \widehat{\delta}_0 + \mathbf{f}'_t\widehat{\boldsymbol{\delta}}_f$ *and* $\widehat{m}^e_t = 1 + \mathbf{f}'_t\widehat{\boldsymbol{\delta}}^e$ *converge to* m_t *and* m^e_t *given in* (2.5) *and* (2.9), *respectively, when the* $((K+1)\times 1)$ *vector of* $\widehat{\boldsymbol{\delta}} = \left[\widehat{\delta}_0\ \widehat{\boldsymbol{\delta}}'_f\right]'$ *and the* $(K \times 1)$ *vector of* $\widehat{\boldsymbol{\delta}}^e$ *are constructed by*

$$\widehat{\boldsymbol{\delta}} = \mathbf{D}^{-1}\mathbf{U} \tag{2.26}$$

$$\widehat{\boldsymbol{\delta}}^e = -\left(\mathbf{D}^e\right)^{-1}\mathbf{U}^e, \tag{2.27}$$

where

$$\mathbf{D} = \left(\frac{\mathbf{F}'_\triangle\mathbf{F}_\triangle}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right)\right], \tag{2.28}$$

$$\mathbf{U} = \frac{1}{B}\sum_{b=1}^{B}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau}, \tag{2.29}$$

$$\mathbf{D}^e = \left(\frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^e_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right)\right], \tag{2.30}$$

$$\mathbf{U}^e = \left(\frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^e_{[b]}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right], \tag{2.31}$$

$$\mathbf{F}_{\triangle,[b]} = \left[\mathbf{1}_\tau\ \mathbf{F}_{[b]}\right], \quad \widehat{\mathbf{V}}_{e,[b]} = diag\left(\widehat{\mathbf{v}}_{e,[b]}\right)$$

and $\widehat{\mathbf{v}}_{e,[b]}$ *is given in* (2.24).

The intuition behind Theorem 2.2 follows. We focus on the case of $\widehat{\boldsymbol{\delta}}$ given by (2.26) because the underlying intuition can be applied to $\widehat{\boldsymbol{\delta}}^e$ given by (2.27) in a similar manner. First, compare the expression of $\widehat{\boldsymbol{\delta}}$ in (2.26) with that of $\widetilde{\boldsymbol{\delta}}$ in (2.13). The matrices $\mathbf{D}$ and $\mathbf{U}$ given by (2.28) and (2.29) are designed to mimic $\left(\frac{\mathbf{F}'_\triangle\mathbf{R}'\mathbf{R}\mathbf{F}_\triangle}{NT^2}\right)$ and $\left(\frac{\mathbf{F}'_\triangle\mathbf{R}'\mathbf{1}_N}{NT}\right)$ in (2.13), respectively. The probability limits of $\mathbf{U}$ and $\frac{\mathbf{F}'_\triangle\mathbf{R}'\mathbf{1}_N}{NT}$ are identical with Assumptions 1- 4. Next, we provide intuition that the matrix $\mathbf{D}$ mimics $\frac{\mathbf{F}'_\triangle\mathbf{R}'\mathbf{R}\mathbf{F}_\triangle}{NT^2}$.

For expositional simplicity, we consider the traded factor case, i.e., $\boldsymbol{\lambda}_f = \boldsymbol{\mu_f}$. Then, the return generating process of (2.3) can be rewritten as follows:

$$\mathbf{R}_{[b]} = \mathbf{X}\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]},$$

where $\mathbf{X} = \left[\lambda_0 \mathbf{1}_{N_{[b]}} \ \mathbf{B}_{[b]}\right]$. Then, with the realized value of the linear SDF over the b-th block, denoted by the $(\tau \times 1)$ vector of $\mathbf{m}_{[b]} = \mathbf{F}_{\triangle,[b]}\boldsymbol{\delta}$, the realized mispricing can be written as

$$\mathbf{1}_{N,[b]} - \frac{\mathbf{R}_{[b]}\mathbf{m}_{[b]}}{\tau} = \mathbf{1}_{N,[b]} - \frac{\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\delta} = \mathbf{1}_{N,[b]} - \left(\mathbf{X}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \frac{\mathbf{E}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)\boldsymbol{\delta}.$$

Hence, if we simply regress the true price of $\mathbf{1}_{N,[b]}$ on $\frac{\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}$ to estimate $\boldsymbol{\delta}$, a bias will be induced by the non-negligible term of $\frac{\mathbf{E}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}$ with the finite τ in the regressor. That is, even though this term has zero expectation, and disappears as τ approaches infinity, it is non-zero for any finite τ. This is why we need to deduct $\frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}$ from $\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2}$ as in (2.28), while we do not need such an adjustment for $\frac{\mathbf{F}'_{\triangle}\mathbf{R}'\mathbf{R}\mathbf{F}_{\triangle}}{NT^2}$ in (2.13). Furthermore, we need a slight adjustment of multiplying the inverse of the sample moment $\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)$ over the b-th block to properly average the adjusted values across blocks, and then we compensate the adjustment by multiplying the sample moment $\left(\frac{\mathbf{F}'_{\triangle}\mathbf{F}_{\triangle}}{T}\right)$ over the whole time series data.

Recall that, in the previous subsection on the balanced panel estimator, we show that the stochastic discount factor can be consistently estimated even without observing the true factors (see Corollary 2.1) through PCA-based methods. It turns out that we can replace $\mathbf{F}$ with $\mathbf{F}^*$ such that $\mathbf{F}^* = \mathbf{F}\mathcal{O} + o_p(1)$ for the unbalanced panel estimator and still consistently estimate the stochastic discount factor.

Corollary 2.2. *With Assumptions 1, 3, and 4, given a consistent factor estimator of* $\mathbf{F}^* = \mathbf{F}\mathcal{O} + o_p(1)$ *for a some rotation matrix of* $\mathcal{O}$, *as* $N, T \to \infty$, $\widehat{m}^*_t = \widehat{\delta}^*_0 + \mathbf{f}^{*\prime}_t\widehat{\boldsymbol{\delta}}^*_f$ *and* $\widehat{m}^{*e}_t = 1 + \mathbf{f}'_t\widehat{\boldsymbol{\delta}}^{*e}$ *converge to* m_t *and* m^e_t *given in* (2.5) *and* (2.9), *respectively, when the* $((K+1)\times 1)$ *vector of* $\widehat{\boldsymbol{\delta}}^* = \left[\widehat{\delta}^*_0 \ \widehat{\boldsymbol{\delta}}^{*\prime}_f\right]'$ *and the* $(K \times 1)$ *vector of* $\widehat{\boldsymbol{\delta}}^{*e}$ *are constructed by*

$$\widehat{\boldsymbol{\delta}}^* = (\mathbf{D}^*)^{-1}\mathbf{U}^* \tag{2.32}$$

$$\widehat{\boldsymbol{\delta}}^{*e} = -(\mathbf{D}^{*e})^{-1}\mathbf{U}^{*e}. \tag{2.33}$$

The matrices of $\mathbf{D}^*$, $\mathbf{U}^*$, $\mathbf{D}^{*e}$ *and* $\mathbf{U}^{*e}$ *are the analogues of* $\mathbf{D}$, $\mathbf{U}$, $\mathbf{D}^e$ *and* $\mathbf{U}^e$ *where* $\mathbf{F}$, $\mathbf{F}_\triangle$, $\mathbf{F}_{[b]}$ *and* $\mathbf{F}_{\triangle,[b]}$ *are replaced by* $\mathbf{F}^*$, $\mathbf{F}^*_\triangle = [\mathbf{1}_\tau \ \mathbf{F}^*]$, $\mathbf{F}^*_{[b]}$ *and* $\mathbf{F}^*_{\triangle,[b]} = \left[\mathbf{1}_\tau \ \mathbf{F}^*_{[b]}\right]$, *respectively, and* $\widehat{\mathbf{V}}^*_{e,[b]}$ *is given in lemma A.8.*

3 Performance of the SDF estimators in a simulated economy

We provide the simulation evidence on the properties of our SDF estimators. We simulate economies that are constructed so that returns follow a strict K-factor model and compare the estimated SDF with the true SDF. The simulation design is similar to that in Chen, Connor, and Korajczyk (2018).

3.1 Calibration

To simulate returns, we need to take a stance on the return generating process in (2.3). We consider three return generating processes implied by the CAPM, the Fama and French (1993) three-factor model (FF3), and the Fama and French (2015) five-factor model (FF5). For monthly factor returns of the three models as well as the risk free return, we use data from Ken French's database.[3] In particular, we use the U.S. value-weighted stock market excess returns for all of the three models, SMB (small minus big) and HML (high minus low) factors for FF3 and FF5, and RMW (robust minus weak) and CMA (conservative minus aggressive) factors for FF5. We use the factor realizations over 600 months (January 1967 to December 2016), to estimate the first two moments of factors: $\boldsymbol{\mu}_f = \frac{1}{600}\sum_{t=1}^{600}\mathbf{f}_t$ and $\Sigma_f = \frac{1}{600-1}\sum_{t=1}^{600}\left(\mathbf{f}_t - \boldsymbol{\mu}_f\right)\left(\mathbf{f}_t - \boldsymbol{\mu}_f\right)'$. The riskless gross return is estimated as the average of the gross realized risk free return over the same period: $\lambda_0 = \frac{1}{600}\sum_{t=1}^{600} R_{f,t}$.

To obtain the parameters for a large number of assets in the simulation we exploit all available individual stock returns over 600 months (January 1967 to December 2016) from the CRSP monthly database. We estimate the factor betas ($\boldsymbol{\beta}_i$) and the variances of residual returns $\left(\sigma_{i,\varepsilon} = \mathbb{E}\left[\varepsilon^2_{i,t}\right]\right)$ of individual stocks by regressing the excess returns of $R_{i,t} - R_{f,t}$ on a constant and a vector of factor returns:

$$R_{i,t} - R_{f,t} = \alpha_i + \boldsymbol{\beta}_i'\mathbf{f}_t + e_{i,t}.$$

After this process, we have the estimated betas ($\boldsymbol{\beta}_i$) and the variance of residual returns $\left(\sigma_{i,e} = \mathbb{E}\left[e^2_{i,t}\right]\right)$ for each 14,277 individual stocks which have more than 60 observations

[3] See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

over our sample period January 1967 to December 2016.

3.2 Simulation Evidence

We simulate economies for the three asset pricing models of CAPM, FF3, and FF5 with N stocks over T periods, where N and T are set by $N =$ 1,000; 2,000; 4,000; and 8,000 and T =60, 120, 240, and 480. The N stocks are randomly selected, without replacement, out of 14,277 stocks available on CRSP over our sample period. If the j-th asset in the simulation is chosen to be asset i from CRSP then it is assigned the beta vector($\boldsymbol{\beta}_i$) and the variance of residual returns $\left(\sigma_{i,e} = \mathbb{E}\left[e_{i,t}^2\right]\right)$ calibrated for the i-th stock in CRSP. We draw $\mathbf{f}_t \sim \mathcal{N}\left(\boldsymbol{\mu}_f, \Sigma_f\right)$ and $e_{i,t} \sim \mathcal{N}\left(0, \sigma_{i,e}^2\right)$ for $t = 1, \cdots, T$ and $j = 1, \cdots, N$ in each repetition. With the calibrated $\boldsymbol{\beta}_i$ and λ_0 and the simulated $\mathbf{f}_t$ and $e_{i,t}$, the return process described in (2.3) can be generated. Note that it holds $\boldsymbol{\lambda}_f = \boldsymbol{\mu}_f$ in this economy because $\mathbf{f}_t$ is traded for the three asset pricing models under consideration.

We examine the performance of our SDF estimator by comparing the estimated SDF with the true SDF given by (2.5) for gross returns and (2.9) for excess returns. In particular, since the estimated SDF is the true SDF plus estimation error, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m:

$$\widehat{m}_t = a + b \cdot m_t + error_t.$$

If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero, and the coefficient on the true SDF (b) is 1. We use these three statistics of R^2, a, and b as metrics for the performance of SDF estimator. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Tables 1 and 2 report the SDF estimator performance in a CAPM economy. We repeat the same exercise for FF3 (Tables 3 and 4) and FF5 (Tables 5 and 6). The results of the SDF estimators for gross (excess) returns are reported in Tables 1, 3 and 5 (Tables 2, 4 and 6). Panel A of each table shows the performance of the balanced panel estimators derived in Theorem 2.1. The results for the unbalanced panel estimators, derived in Theorem 2.2, follow in Panel B of each table. Furthermore, to investigate the implication of Corollary 2.1, stating that our SDF estimators are robust to the case of using estimated factors, we consider both cases of using true factors (Panels A-1 and B-1) and estimated factors (Panels A-2 and B-2). To estimate pervasive factors, we use the asymptotic principal components (APC) method of Connor and Korajczyk (1986) applied to the simulated returns. Lastly, for comparison with other SDF estimators, we report the performance of other estimators in Panel C of each table. For gross returns, we also consider the Pukthuanthong and Roll's (2017) estimator, which is appropriate for a large-N setting, and the conventional GMM estimator, requiring a small-N setting. Because Pukthuanthong and Roll's (2017) estimator is not applicable to excess returns, we only compare our estimators to the GMM estimator when we estimate the SDF from excess returns. We use a two-step efficient GMM estimator. For the case of gross returns, the small-N GMM estimator is obtained as follows. In the first step, we find $\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{(1)} = \arg\min_{\boldsymbol{\delta}} \mathbf{g}(\boldsymbol{\delta})' \mathbf{I}_N \mathbf{g}(\boldsymbol{\delta})$, where $\mathbf{g}(\boldsymbol{\delta}) = \mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_{\triangle}\boldsymbol{\delta}}{T}$. In the second step, we estimate $\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{(2)} = \arg\min_{\boldsymbol{\delta}} \mathbf{g}(\boldsymbol{\delta})' \mathbf{W}^{-1} \mathbf{g}(\boldsymbol{\delta})$, where $\mathbf{W} = \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{1}_N - \mathbf{r}_t \left[1 \ \ \mathbf{f}_t'\right] \widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{(1)}\right)\left(\mathbf{1}_N - \mathbf{r}_t \left[1 \ \ \mathbf{f}_t'\right] \widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{(1)}\right)'$, $\mathbf{r}_t$ is the t-th column of $\mathbf{R}$, and $\mathbf{f}_t'$ is the t-th row of $\mathbf{F}$. The SDF m_t is estimated as $\widehat{m}_{t,\texttt{GMM}} = \left[1 \ \ \mathbf{f}_t'\right] \widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{(2)}$.[4] For the small-$N$ GMM estimator, we utilize the following three sets of portfolios: (i) 10 portfolios formed on Market Beta, (ii) 25 Portfolios formed on Size and Book-to-Market, (iii) 25 Portfolios Formed on Operating Profitability and Investment. Note that the three sets of portfolios are motivated by the corresponding asset pricing models of CAPM, FF3 and FF5. We obtain the realized returns of the three sets of portfolios over January 1967 to December 2016 from Ken French's database (cited above) and estimate the beta and residual variance of each portfolio for each of the three asset pricing models with the identical methods used for individual stocks.

We start with the CAPM results in Tables 1 and 2. Because the SDF estimators using the true factor are linear combinations of a constant and the market excess returns, R^2 is always 1 by construction. Hence, we do not report R^2 for such cases as in Panel A-1 or Panel B-1. Interestingly, even for the cases using estimated factors, R^2 is very close to 1 as shown in Panels A-2 and B-2. In contrast to this, Panel C-1 of Table 1 show lower R^2 values for the Pukthuanthong and Roll's (2017) estimator. For example, for $N = 4000$ and $T = 120$, the average R^2 is less than 10%.[5] This evidence shows that being slightly less agnostic by imposing a more restrictive factor structure on asset returns and using a small number of extracted factors leads to significant improvement

[4]For the case of excess returns, the small-N GMM estimator is obtained as follows. In the first step, we find $\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{e,(1)} = \arg\min_{\boldsymbol{\delta}^e} \mathbf{g}^e(\boldsymbol{\delta}^e)' \mathbf{I}_N \mathbf{g}^e(\boldsymbol{\delta}^e)$, where $\mathbf{g}^e(\boldsymbol{\delta}^e) = \frac{\mathbf{R}\mathbf{1}_T}{T} - \frac{\mathbf{R}\mathbf{F}\boldsymbol{\delta}^e}{T}$. In the second step, we estimate $\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{e,(2)} = \arg\min_{\boldsymbol{\delta}^e} \mathbf{g}^e(\boldsymbol{\delta}^e)' \mathbf{W}^{-1} \mathbf{g}^e(\boldsymbol{\delta}^e)$, where $\mathbf{W} = \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{r}_t^e - \mathbf{f}_t'\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{e,(1)}\right)\left(\mathbf{r}_t^e - \mathbf{f}_t'\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{e,(1)}\right)'$, $\mathbf{r}_t^e$ is the t-th column of $\mathbf{R}^e$, and $\mathbf{f}_t'$ is the t-th row of $\mathbf{F}$. The SDF m_t^e is estimated as $\widehat{m}_{t,\texttt{GMM}}^e = 1 + \mathbf{f}_t'\widehat{\boldsymbol{\delta}}_{\texttt{GMM}}^{e,(2)}$. This estimator is not amenable to the case in which $N >> T$ due to the fact that the weighting matrix, $\mathbf{W}$, is not invertible in for that case.

[5]Some readers may find this result puzzling, given Proposition 2.1. However, untabulated simulation results show that increasing N to 1 million yields R^2 values for the Pukthuanthong and Roll (2017) estimator is close to 1.

in the performance of the estimated SDF, compared to the fully agnostic approach by Pukthuanthong and Roll (2017). In terms of the intercept (a) and slope (b) , Panel A of Table 1 shows that the balanced panel estimators have a clear bias although the bias decreases as T increases. For the unbalanced panel estimator with the bias correction, reported in Panel B, we set $\tau = 60$ and find that the correction clearly eliminates the bias even with short T. We observe this pattern for both cases of using the true factors (A-1 vs B-1) or the estimated factors (A-2 vs B-2). It is worth noting that the bias in the balanced panel estimator for excess returns disappears much faster than that for gross returns by comparing Panel A of Table 1 to that of Table 2. Lastly, we find that the performance of the small-N GMM estimator heavily depends on the choice of assets. The performance of estimates is obviously the best for 10 portfolios formed on Market Beta. This should be expected since it provides an ideal environment of wide cross-sectional variation of market beta.

The results for the FF3 and FF5 models are qualitatively similar across the two models, so we focus on FF3 (Tables 3 and 4). In terms of R^2, the performance using true factors and estimated factors become similar as N and T increase. For example, in Panel B of Table 4, when $N = 4000$ and $T = 480$, the average R^2 is 0.81 for the case using the true factors and 0.78 for the case of using estimated factors. The bias for the intercept (a) and slope (b) in Panel A of Table 3 is attenuated in Panel B due to the correction terms in the unbalanced panel estimator. However, when either N or T is small, the unbalanced panel estimator still suffers from the bias due to the bias in the estimated factors (Panel B-2). The performance of the small-N GMM estimator changes with test portfolios. It does not work well with the 10 portfolios formed on Market Beta, for which an econometrician would not find sufficient variation in the cross section of exposures to some factors in the FF3 or FF5 models. Although– the GMM estimator performs better with 25 portfolios formed by Size and Book-to-Market ratio or Operating Profitability and Investment, our large sample estimators perform at least as well as the GMM estimator when N is sufficiently large.

Lastly, we evaluate how well the penalized estimators using Lasso, given by (2.17) and (2.18), select priced factors. We construct returns of assets in the economy be priced by a certain K-factor model but estimate the stochastic discount factor from a larger set of factors. In particular, for Panel A (B) of Table 7, we set the returns of assets in the economy follow CAPM (FF3) but admit the stochastic discount factors to be estimated from a larger set of factors (MKT, SMB, and HML for the CAPM and MKT, SMB, HML, RMA, and CMW for the FF3 model). Table 7 reports the percentage of selected frequency of a given factor over 10,000 repetitions. We set $N = 2000$ and $T =$

480. Overall, the penalized estimators select the priced factors correctly in most cases. The priced factors in the simulation are selected at least 93%, and usually selected in excess of 98% of the simulations. Using gross returns, there is a tendency to over select the non-priced factors (8%-13% of the time). However, using excess returns the unpriced factors are chosen between 0% and 0.6% of the simulations. The Lasso-based selection seems to be very accurate for the SDF estimators using excess returns.

The simulation exercise shows that our SDF estimators have some desirable properties. As N and T increase to a size typical of financial panel data in developed markets, R^2 in the regression of the estimated SDF on the true SDF approaches 1. Furthermore, the intercept (slope) converges to 0 (1) with large, but empirically relevant, values of N and T. Also, we find that SDFs based on APC factor estimates perform similarly to those using known factors when N is large. The SDF estimators for the excess returns suffer less from the small T bias and show faster convergence to the true SDF. In addition, the Lasso-penalized estimator using excess returns selects price factors quite accurately. Resorting to the superior performance of SDF estimators using excess returns relative to those using gross returns, we mainly focus on the estimators utilizing excess returns in our empirical results, reported below.

4 Empirical Application

In this section, we apply our SDF estimators to U.S. equity return data. Recall that our SDF estimator can be used either for a set of factors proposed by a specific asset pricing model (e.g., Sharpe's (1964) CAPM) or a set of statistical factors (e.g, Connor and Korajczyk (1986)). Hence, we consider both cases.

In particular, we consider six asset pricing models: CAPM, FF3, HXZ4 (Hou, Xue and Zhang (2015)), FF5 (Fama and French (2015)), BS6 (Barillas and Shanken (2017)). As noted above, the CAPM is a model with a single factor of market excess return. FF3 considers two additional factors of size (SMB) and value (HML). HXZ4 augment the set of factors by adding profitability (ROE) and investment (I/A). However, they drop the value factor with the claim that the value factor becomes redundant with their two new factors. FF5 use different factors for profitability (RMW) and investment (CMA). BS6 revive the value factor by using the monthly updated version (HML devil) in conjunction with the momentum (MOM) factor. Barillas and Shanken (2017) argue that the model with the six factors of market, size, value, momentum, profitability (ROE) and investment (I/A) performs the best relative to other potential combinations.

To obtain statistical factors, we apply two methodologies. First, we apply APC by

Connor and Korajczyk (1986) to excess returns of a large cross section of individual stocks from CRSP. To manage missing data in individual stock returns, we use the expectation maximization (EM) algorithm. From this method, we obtain four systematic factors over the 600 months from January 1967 to December 2016. Second, we utilize a recent technique developed by Pelger and Lettau (2017). They propose applying PCA to a matrix strengthened by a signal on the average returns, $\left(\frac{\mathbf{R}^e\mathbf{R}^{e\prime}}{T} + \lambda\left(\frac{\mathbf{R}^e\mathbf{1}_T}{T}\right)\left(\frac{\mathbf{R}^e\mathbf{1}_T}{T}\right)'\right)$ with some $\lambda \geq 0$. We apply their method to the returns of 209 portfolios, composed of 16 sets of characteristic-based decile portfolios plus 49 industry portfolios,[6] and extract four systematic factors from the realized returns of the 209 portfolios over the 600 months.

4.1 Application to large cross section of portfolios

We start with a large cross section of portfolios, which are appropriate for our balanced panel estimators. We consider two different sets of portfolios. The first set is comprised of 1200 portfolios sorted by estimated expected returns over the sample period of 293 months from January 1990 to May 2014.[7] Freyberger, Neuhierl and Weber (2017) estimate expected returns of individual stocks by using a non-linear function of twenty one characteristics of individual stocks.[8] The rational behind our choice of portfolio construction is that a wide cross section of expected return should be aligned with a wide cross section of true factor loadings. The second set of portfolios is a collection of decile portfolios sorted independently on 16 characteristics plus 49 industry portfolios. In particular, we consider sixteen sets of decile portfolios sorted on various stock characteristics: accruals (Sloan (1996)), book-to-market ratio (Fama and French (1992, 1993)), cash flow-to-price ratio (Chan, et al. (1991)), dividend-to-price ratio (Litzenberger and Ramaswamy (1982)), earnings-to-price ratio (Basu (1983)), investment (Chen, Novy-Marx, and Zhang (2010)), long-term reversal (DeBondt and Thaler (1985)), market beta (Frazzini and Pederson (2011)), 12-2 past return (Jegadeesh and Titman (1993)), net share issues (Ikenberry, et al. (1995), Fama and French (2008)), operating profitability (Hou, et al. (2015), Fama and French (2015)), quality minus junk (Asness, et al. (2014)), residual variance (Ang, et al. (2006)), short-term reversal

[6]The 209 portfolios will be explained in details in subsection 4.1.2 where we use those portfolios as test assets.

[7]The number of portfolios is determined by the smallest number of individual stocks over the sample period so that there is at least one stock allocated to each portfolio in every time period.

[8]We thank Andreas Neuhierl for providing the data on the estimated expected returns of individual stocks.

(Jegadeesh (1990)), aggregate variance (Ang, et al. (2006)). For industry portfolios, we consider 49 Fama-French industry portfolios. All of these portfolio returns are obtained from Ken French's database (cited above) except the 10 quality minus junk portfolio returns, which are from the AQR Capital Management data library.[9]

4.1.1 1200 portfolios sorted by estimated expected returns

Table 8 reports the estimated values of $\boldsymbol{\delta}^e$, when we apply our balanced panel estimator using 1200 portfolios sorted by expected returns. The associated standard errors are estimated by bootstrap method of resampling assets with replacement and reported in parenthesis.[10] The sample periods are 293 months over January 1990 to May 2014.

The direction of weighting of factors in the estimated stochastic discount factor is pretty well aligned with the implications of each model. In CAPM, the coefficient for market excess returns is significantly negative (-8.07 with a standard error of 0.08). To give some context for this estimate, consider a CAPM world in which λ_m is 0.5% per month (6% per year) and $\sigma_m = 5.77\%$ per month (20% year). These parameters are close to historical averages. For this economy, (2.10) implies the true value of δ^e equals -8.66, which is close to the -8.07 reported in the first row of Table 8 For FF3 and HXZ4, every factor receives significant weighting in the estimated SDF. The Lasso estimator excludes HML (I/A) in the estimated SDF for the FF3 (HXZ4) model. In FF5, the weighting of HML is opposite in sign from the implications of the model. This might be due to the redundancy of HML factor in conjunction with investment and profitability factors as mentioned in Hou, Xue, and Zhang (2015) and Fama and French (2015). This conjecture is also consistent with the results from Lasso estimator. In FF5, HML is dropped from the estimated SDF using the Lasso penalty and the sizes of coefficients on CMA and RMW shrink substantially. Although the monthly updated value factor of HML (devil) still appears to be significantly priced in BS6, HML (devil) does not survive in Lasso estimation. Also, the performance of statistical factors is impressive. The first principal components contribute significantly to the estimated discount factor under both methods. Two out of the remaining three principal components affect the discount factor.

[9] See `https://www.aqr.com/library/data-sets`.

[10] Given N assets, we resample $N_{(s)} = N$ assets with replacement and estimate $\boldsymbol{\delta}_{(s)}$ for $s = 1, \cdots, 1000$. The standard errors are computed as the standard deviation of $\boldsymbol{\delta}_{(s)}$ over $s = 1, \cdots, 1000$.

4.1.2 209 portfolios (16 decile portfolios and 49 industry portfolios)

We proceed to the next set of 209 portfolios. For this set of portfolios, we have a longer sample period of 600 months from January 1967 to December 2016. The estimated coefficients, $\boldsymbol{\delta}^e$, along with their standard errors in parenthesis (calculated using the same bootstrap procedure described above) are reported in Table 9.

Results for the five asset pricing models are given in Panel A. For CAPM, the coefficient for market excess returns is still consistently negative but the magnitude is reduced by less then one third relative to the case of using 1200 portfolios sorted on expected returns. One explanation would be that most of the decile portfolios are constructed on the basis of return differences that are not explained by CAPM. Across FF3, HXZ4, FF5, and BS6, SMB does not appear to be important in the stochastic discount factor, and it is also dropped in the Lasso estimator. This confirms that the most of these anomalies are somewhat independent phenomena apart from size effects. Given the choice of 209 test portfolios, HXZ4 is the only model for which every factor significantly affects the discount factor with a positive risk premium (i.e., negative estimated δ). Although I/A is very important in the construction of the SDF in HXZ4, I/A becomes almost redundant and excluded with Lasso penalty in BS6. In contrast to the results in the in Table 8, MOM and HML (devil) are now aligned well with their historical returns, hinting that these two factors may play an important role in jointly explaining various anomalies. All of the four APC-based factors estimated from individual stocks receive significant weighting in the stochastic discount factor. However, for this set of portfolios, the SDF estimated via RP-PCA only has significant weighting of the first principal component.

4.2 Application to individual stock returns

In this subsection, we apply our unbalanced panel estimators to individual stock returns from CRSP. We consider all individual stocks which were traded in the three main exchanges of NYSE, AMEX, and NASDAQ over our sample period of 50 years from January 1967 to December 2016. The share code is required to be 10 or 11 so that only common stocks are included in our sample. We apply price filter of five dollars at the beginning of each month. Also, we drop individual stocks with a lifespan of less than 5 years. After applying the three filters, we obtain 10,112 individual stocks.

Table 10 reports the estimates, $\hat{\delta}$, from the unbalanced panel data from all individual stocks available in CRSP over our sample period. Standard errors are computed by bootstrap method and reported in parenthesis.

In Panel A, the behavior of the estimated stochastic discount factor aligns with intuition. Across all models, the coefficient on MKT is significant, with the expected negative sign. For FF3 and HXZ4, every factor is significant in the estimated SDF. Consistent with the findings from a large cross section of portfolios, the weight on HML is significantly positive. Most of the statistical factors appear to be important as reported in Panel B. In particular, all of the four principal components extracted from individual stocks are significantly related to the estimated SDF.

5 Conclusion

While a large panel of asset return data is available, the empirical asset pricing literature has tended to utilize small numbers of test assets in the cross section to test specific asset pricing models. We propose novel estimators of the stochastic discount factor (SDF) which can exploit a large panel data. Simulation evidence shows that our SDF estimators perform better than some other methods in an economy with risk matching that of the U.S. equity market. The bias correction for unbalanced panels works well in eliminating the bias associated with small time-series samples. A Lasso version of the estimator is able to exclude unpriced factors. When applied to actual return data, the relation between the estimated SDF and the pervasive factors is in line with long-run estimates of risk premia and factor risks. We find that the value factor, HML, from Fama and French (1993, 2015) is subsumed by investment and profitability factors for a number of sets of test assets. The statistical significance of OLS estimated risk premia and Lasso estimates often agree on which factors should not enter the SDF. For the Fama-French 3-factor model, both OLS and Lasso exclude HML while only the Lasso estimator excludes HML in the Fama-French 5-factor model. Both approaches exclude the investment factor in the HXZ4 model, but only Lasso excludes the investment factor in the BS6 model.

A Proofs

Lemma A.1. *With Assumption 1, it holds that as N, T increases,*

$$\frac{\mathbf{B}'_{\triangle}\mathbf{B}_{\triangle}}{N} \to \begin{bmatrix} 1 & \boldsymbol{\mu}'_{\beta} \\ \boldsymbol{\mu}_{\beta} & \mathbf{V}_{\beta} \end{bmatrix} = \mathbf{V}_{\beta_{\triangle}} \tag{A.1}$$

$$\frac{\mathbf{F}'_{\triangle}\mathbf{F}_{\triangle}}{T} \xrightarrow{p} \begin{bmatrix} 1 & \boldsymbol{\mu}'_{f} \\ \boldsymbol{\mu}_{f} & \mathbf{V}_{f} \end{bmatrix} = \mathbf{V}_{f_{\triangle}}. \tag{A.2}$$

Proof Assumption 1 implies that

$$\frac{\mathbf{B}'_{\triangle}\mathbf{B}_{\triangle}}{N} = \begin{bmatrix} 1 & \frac{\mathbf{1}'_N\mathbf{B}}{N} \\ \frac{\mathbf{B}'\mathbf{1}_N}{N} & \frac{\mathbf{B}'\mathbf{B}}{N} \end{bmatrix} \to \begin{bmatrix} 1 & \boldsymbol{\mu}'_{\beta} \\ \boldsymbol{\mu}_{\beta} & \mathbf{V}_{\beta} \end{bmatrix}$$

$$\frac{\mathbf{F}'_{\triangle}\mathbf{F}_{\triangle}}{T} = \begin{bmatrix} 1 & \frac{\mathbf{1}'_T\mathbf{F}}{T} \\ \frac{\mathbf{F}'\mathbf{1}_T}{T} & \frac{\mathbf{F}'\mathbf{F}}{T} \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 1 & \boldsymbol{\mu}'_{f} \\ \boldsymbol{\mu}_{f} & \mathbf{V}_{f} \end{bmatrix},$$

which completes the proof of the lemma. □

For simplicity, we use the following notation:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_0 & \mathbf{0}'_K \\ (\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f) & \mathbf{I}_K \end{bmatrix} \tag{A.3}$$

$$\boldsymbol{\Lambda}^e = \begin{bmatrix} (\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f) & \mathbf{I}_K \end{bmatrix}. \tag{A.4}$$

Proof of Theorem 2.1 Define $\mathbf{B}_{\triangle} = [\mathbf{1}_N\mathbf{B}]$. In addition, from Assumption 2, we have that

$$\frac{\mathbf{B}'_{\triangle}\mathbf{E}\mathbf{F}_{\triangle}}{NT} = \begin{bmatrix} \frac{\mathbf{1}'_N\mathbf{E}\mathbf{1}_T}{NT} & \frac{\mathbf{1}'_N\mathbf{E}\mathbf{F}}{NT} \\ \frac{\mathbf{B}'\mathbf{E}\mathbf{1}_T}{NT} & \frac{\mathbf{B}'\mathbf{E}\mathbf{F}}{NT} \end{bmatrix} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1)} \tag{A.5}$$

$$\frac{\mathbf{F}'_{\triangle}\mathbf{E}'\mathbf{E}\mathbf{F}_{\triangle}}{NT^2} = \begin{bmatrix} \frac{\mathbf{1}'_T\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2} & \frac{\mathbf{1}'_T\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2} \\ \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2} & \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2} \end{bmatrix} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1).} \tag{A.6}$$

First, we establish that $\widetilde{\boldsymbol{\delta}} \xrightarrow{p} \boldsymbol{\delta}$, implying $\widetilde{m}_t \xrightarrow{p} m_t$. Note that

$$\widetilde{\boldsymbol{\delta}} = \left(\frac{\mathbf{F}'_{\triangle}\mathbf{R}'\mathbf{R}\mathbf{F}_{\triangle}}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle}\mathbf{R}'\mathbf{1}_N}{NT}\right). \tag{A.7}$$

We rewrite the return generating process of $\mathbf{R}$ in (2.11) as

$$\begin{aligned}\mathbf{R} &= \mathbf{1}_N \lambda_0 \mathbf{1}_T' + \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E} \\ &= \mathbf{B}_\triangle \boldsymbol{\Lambda} \mathbf{F}_\triangle' + \mathbf{E}, \end{aligned} \tag{A.8}$$

where $\boldsymbol{\Lambda}$ is given by (A.3). From Assumptions 1, 2, and the limits of (A.1)-(A.6), we have that

$$\begin{aligned}\frac{\mathbf{F}_\triangle'\mathbf{R}'\mathbf{R}\mathbf{F}_\triangle}{NT^2} &= \frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{B}_\triangle}{N}\boldsymbol{\Lambda}\frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T} + \frac{\mathbf{F}_\triangle'\mathbf{E}'\mathbf{E}\mathbf{F}_\triangle}{NT^2} \\ &\quad + \frac{\mathbf{F}_\triangle'\mathbf{E}'\mathbf{B}_\triangle}{NT}\boldsymbol{\Lambda}\frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T} + \frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{E}\mathbf{F}_\triangle}{NT} \\ &\xrightarrow{p} \mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\mathbf{V}_{\beta_\triangle}\boldsymbol{\Lambda}\mathbf{V}_{f_\triangle}\end{aligned}$$

and that

$$\begin{aligned}\frac{\mathbf{F}_\triangle'\mathbf{R}'\mathbf{1}_N}{NT} &= \frac{\mathbf{F}_\triangle'\left(\mathbf{F}_\triangle\boldsymbol{\Lambda}'\mathbf{B}_\triangle' + \mathbf{E}'\right)\mathbf{1}_N}{NT} = \frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{1}_N}{N} + \frac{\mathbf{F}_\triangle'\mathbf{E}'\mathbf{1}_N}{NT} \\ &\xrightarrow{p} \mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\left[1\boldsymbol{\mu}_\beta'\right]'.\end{aligned}$$

Hence, from (A.7), it follows that

$$\widetilde{\boldsymbol{\delta}} \xrightarrow{p} \left(\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\mathbf{V}_{\beta_\triangle}\boldsymbol{\Lambda}\mathbf{V}_{f_\triangle}\right)^{-1}\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\left[1\boldsymbol{\mu}_\beta'\right]' = \mathbf{V}_{f_\triangle}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}_{\beta_\triangle}^{-1}\left[1\ \ \boldsymbol{\mu}_\beta'\right]'. \tag{A.9}$$

Since

$$\begin{aligned}\mathbf{V}_{f_\triangle}^{-1} &= \begin{bmatrix} \left(1 + \boldsymbol{\mu}_f'\Sigma_f^{-1}\boldsymbol{\mu}_f\right) & -\boldsymbol{\mu}_f'\boldsymbol{\Sigma}_f^{-1} \\ -\boldsymbol{\Sigma}_f^{-1}\boldsymbol{\mu}_f & \boldsymbol{\Sigma}_f^{-1}\end{bmatrix} \\ \boldsymbol{\Lambda}^{-1} &= \begin{bmatrix} \frac{1}{\lambda_0} & \mathbf{0}_K' \\ \frac{\mathbf{1}}{\boldsymbol{\lambda}_0}\left(\boldsymbol{\mu}_f - \boldsymbol{\lambda}_f\right) & \mathbf{I}_K\end{bmatrix} \\ \mathbf{V}_{\beta_\triangle}^{-1} &= \begin{bmatrix} \left(1 + \boldsymbol{\mu}_\beta'\Sigma_\beta^{-1}\boldsymbol{\mu}_\beta\right) & -\boldsymbol{\mu}_\beta'\Sigma_\beta^{-1} \\ -\boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta & \Sigma_\beta^{-1}\end{bmatrix},\end{aligned}$$

it follows that

$$\mathbf{V}_{f_\triangle}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}_{\beta_\triangle}^{-1}\left[1\boldsymbol{\mu}_\beta'\right]' = \frac{1}{\lambda_0}\begin{bmatrix}\left(1 + \boldsymbol{\mu}_f'\Sigma_f^{-1}\boldsymbol{\lambda}_f\right) \\ -\boldsymbol{\Sigma}_f^{-1}\boldsymbol{\lambda}_f\end{bmatrix} = \boldsymbol{\delta}. \tag{A.10}$$

Combining (A.9) and (A.10), we prove the first claim in the theorem.

Next, in a similar manner, we show that $\widetilde{\boldsymbol{\delta}}^e \xrightarrow{p} \boldsymbol{\delta}^e$, implying $\widetilde{m}_t^e \xrightarrow{p} m_t^e$. Note that

$$\widetilde{\boldsymbol{\delta}}^e = -\left(\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{F}}{NT^2}\right)^{-1}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{1}_T}{NT^2}. \tag{A.11}$$

The return generating process of $\mathbf{R}^e$ in (2.12) as

$$\begin{aligned}\mathbf{R}^e &= \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_{\boldsymbol{f}}\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E}\\ &= \mathbf{B}\boldsymbol{\Lambda}^e\mathbf{F}_\triangle' + \mathbf{E},\end{aligned} \tag{A.12}$$

where $\boldsymbol{\Lambda}^e$ is given by (A.4). From Assumptions 1 and 2, we have that

$$\begin{aligned}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{F}}{NT^2} &= \frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{B}}{N}\boldsymbol{\Lambda}^e\frac{\mathbf{F}_\triangle'\mathbf{F}}{T} + \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2}\\ &\quad + \frac{\mathbf{F}'\mathbf{E}'\mathbf{B}}{NT}\boldsymbol{\Lambda}^e\frac{\mathbf{F}_\triangle'\mathbf{F}}{T} + \frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{E}\mathbf{F}}{NT}\\ &\xrightarrow{p} \begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}'\end{aligned}$$

and that

$$\begin{aligned}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{1}_T}{NT^2} &= \frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{B}}{N}\boldsymbol{\Lambda}^e\frac{\mathbf{F}_\triangle'\mathbf{1}_T}{T} + \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2}\\ &\quad + \frac{\mathbf{F}'\mathbf{E}'\mathbf{B}}{NT}\boldsymbol{\Lambda}^e\frac{\mathbf{F}_\triangle'\mathbf{1}_T}{T} + \frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{E}\mathbf{1}_T}{NT}\\ &\xrightarrow{p} \begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix}1 & \boldsymbol{\mu}_f'\end{bmatrix}'.\end{aligned}$$

Hence, from (A.11), it follows that

$$\begin{aligned}\widetilde{\boldsymbol{\delta}}^e &\xrightarrow{p} -\left(\begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}'\right)^{-1}\begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix}1 & \boldsymbol{\mu}_f'\end{bmatrix}'\\ &= -\left(\boldsymbol{\Lambda}^e\begin{bmatrix}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f\end{bmatrix}'\right)^{-1}\boldsymbol{\Lambda}^e\begin{bmatrix}1 & \boldsymbol{\mu}_f'\end{bmatrix}' = -\left(\left(\boldsymbol{\lambda}_{\boldsymbol{f}} - \boldsymbol{\mu}_{\boldsymbol{f}}\right)\boldsymbol{\mu}_{\boldsymbol{f}}' + \mathbf{V}_f\right)^{-1}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f + \boldsymbol{\mu}_f\right)\\ &= -\left(\boldsymbol{\lambda}_{\boldsymbol{f}}\boldsymbol{\mu}_{\boldsymbol{f}}' + \boldsymbol{\Sigma}_f\right)^{-1}\boldsymbol{\lambda}_f = \boldsymbol{\delta}^e.\end{aligned} \tag{A.13}$$

From (A.10) and (A.13), we complete the proof of the theorem. □

Proof of Corollary 2.1 Note that it suffices to establish that $\widetilde{\boldsymbol{\delta}}^* \xrightarrow{p} \begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O}' \end{bmatrix} \boldsymbol{\delta}$, implying $\widetilde{m}_t^{*e} \xrightarrow{p} m_t$. Because $\mathbf{F}_\triangle^* = \mathbf{F}_\triangle \begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O} \end{bmatrix} + o_p(1)$, we have that

$$
\begin{aligned}
\widetilde{\boldsymbol{\delta}}^* &= \left(\frac{\mathbf{F}_\triangle^{*\prime} \mathbf{R}' \mathbf{R} \mathbf{F}_\triangle^*}{NT^2} \right)^{-1} \left(\frac{\mathbf{F}_\triangle^{*\prime} \mathbf{R}' \mathbf{1}_N}{NT} \right) \\
&= \left(\begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O}' \end{bmatrix} \frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{R} \mathbf{F}_\triangle}{NT^2} \begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O} \end{bmatrix} + o_p(1) \right)^{-1} \left(\begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O}' \end{bmatrix} \frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{1}_N}{NT} + o_p(1) \right) \\
&\xrightarrow{p} \begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O} \end{bmatrix} \lim_{N,T\to\infty} \left(\frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{R} \mathbf{F}_\triangle}{NT^2} \right)^{-1} \left(\frac{\mathbf{F}_\triangle' \mathbf{R}' \mathbf{1}_N}{NT} \right) = \begin{bmatrix} 1 & \mathbf{0}_K' \\ \mathbf{0}_K & \mathcal{O}' \end{bmatrix} \boldsymbol{\delta},
\end{aligned}
$$

where the last equality is from Theorem 2.1. This proves that $\widetilde{m}_t^{*e} \xrightarrow{p} m_t$.

Next, in a similar manner, we show that $\widetilde{\boldsymbol{\delta}}^{*e} \xrightarrow{p} \mathcal{O}' \boldsymbol{\delta}^e$, implying $\widetilde{m}_t^{*e} \xrightarrow{p} m_t^e$. Note that

$$
\begin{aligned}
\widetilde{\boldsymbol{\delta}}^{*e} &= - \left(\frac{\mathbf{F}^{*\prime} \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{F}^*}{NT^2} \right)^{-1} \left(\frac{\mathbf{F}^{*\prime} \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{1}_T}{NT^2} \right), \\
&= - \left(\mathcal{O}' \frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{F}}{NT^2} \mathcal{O}' + o_p(1) \right)^{-1} \left(\mathcal{O}' \frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{1}_T}{NT^2} + o_p(1) \right) \\
&\xrightarrow{p} \mathcal{O}' \lim_{N,T\to\infty} \left(\frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{F}}{NT^2} \right)^{-1} \left(\frac{\mathbf{F}' \mathbf{R}^{e\prime} \mathbf{R}^e \mathbf{1}_T}{NT^2} \right) = \mathcal{O}' \boldsymbol{\delta}^e,
\end{aligned}
$$

where the last equality is from Theorem 2.1. This proves that $\widetilde{m}_t^{*e} \xrightarrow{p} m_t^e$. The proof is complete. □

Proof of Proposition 2.1 Note that from Assumptions 1, 2, and the limit of (A.1) and the homoskedasticity condition, as $N \to \infty$,

$$
\begin{aligned}
\frac{\mathbf{R}'\mathbf{R}}{N} &= \mathbf{F}_\triangle \boldsymbol{\Lambda}' \frac{\mathbf{B}_\triangle' \mathbf{B}_\triangle}{N} \boldsymbol{\Lambda} \mathbf{F}_\triangle' + \frac{\mathbf{E}'\mathbf{E}}{N} + \frac{\mathbf{E}'\mathbf{B}_\triangle}{N} \boldsymbol{\Lambda} \mathbf{F}_\triangle' + \mathbf{F}_\triangle \boldsymbol{\Lambda}' \frac{\mathbf{B}_\triangle' \mathbf{E}}{N} \\
&\xrightarrow{p} \mathbf{F}_\triangle \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle} \boldsymbol{\Lambda} \mathbf{F}_\triangle' + s \mathbf{I}_T
\end{aligned}
\tag{A.14}
$$

$$
\frac{\mathbf{R}'\mathbf{1}_N}{N} = \mathbf{F}_\triangle \boldsymbol{\Lambda}' \frac{\mathbf{B}_\triangle' \mathbf{1}_N}{N} + \frac{\mathbf{E}'\mathbf{1}_N}{N} \xrightarrow{p} \mathbf{F}_\triangle \boldsymbol{\Lambda}' \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \end{bmatrix}'. \tag{A.15}
$$

From the N-limits of (A.14) and (A.15), some algebra shows that as $N \to \infty$,

$$\begin{aligned}
\breve{m}_t =& \iota_t' \left(\frac{\mathbf{R}'\mathbf{R}}{NT}\right)^{-1} \left(\frac{\mathbf{R}'\mathbf{1}_N}{N}\right) \\
\xrightarrow{p}& T\iota_t' \left(\left[\mathbf{F}_\triangle \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle} \boldsymbol{\Lambda} \mathbf{F}_\triangle'\right] + s\mathbf{I}_T\right)^{-1} \left(\mathbf{F}_\triangle \boldsymbol{\Lambda}' \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \end{bmatrix}'\right) \\
=& \iota_t' \left(\mathbf{F}_\triangle \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle}^{1/2} \left(\frac{s}{T}\mathbf{I}_{K+1} + \mathbf{V}_{\beta_\triangle}^{1/2} \boldsymbol{\Lambda} \frac{\mathbf{F}_\triangle' \mathbf{F}_\triangle}{T} \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle}^{1/2}\right)^{-1}\right) \left(\mathbf{V}_{\beta_\triangle}^{-1/2} \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \end{bmatrix}'\right) \\
=& \begin{bmatrix} 1 & \mathbf{f}_t' \end{bmatrix} \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle}^{1/2} \left(\frac{s}{T}\mathbf{I}_{K+1} + \mathbf{V}_{\beta_\triangle}^{1/2} \boldsymbol{\Lambda} \frac{\mathbf{F}_\triangle' \mathbf{F}_\triangle}{T} \boldsymbol{\Lambda}' \mathbf{V}_{\beta_\triangle}^{1/2}\right)^{-1} \mathbf{V}_{\beta_\triangle}^{-1/2} \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \end{bmatrix}'.
\end{aligned}$$

Hence, as $N \to \infty$ and then $T \to \infty$,

$$\begin{aligned}
\breve{m}_t &= \iota_t' \left(\frac{\mathbf{R}'\mathbf{R}}{NT^2}\right)^{-1} \left(\frac{\mathbf{R}'\mathbf{1}_N}{NT}\right) \\
&\xrightarrow{p} \begin{bmatrix} 1 & \mathbf{f}_t' \end{bmatrix} \left(\mathbf{V}_{f_\triangle}^{-1} \boldsymbol{\Lambda}^{-1}\right) \left(\mathbf{V}_{\beta_\triangle}^{-1} \begin{bmatrix} 1 & \boldsymbol{\mu}_\beta' \end{bmatrix}'\right) = \begin{bmatrix} 1 & \mathbf{f}_t' \end{bmatrix} \boldsymbol{\delta} = m_t,
\end{aligned}$$

where the next to the last equality is from (A.10). This completes the proof of the proposition. $\square$

For the proof of the rest, we define $\mathcal{S}$ as the $\left(\tau^2 \times \tau\right)$ selection matrix such that the $(\tau(s-1)+1, s)$ element of $\mathcal{S}$ is 1, for $s = 1, \cdots, \tau$ and all other elements are zero.

Proof of Lemma 2.1 Define the $(\tau \times 1)$ vector of $\mathbf{v}_{e,[b]}$ such that $\mathbf{V}_{e,[b]} = diag\left(\mathbf{v}_{e,[b]}\right)$. From the expression of $\widehat{\mathbf{V}}_{e,[b]}$ given in (2.24), we have that $\widehat{\mathbf{V}}_{e,[b]} = \text{diag}\left(\widehat{\mathbf{v}}_{e,[b]}\right)$, where

$$\widehat{\mathbf{v}}_{e,[b]} = \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1} \mathcal{S}' \text{vec}\left(\frac{\widehat{\mathbf{E}}_{[b]}' \widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right).$$

Hence, it suffices to show $\widehat{\mathbf{v}}_{e,[b]} \xrightarrow{p} \mathbf{v}_{e,[b]}$. The invertibility of $\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)$ is discussed in footnote 7 of Kim and Skoulakis (2017).

First, we verify the N-limit of vec $\left(\frac{\widehat{\mathbf{E}}_{[b]}' \widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right)$. Since $\mathbf{1}_\tau' \mathbf{H}_{[b]} = \mathbf{0}_\tau'$ and $\mathbf{F}_{[b]}' \mathbf{H}_{[b]} = \mathbf{0}_{K \times \tau}$, for both the gross returns case of (2.22) and the excess return case of (2.23), it holds that

$$\widehat{\mathbf{E}}_{[b]} = \mathbf{E}_{[b]} \mathbf{H}_{[b]}.$$

Using the property of vec$(\cdot)$ operator, we have that

$$\text{vec}\left(\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right) = \text{vec}\left(\mathbf{H}_{[b]}\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}\mathbf{H}_{[b]}\right) = \left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\text{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}\right)$$
$$\xrightarrow{p} \left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\text{vec}\left(\mathbf{V}_{e,[b]}\right),$$

where the last limit is from Assumption 4.

Hence, from the above limit and the properties of selection matrix of $\mathcal{S}$ such that vec $\left(\mathbf{V}_{e,[b]}\right) = \mathcal{S}\mathbf{v}_{e,[b]}$ and that $\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]} = \mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\mathcal{S}$, we have that

$$\begin{aligned}
\widehat{\mathbf{v}}_{e,[b]} &= \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\text{vec}\left(\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right) \\
&\xrightarrow{p} \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\text{vec}\left(\mathbf{V}_{e,[b]}\right) \\
&= \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\mathcal{S}\mathbf{v}_{e,[b]} \\
&= \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)\mathbf{v}_{e,[b]} = \mathbf{v}_{e,[b]},
\end{aligned}$$

completing the proof of the lemma. □

We use the following lemmas to prove Theorem 2.2.

Lemma A.2. *Let Assumptions 3 and 4 be in effect. Consider any set of functions of* $\mathbf{F}_{[b]}$*;* $f_{(1)} : \mathbb{R}^{\tau\times K} \to \mathbb{R}^{K}$*,* $f_{(2)} : \mathbb{R}^{\tau\times K} \to \mathbb{R}^{K^2}$*,* $f_{(3)} : \mathbb{R}^{\tau\times K} \to \mathbb{R}^{\tau}$*,* $f_{(4)} : \mathbb{R}^{\tau\times K} \to \mathbb{R}^{K\tau}$*, and* $f_{(4)} : \mathbb{R}^{\tau\times K} \to \mathbb{R}^{\tau^2}$*. Then, as* $N, T \to \infty$*, it holds that*

$$\frac{1}{B}\sum_{b=1}^{B} f_{(1)}\left(\mathbf{F}_{[b]}\right)'\left(\frac{1}{N_{[b]}}\mathbf{B}'_{[b]}\mathbf{1}_{N_{[b]}} - \boldsymbol{\mu}_{\beta}\right) \xrightarrow{p} 0$$

$$\frac{1}{B}\sum_{b=1}^{B} f_{(2)}\left(\mathbf{F}_{[b]}\right)' vec\left(\frac{1}{N_{[b]}}\mathbf{B}'_{[b]}\mathbf{B}_{[b]} - \mathbf{V}_{\beta}\right) \xrightarrow{p} 0$$

$$\frac{1}{B}\sum_{b=1}^{B} f_{(3)}\left(\mathbf{F}_{[b]}\right)'\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{1}_{N_{[b]}}\right) \xrightarrow{p} 0$$

$$\frac{1}{B}\sum_{b=1}^{B} f_{(4)}\left(\mathbf{F}_{[b]}\right)' vec\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{B}_{[b]}\right) \xrightarrow{p} 0$$

$$\frac{1}{B}\sum_{b=1}^{B} f_{(5)}\left(\mathbf{F}_{[b]}\right)' vec\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{E}_{[b]} - \mathbf{V}_{e,[b]}\right) \xrightarrow{p} 0.$$

Proof Let the $(m \times 1)$ vector of $e_{[b]}$ denote any of the five errors: $\left(\frac{1}{N_{[b]}}\mathbf{B}'_{[b]}\mathbf{1}_{N_{[b]}} - \boldsymbol{\mu}_{\beta}\right)$, vec $\left(\frac{1}{N_{[b]}}\mathbf{B}'_{[b]}\mathbf{B}_{[b]} - \mathbf{V}_{\beta}\right)$, $\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{1}_{N_{[b]}}\right)$, vec $\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{B}_{[b]}\right)$, and vec $\left(\frac{1}{N_{[b]}}\mathbf{E}'_{[b]}\mathbf{E}_{[b]} - \mathbf{V}_{e,[b]}\right)$. Note that from Assumption 3,

$$\max_{b\in 1,\cdots,B}\sqrt{e_{[b]}'e_{[b]}}\xrightarrow{p}0. \tag{A.16}$$

Let $f_{(k)}$ be the corresponding function. As T increases, it follows that

$$\begin{aligned}\frac{1}{B}\sum_{b=1}^{B}f_{(k)}\left(\mathbf{F}_{[b]}\right)'e_{[b]} &\leq \frac{1}{B}\sum_{b=1}^{B}\sqrt{f_{(k)}\left(\mathbf{F}_{[b]}\right)'f_{(k)}\left(\mathbf{F}_{[b]}\right)}\sqrt{e_{[b]}'e_{[b]}}\\ &\leq\left(\frac{1}{B}\sum_{b=1}^{B}\sqrt{f_{(k)}\left(\mathbf{F}_{[b]}\right)'f_{(k)}\left(\mathbf{F}_{[b]}\right)}\right)\max_{b\in 1,\cdots,B}\sqrt{e_{[b]}'e_{[b]}}\\ &\leq M\max_{b\in 1,\cdots,B}\sqrt{e_{[b]}'e_{[b]}}\xrightarrow{p}0, \end{aligned}\tag{A.17}$$

where the first inequality is from the Cauchy-Schwarz inequality, the third inequality is from Assumption 4, and the last limit is from (A.16). In a similar manner, we can show that as T increases,

$$-\frac{1}{B}\sum_{b=1}^{B}f_{(k)}\left(\mathbf{F}_{[b]}\right)'e_{[b]}\leq M\max_{b\in 1,\cdots,B}\sqrt{e_{[b]}'e_{[b]}}\xrightarrow{p}0. \tag{A.18}$$

Lastly, combining (A.17) and (A.18) in conjunction with the squeeze theorem, we have that as $N,T\to\infty$,

$$\frac{1}{B}\sum_{b=1}^{B}f_{(k)}\left(\mathbf{F}_{[b]}\right)'e_{[b]}\xrightarrow{p}0,$$

completing the proof of the lemma. □

Lemma A.3. *It holds that*

$$vec\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}}-diag\left(\widehat{\mathbf{v}}_{e,[b]}\right)\right)=\mathbf{K}_{[b]}vec\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}}-\mathbf{V}_{e,[b]}\right),$$

where

$$\mathbf{K}_{[b]}=\left(\mathbf{I}_{\tau^2}-\mathcal{S}\left(\mathbf{H}_{[b]}\odot\mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]}\otimes\mathbf{H}_{[b]}\right)\right),$$

and $\mathbf{H}_{[b]}$ *is given in* (2.25).

Proof From (2.24),

$$\text{vec}\left(diag\left(\widehat{\mathbf{v}}_{e,[b]}\right)\right)=\mathcal{S}\widehat{\mathbf{v}}_{e,[b]}=\mathcal{S}\left(\mathbf{H}_{[b]}\odot\mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]}\otimes\mathbf{H}_{[b]}\right)\text{vec}\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}}\right).$$

Hence,

$$
\begin{aligned}
&\operatorname{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - diag\left(\widehat{\mathbf{v}}_{e,[b]}\right)\right)\\
=&\left(\mathbf{I}_{\tau^2} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\right)\operatorname{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}\right)\\
=&\left(\mathbf{I}_{\tau^2} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\right)\operatorname{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \mathbf{V}_{e,[b]}\right),
\end{aligned}
$$

where the last equality is from

$$
\begin{aligned}
&\left(\mathbf{I}_{\tau^2} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\right)\operatorname{vec}\left(\mathbf{V}_{e,[b]}\right)\\
=&\left(\mathbf{I}_{\tau^2} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\right)\mathcal{S}\mathbf{v}_{e,[b]}\\
=&\mathcal{S}\mathbf{v}_{e,[b]} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\mathcal{S}\mathbf{v}_{e,[b]} = \mathbf{0}_{\tau^2}.
\end{aligned}
$$

This completes the proof of the lemma. □

Lemma A.4. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{D} \xrightarrow{p} \mathbf{V}_{f\triangle}\mathbf{\Lambda}'\mathbf{V}_{\beta\triangle}\mathbf{\Lambda}\mathbf{V}_{f\triangle},$$

where $\mathbf{D}$, $\mathbf{V}_{f\triangle}$, $\mathbf{V}_{\beta\triangle}$, and $\mathbf{\Lambda}$ are given in (2.28), (A.2), (A.1), *and* (A.3), *respectively.*

Proof We decompose $\mathbf{D}$ in (2.28) as

$$\mathbf{D} = \mathbf{D}_1\mathbf{D}_2,$$

where

$$
\begin{aligned}
\mathbf{D}_1 &= \frac{\mathbf{F}'_{\triangle}\mathbf{F}_{\triangle}}{T},\\
\mathbf{D}_2 &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right)\right].
\end{aligned}
$$

From (A.2),

$$\mathbf{D}_1 \xrightarrow{p} \mathbf{V}_{f\triangle}. \tag{A.19}$$

Hence, it suffices to show that $\mathbf{D}_2 \xrightarrow{p} \mathbf{\Lambda}'\mathbf{V}_{\beta\triangle}\mathbf{\Lambda}\mathbf{V}_{f\triangle}$. Rewrite $\mathbf{R}_{[b]}$ in (2.22) as

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]}\mathbf{\Lambda}\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}, \tag{A.20}$$

where $\mathbf{B}_{\triangle,[b]} = \left[\mathbf{1}_N \mathbf{B}_{[b]}\right]$ and $\boldsymbol{\Lambda}$ is given in (A.3). Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2} \\
=&\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}'\frac{\mathbf{B}'_{\triangle,[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\boldsymbol{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{F}_{\triangle,[b]}}{\tau} \\
+&\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}'\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}_{\triangle,[b]}}{\tau},
\end{aligned}$$

yielding

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right) = \boldsymbol{\Lambda}'\mathbf{V}_{\triangle,\beta}\boldsymbol{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \mathcal{E}_{D,[b]}, \tag{A.21}$$

where

$$\begin{aligned}
\mathcal{E}_{D,[b]} = &\boldsymbol{\Lambda}'\left(\frac{\mathbf{B}'_{\triangle,[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} - \mathbf{V}_{\triangle,\beta}\right)\boldsymbol{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} \\
&+ \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{F}_{\triangle,[b]}}{\tau} \\
&+ \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \boldsymbol{\Lambda}'\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}_{\triangle,[b]}}{\tau}.
\end{aligned}$$

Using the property of $\operatorname{vec}(ABC) = (C' \otimes A)\operatorname{vec}(B)$ and Lemma A.3, we have

$$\begin{aligned}
\operatorname{vec}\left(\mathcal{E}_{D,[b]}\right) = &\left(\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}'\right) \otimes \boldsymbol{\Lambda}'\right)\operatorname{vec}\left(\frac{\mathbf{B}'_{\triangle,[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} - \mathbf{V}_{\triangle,\beta}\right) \\
&+ \left(\frac{\mathbf{F}'_{\triangle,[b]}}{\tau} \otimes \left(\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\right)\mathbf{K}_{[b]}\operatorname{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \mathbf{V}_{e,[b]}\right) \\
&+ \left(\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}'\right) \otimes \left(\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\right)\operatorname{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right) \\
&+ \left(\frac{\mathbf{F}'_{\triangle,[b]}}{\tau} \otimes \boldsymbol{\Lambda}'\right)\operatorname{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}}\right)'\right),
\end{aligned}$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{D,[b]} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1)}$$

from Lemma A.2. Hence, from the expression of (A.21), we have

$$\begin{aligned}
\mathbf{D}_2 &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right)\right]. \\
&= \mathbf{\Lambda}'\mathbf{V}_\beta\mathbf{\Lambda}\frac{1}{B}\sum_{b=1}^{B}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{[b]} \\
&\xrightarrow{p} \mathbf{\Lambda}'\mathbf{V}_\beta\mathbf{\Lambda}\mathbf{V}_{f\triangle}. \qquad \text{(A.22)}
\end{aligned}$$

Combining (A.19) and (A.22), we have

$$\mathbf{D} = \mathbf{D}_1\mathbf{D}_2 \xrightarrow{p} \mathbf{V}_{f\triangle}\mathbf{\Lambda}'\mathbf{V}_{\beta\triangle}\mathbf{\Lambda}\mathbf{V}_{f\triangle},$$

completing the proof of the lemma. □

Lemma A.5. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{U} \xrightarrow{p} \mathbf{V}_{f_\triangle}\mathbf{\Lambda}'\left[1\ \boldsymbol{\mu}'_\beta\right]',$$

where $\mathbf{U}$ and $\mathbf{V}_{f\triangle}$ are given in (2.29) and (A.2), respectively.

Proof Recall the expression of (A.20):

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]}\mathbf{\Lambda}\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}. \qquad \text{(A.23)}$$

Note that

$$\begin{aligned}
\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau} &= \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}\mathbf{\Lambda}'}{\tau^2}\right)\left(\frac{\mathbf{B}_{\triangle,[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}}\right) + \frac{\mathbf{F}'_{\triangle,[b]}\mathbf{E}_{[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau^2} \\
&= \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)\mathbf{\Lambda}'\boldsymbol{\mu}_\beta + \mathcal{E}_{U,[b]}, \qquad \text{(A.24)}
\end{aligned}$$

where

$$\mathcal{E}_{U,[b]} = \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}\mathbf{\Lambda}'}{\tau^2}\right)\left(\frac{\mathbf{B}_{\triangle,[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}} - \boldsymbol{\mu}_\beta\right) + \left(\frac{\mathbf{F}'_{\triangle,[b]}}{\tau^2}\right)\left(\frac{\mathbf{E}_{[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}}\right).$$

From Lemma A.2,

$$\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{U,[b]} \xrightarrow{p} \mathbf{0}_{K+1}. \qquad \text{(A.25)}$$

Hence, from (A.24) and (A.25),

$$\mathbf{U} = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau}\right) \xrightarrow{p} \mathbf{V}_{f_\triangle}\mathbf{\Lambda}'\left[1\boldsymbol{\mu}'_\beta\right]',$$

completing the proof of the lemma. □

Lemma A.6. *Let Assumptions 1-4 be in effect. As* $N, T \to \infty$,

$$\mathbf{D}^e \xrightarrow{p} \left[\begin{array}{cc} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{array}\right] \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e \left[\begin{array}{cc} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{array}\right]',$$

where $\mathbf{D}^e$ *and* $\mathbf{\Lambda}^e$ *are given in* (2.30) *and* (A.4)*, respectively.*

Proof We decompose $\mathbf{D}^e$ in as

$$\mathbf{D}^e = \mathbf{D}^e_1\mathbf{D}^e_2,$$

where

$$\begin{aligned}
\mathbf{D}^e_1 &= \frac{\mathbf{F}'\mathbf{F}_\triangle}{T},\\
\mathbf{D}^e_2 &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^e_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right)\right].
\end{aligned}$$

From Assumption 1,

$$\mathbf{D}^e_1 \xrightarrow{p} \left[\begin{array}{cc} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{array}\right]. \tag{A.26}$$

Hence, it suffices to show that $\mathbf{D}^e_2 \xrightarrow{p} \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\left[\begin{array}{cc} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{array}\right]'$. Rewrite $\mathbf{R}_{[b]}$ in (2.23) as

$$\mathbf{R}^e_{[b]} = \mathbf{B}_{[b]}\mathbf{\Lambda}^e\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}. \tag{A.27}$$

Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\\
=&\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\mathbf{\Lambda}^e\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau} + \frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{F}_{[b]}}{\tau}\\
&+\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\mathbf{\Lambda}^e\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau} + \frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}_{[b]}}{\tau},
\end{aligned}$$

yielding

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2}-\frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right)=\mathbf{\Lambda}^{e\prime}\mathbf{V}_{\beta}\mathbf{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}+\mathcal{E}_{D^e,[b]}, \tag{A.28}$$

where

$$\begin{aligned}\mathcal{E}_{D^e,[b]}&=\mathbf{\Lambda}^{e\prime}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}-\mathbf{V}_{\beta}\right)\mathbf{\Lambda}^{e}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}\\&+\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}-\widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{F}_{[b]}}{\tau}\\&+\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\mathbf{\Lambda}^{e}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}+\mathbf{\Lambda}^{e\prime}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}_{[b]}}{\tau}.\end{aligned}$$

Using the property of $\mathrm{vec}\,(ABC)=(C'\otimes A)\,\mathrm{vec}\,(B)$ and Lemma A.3, we have

$$\begin{aligned}\mathrm{vec}\left(\mathcal{E}_{D^e,[b]}\right)&=\left(\frac{\mathbf{F}'_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\otimes\mathbf{\Lambda}^{e\prime}\right)\mathrm{vec}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}-\mathbf{V}_{\beta}\right)\\&+\left(\frac{\mathbf{F}'_{[b]}}{\tau}\otimes\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\mathbf{K}_{[b]}\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}-\mathbf{V}_{e,[b]}\right)\\&+\left(\frac{\mathbf{F}'_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\otimes\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\\&+\left(\frac{\mathbf{F}'_{[b]}}{\tau}\otimes\mathbf{\Lambda}^{e\prime}\right)\mathrm{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\right),\end{aligned}$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{D^e,[b]}\xrightarrow{p}\mathbf{0}_{(K+1)\times K}$$

from Lemma A.2. Hence, from the expression of (A.28), we have

$$\begin{aligned}\mathbf{D}_2^e&=\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2}-\frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right)\right].\\&=\mathbf{\Lambda}^{e\prime}\mathbf{V}_{\beta}\mathbf{\Lambda}^{e}\frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}\right)+\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{[b]}\\&\xrightarrow{p}\mathbf{\Lambda}^{e\prime}\mathbf{V}_{\beta}\mathbf{\Lambda}^{e}\left[\begin{array}{cc}\boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_{f}\end{array}\right]'.\end{aligned} \tag{A.29}$$

Combining (A.26) and (A.29), we have

$$\mathbf{D}^e = \mathbf{D}_1^e \mathbf{D}_2^e \xrightarrow{p} \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix} \boldsymbol{\Lambda}^{e\prime} \mathbf{V}_\beta \boldsymbol{\Lambda}^e \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}',$$

completing the proof of the lemma. □

Lemma A.7. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{U}^e \xrightarrow{p} \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix} \boldsymbol{\Lambda}^{e\prime} \mathbf{V}_\beta \boldsymbol{\Lambda}^e \begin{bmatrix} 1 & \boldsymbol{\mu}_{\boldsymbol{f}}' \end{bmatrix}',$$

where $\mathbf{U}^e$ and $\boldsymbol{\Lambda}^e$ are given in (2.31) *and* (A.4), *respectively.*

Proof We decompose $\mathbf{U}^e$ in as

$$\mathbf{U}^e = \mathbf{U}_1^e \mathbf{U}_2^e,$$

where

$$\mathbf{U}_1^e = \frac{\mathbf{F}'\mathbf{F}_\triangle}{T},$$
$$\mathbf{U}_2^e = \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right].$$

From Assumption 1,

$$\mathbf{U}_1^e \xrightarrow{p} \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}. \tag{A.30}$$

Hence, it suffices to show that $\mathbf{U}_2^e \xrightarrow{p} \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e \begin{bmatrix} 1 & \boldsymbol{\mu}_{\boldsymbol{f}}' \end{bmatrix}'$. Rewrite $\mathbf{R}_{[b]}$ in (2.23) as

$$\mathbf{R}_{[b]}^e = \mathbf{B}_{[b]}\boldsymbol{\Lambda}^e\mathbf{F}_{\triangle,[b]}' + \mathbf{E}_{[b]}. \tag{A.31}$$

Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\\
=&\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{1}_\tau}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}'}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{1}_\tau}{\tau}\\
+&\frac{\mathbf{F}_{\triangle,[b]}'}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{1}_\tau}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{1}_\tau}{\tau},
\end{aligned}$$

yielding

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{1}_\tau}{N_{[b]}\tau^2}-\frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)=\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{1}_\tau}{\tau}+\mathcal{E}_{U^e,[b]}, \tag{A.32}$$

where

$$\begin{aligned}\mathcal{E}_{U^e,[b]}=&\,\mathbf{\Lambda}^{e\prime}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}-\mathbf{V}_\beta\right)\mathbf{\Lambda}^e\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{1}_\tau}{\tau}\\&+\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}-\widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{1}_\tau}{\tau}\\&+\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\mathbf{\Lambda}^e\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{1}_\tau}{\tau}+\mathbf{\Lambda}^{e\prime}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{1}_\tau}{\tau}.\end{aligned}$$

Using the property of $\mathrm{vec}\,(ABC)=(C'\otimes A)\,\mathrm{vec}\,(B)$ and Lemma A.3, we have

$$\begin{aligned}\mathcal{E}_{U^e,[b]}=&\left(\frac{\mathbf{1}'_\tau\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\otimes\mathbf{\Lambda}^{e\prime}\right)\mathrm{vec}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}-\mathbf{V}_\beta\right)\\&+\left(\frac{\mathbf{1}'_\tau}{\tau}\otimes\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\mathbf{K}_{[b]}\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}-\mathbf{V}_{e,[b]}\right)\\&+\left(\frac{\mathbf{1}'_\tau\mathbf{F}_{\triangle,[b]}}{\tau}\mathbf{\Lambda}^{e\prime}\otimes\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}'_{\triangle,[b]}}{\tau}\right)\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\\&+\left(\frac{\mathbf{1}'_\tau}{\tau}\otimes\mathbf{\Lambda}^{e\prime}\right)\mathrm{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\right),\end{aligned}$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{D^e,[b]}\xrightarrow{p}\mathbf{0}_K$$

from Lemma A.2. Hence, from the expression of (A.32), we have

$$\begin{aligned}\mathbf{U}^e_2=&\,\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{1}_\tau}{N_{[b]}\tau^2}-\frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right].\\&=\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}\right)+\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{U^e,[b]}\\&\xrightarrow{p}\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\left[\begin{array}{cc}1&\boldsymbol{\mu}'_{\boldsymbol{f}}\end{array}\right]'.\end{aligned}\tag{A.33}$$

Combining (A.30) and (A.33), we have

$$\mathbf{U}^e = \mathbf{U}_1^e\mathbf{U}_2^e \xrightarrow{p} \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix} \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e \begin{bmatrix} 1 & \boldsymbol{\mu_f}' \end{bmatrix}',$$

completing the proof of the lemma. □

Now, we prove Theorem 2.2 using the above lemmas.

Proof of Theorem 2.2 We will show that $\widehat{\boldsymbol{\delta}} \xrightarrow{p} \boldsymbol{\delta}$ and $\widehat{\boldsymbol{\delta}}^e \xrightarrow{p} \boldsymbol{\delta}^e$, implying $\widehat{m}_t \xrightarrow{p} m_t$ and $\widehat{m}_t^e \xrightarrow{p} m_t^e$, respectively. completing the proof of the lemma. From Lemmas A.4 and A.5, we have that

$$\begin{aligned}\widehat{\boldsymbol{\delta}} = \mathbf{D}^{-1}\mathbf{U} &\xrightarrow{p} \left(\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\mathbf{V}_{\beta_\triangle}\boldsymbol{\Lambda}\mathbf{V}_{f_\triangle}\right)^{-1}\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\left[1\boldsymbol{\mu}_\beta'\right]' \\ &= \frac{1}{\lambda_0}\begin{bmatrix} \left(1+\boldsymbol{\mu}_f\boldsymbol{\Sigma}_f^{-1}\boldsymbol{\lambda}_f\right) \\ -\boldsymbol{\Sigma}_f^{-1}\boldsymbol{\lambda}_f \end{bmatrix} = \boldsymbol{\delta},\end{aligned}$$

where the next to the last equality is from (A.10). From Lemmas A.6 and A.7, we have that

$$\begin{aligned}\widehat{\boldsymbol{\delta}}^e &= -\left(\mathbf{D}^e\right)^{-1}\mathbf{U}^e \\ &\xrightarrow{p} -\left(\begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}'\right)^{-1}\begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix} 1 & \boldsymbol{\mu}_f' \end{bmatrix}' \\ &= -\left(\boldsymbol{\Lambda}^e\begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}'\right)^{-1}\boldsymbol{\Lambda}^e\begin{bmatrix} 1 & \boldsymbol{\mu}_f' \end{bmatrix}' = -\left(\left(\boldsymbol{\lambda_f}-\boldsymbol{\mu_f}\right)\boldsymbol{\mu_f}'+\mathbf{V}_f\right)^{-1}\left(\boldsymbol{\lambda}_f-\boldsymbol{\mu}_f+\boldsymbol{\mu}_f\right) \\ &= -\left(\boldsymbol{\lambda}_f\boldsymbol{\mu}_f+\boldsymbol{\Sigma}_f\right)^{-1}\boldsymbol{\lambda}_f = \boldsymbol{\delta}^e.\end{aligned}$$

The above two limits complete the proof of the theorem. □

For the rest of proofs, it is convenient to express $\mathbf{D}^*$, $\mathbf{D}^*$, $\mathbf{D}^*$ and $\mathbf{D}^*$ in terms of $\mathbf{F}^*$:

$$\mathbf{D}^* = \left(\frac{\mathbf{F}_\triangle^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}'\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}^*}{N_{[b]}\tau^2}-\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{F}_{\triangle,[b]}^*}{\tau^2}\right)\right], \tag{A.34}$$

$$\mathbf{U}^* = \frac{1}{B}\sum_{b=1}^{B}\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}'\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau}, \tag{A.35}$$

$$\mathbf{D}^{*e} = \left(\frac{\mathbf{F}^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^e\mathbf{F}_{[b]}^*}{N_{[b]}\tau^2}-\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{F}_{[b]}^*}{\tau^2}\right)\right], \tag{A.36}$$

$$\mathbf{U}^{*e} = \left(\frac{\mathbf{F}^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^e\mathbf{1}_\tau}{N_{[b]}\tau^2}-\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{1}_\tau}{\tau^2}\right)\right]. \tag{A.37}$$

Also, to prove Corollary 2.2, we need following lemmas, which confirm that Lemmas 2.1 and A.3-A.7 still hold for the estimated factors.

Lemma A.8. *With Assumptions 1, 3, and 4, as* $N, T \to \infty$, $\widehat{\mathbf{V}}^*_{e,[b]} = diag\left(\widehat{\mathbf{v}}^*_{e,[b]}\right) \xrightarrow{p} \mathbf{V}_{e,[b]}$ *for each* $b = 1, \cdots, B$, *where* $\widehat{\mathbf{v}}^*_{e,[b]}$ *is given by*

$$\widehat{\mathbf{v}}^*_{e,[b]} = \left(\mathbf{H}^*_{[b]} \odot \mathbf{H}^*_{[b]}\right)^{-1} \mathcal{S}' vec\left(\frac{\widehat{\mathbf{E}}^{*\prime}_{[b]}\widehat{\mathbf{E}}^*_{[b]}}{N_{[b]}}\right), \tag{A.38}$$

where $\widehat{\mathbf{E}}^*_{[b]}$ is defined by for the case of using the gross returns $\widehat{\mathbf{E}}^*_{[b]} = \mathbf{R}_{[b]}\mathbf{H}^*_{[b]}$ and for the case of using excess returns $\widehat{\mathbf{E}}_{[b]} = \mathbf{R}^e_{[b]}\mathbf{H}^*_{[b]}$and

$$\begin{aligned} \mathbf{H}^*_{[b]} &= \mathbf{J}_\tau - \mathbf{J}_\tau \mathbf{F}^*_{[b]}\left(\mathbf{F}^{*\prime}_{[b]}\mathbf{J}_\tau\mathbf{F}^*_{[b]}\right)^{-1}\mathbf{F}^{*\prime}_{[b]}\mathbf{J}_\tau \\ \mathbf{J}_\tau &= \mathbf{I}_\tau - \frac{1}{\tau}\mathbf{1}_{\tau\times\tau}. \end{aligned} \tag{A.39}$$

Proof Because $\mathbf{F}^*_{[b]} = \mathbf{F}_{[b]}\mathcal{O} + o_p(1)$, it follows that $\mathbf{H}^*_{[b]} = \mathbf{H}_{[b]} + o_p(1)$ and $\widehat{\mathbf{E}}^*_{[b]} = \widehat{\mathbf{E}}_{[b]} + o_p(1)$, implying that $\widehat{\mathbf{v}}^*_{e,[b]} = \widehat{\mathbf{v}}_{e,[b]} + o_p(1)$. Hence, by applying Lemma 2.1, we prove that $\widehat{\mathbf{V}}^*_{e,[b]} = diag\left(\widehat{\mathbf{v}}^*_{e,[b]}\right) \xrightarrow{p} \mathbf{V}_{e,[b]}$. □

Lemma A.9. *It holds that*

$$vec\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - diag\left(\widehat{\mathbf{v}}^*_{e,[b]}\right)\right) = \mathbf{K}_{[b]} vec\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \mathbf{V}_{e,[b]}\right) + o_p(1),$$

where

$$\mathbf{K}_{[b]} = \left(\mathbf{I}_{\tau^2} - \mathcal{S}\left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1}\mathcal{S}'\left(\mathbf{H}_{[b]} \otimes \mathbf{H}_{[b]}\right)\right),$$

and $\mathbf{H}_{[b]}$ *is given in* (2.25).

Proof From Lemmas 2.1 and A.8, we have $\widehat{\mathbf{v}}_{e,[b]} = \mathbf{v}_{e,[b]} + o_p(1)$ and $\widehat{\mathbf{v}}^*_{e,[b]} = \mathbf{v}_{e,[b]} + o_p(1)$, implying that

$$\widehat{\mathbf{v}}^*_{e,[b]} = \widehat{\mathbf{v}}_{e,[b]} + o_p(1),$$

which in conjunction with Lemma A.3 proves the claim of the lemma. □

Lemma A.10. *With Assumptions 1, 3, and 4, as* $N, T \to \infty$,

$$\mathbf{D}^* \xrightarrow{p} \mathcal{O}'_\triangle \mathbf{V}_{f\triangle}\mathbf{\Lambda}'\mathbf{V}_{\beta\triangle}\mathbf{\Lambda}\mathbf{V}_{f\triangle}\mathcal{O}_\triangle,$$

where

$$\mathcal{O}_\triangle = \begin{bmatrix} 1 & \mathbf{0}'_K \\ \mathbf{0}_K & \mathcal{O} \end{bmatrix}, \tag{A.40}$$

and $\mathbf{D}^*$, $\mathbf{V}_{f\triangle}$, $\mathbf{V}_{\beta\triangle}$, *and* $\mathbf{\Lambda}$ *are given in* (A.34), (A.2), (A.1), *and* (A.3), *respectively.*

Proof We decompose $\mathbf{D}^*$ in (A.34) as

$$\mathbf{D}^* = \mathbf{D}_1^* \mathbf{D}_2^*,$$

where

$$\mathbf{D}_1^* = \frac{\mathbf{F}_\triangle^{*\prime} \mathbf{F}_\triangle^*}{T},$$
$$\mathbf{D}_2^* = \frac{1}{B} \sum_{b=1}^{B} \left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} \right)^{-1} \left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{R}_{[b]}^\prime \mathbf{R}_{[b]} \mathbf{F}_{\triangle,[b]}^*}{N_{[b]} \tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \widehat{\mathbf{V}}_{e,[b]}^* \mathbf{F}_{\triangle,[b]}^*}{\tau^2} \right) \right].$$

Because

$$\mathbf{D}_1^* = \frac{\mathbf{F}_\triangle^{*\prime} \mathbf{F}_\triangle^*}{T} = \mathcal{O}_\triangle^\prime \frac{\mathbf{F}_\triangle^\prime \mathbf{F}_\triangle}{T} \mathcal{O}_\triangle + o_p(1),$$

from (A.2),

$$\mathbf{D}_1^* \xrightarrow{p} \mathcal{O}_\triangle^\prime \mathbf{V}_{f\triangle} \mathcal{O}_\triangle. \tag{A.41}$$

Hence, it suffices to show that $\mathbf{D}_2 \xrightarrow{p} \mathcal{O}_\triangle^\prime \mathbf{\Lambda}^\prime \mathbf{V}_{\beta\triangle} \mathbf{\Lambda} \mathbf{V}_{f\triangle} \mathcal{O}_\triangle$. Rewrite $\mathbf{R}_{[b]}$ in (2.22) using $\mathbf{F}_{\triangle,[b]} = \mathbf{F}_{\triangle,[b]}^* \mathcal{O}_\triangle^\prime + o_p(1)$ as

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]} \mathbf{\Lambda} \mathcal{O}_\triangle^\prime \mathbf{F}_{\triangle,[b]}^{*\prime} + \left(\mathbf{B}_{\triangle,[b]} \mathbf{\Lambda} \right) o_p(1) + \mathbf{E}_{[b]}, \tag{A.42}$$

where $\mathbf{B}_{\triangle,[b]} = \left[\mathbf{1}_N \mathbf{B}_{[b]} \right]$ and $\mathbf{\Lambda}$ is given in (A.3). Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{R}_{[b]}^\prime \mathbf{R}_{[b]} \mathbf{F}_{\triangle,[b]}^*}{N_{[b]} \tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \widehat{\mathbf{V}}_{e,[b]}^* \mathbf{F}_{\triangle,[b]}^*}{\tau^2} \\
=& \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} \mathcal{O}_\triangle^\prime \mathbf{\Lambda}^\prime \frac{\mathbf{B}_{\triangle,[b]}^\prime \mathbf{B}_{\triangle,[b]}}{N_{[b]}} \mathbf{\Lambda} \mathcal{O}_\triangle \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau} \left(\frac{\mathbf{E}_{[b]}^\prime \mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}^* \right) \frac{\mathbf{F}_{\triangle,[b]}^*}{\tau} \\
&+ \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau} \left(\frac{\mathbf{E}_{[b]}^\prime \mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right) \mathbf{\Lambda} \mathcal{O}_\triangle \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} \mathcal{O}_\triangle^\prime \mathbf{\Lambda}^\prime \left(\frac{\mathbf{E}_{[b]}^\prime \mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right)^\prime \frac{\mathbf{F}_{\triangle,[b]}^*}{\tau} + o_p(1),
\end{aligned}$$

yielding

$$\begin{aligned}
&\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} \right)^{-1} \left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{R}_{[b]}^\prime \mathbf{R}_{[b]} \mathbf{F}_{\triangle,[b]}^*}{N_{[b]} \tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \widehat{\mathbf{V}}_{e,[b]}^* \mathbf{F}_{\triangle,[b]}^*}{\tau^2} \right) \\
=& \mathcal{O}_\triangle^\prime \mathbf{\Lambda}^\prime \mathbf{V}_{\triangle,\beta} \mathbf{\Lambda} \mathcal{O}_\triangle \frac{\mathbf{F}_{\triangle,[b]}^{*\prime} \mathbf{F}_{\triangle,[b]}^*}{\tau} + \mathcal{E}_{D,[b]}^* + o_p(1),
\end{aligned} \tag{A.43}$$

where

$$\mathcal{E}^*_{D,[b]} = \mathcal{O}'_\triangle \mathbf{\Lambda}' \left(\frac{\mathbf{B}'_{\triangle,[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} - \mathbf{V}_{\triangle,\beta} \right) \mathbf{\Lambda} \mathcal{O}_\triangle \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}$$
$$+ \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} \right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}^*_{e,[b]} \right) \frac{\mathbf{F}^*_{\triangle,[b]}}{\tau}$$
$$+ \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} \right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right) \mathbf{\Lambda} \mathcal{O}_\triangle \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} + \mathcal{O}'_\triangle \mathbf{\Lambda}' \left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right)' \frac{\mathbf{F}^*_{\triangle,[b]}}{\tau}.$$

Using the property of $\mathrm{vec}\,(ABC) = (C' \otimes A)\,\mathrm{vec}\,(B)$ and Lemma A.9, we have

$$\mathrm{vec}\left(\mathcal{E}^*_{D,[b]}\right) = \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} (\mathbf{\Lambda}\mathcal{O}_\triangle)' \right) \otimes (\mathbf{\Lambda}\mathcal{O}_\triangle)' \right) \mathrm{vec}\left(\frac{\mathbf{B}'_{\triangle,[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} - \mathbf{V}_{\triangle,\beta} \right)$$
$$+ \left(\frac{\mathbf{F}^*_{\triangle,[b]}}{\tau} \otimes \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} \right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \right) \right) \mathbf{K}_{[b]} \mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \mathbf{V}_{e,[b]} \right)$$
$$+ \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} (\mathbf{\Lambda}\mathcal{O}_\triangle)' \right) \otimes \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} \right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \right) \mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right)$$
$$+ \left(\frac{\mathbf{F}^*_{\triangle,[b]}}{\tau} \otimes (\mathbf{\Lambda}\mathcal{O}_\triangle)' \right) \mathrm{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{\triangle,[b]}}{N_{[b]}} \right)' \right) + o_p(1),$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B} \mathcal{E}^*_{D,[b]} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1)}$$

from Lemma A.2. Hence, from the expression of (A.43), we have

$$\mathbf{D}^*_2 = \frac{1}{B}\sum_{b=1}^{B} \left[\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} \right)^{-1} \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}^*_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau^2} \right) \right].$$
$$= \mathcal{O}'_\triangle \mathbf{\Lambda}' \mathbf{V}_{\triangle,\beta} \mathbf{\Lambda} \mathcal{O}_\triangle \frac{1}{B}\sum_{b=1}^{B} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau} + \frac{1}{B}\sum_{b=1}^{B} \mathcal{E}^*_{D,[b]} + o_p(1)$$
$$\xrightarrow{p} \mathcal{O}'_\triangle \mathbf{\Lambda}' \mathbf{V}_{\beta\triangle} \mathbf{\Lambda} \mathbf{V}_{f\triangle} \mathcal{O}_\triangle. \tag{A.44}$$

Combining (A.41) and (A.44), we have

$$\mathbf{D}^* = \mathbf{D}^*_1\mathbf{D}^*_2 \xrightarrow{p} \mathcal{O}'_\triangle \mathbf{V}_{f\triangle} \mathbf{\Lambda}' \mathbf{V}_{\beta\triangle} \mathbf{\Lambda} \mathbf{V}_{f\triangle} \mathcal{O}_\triangle,$$

completing the proof of the lemma. □

Lemma A.11. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{U}^* \xrightarrow{p} \mathcal{O}'_{\triangle} \mathbf{V}_{f\triangle} \mathbf{\Lambda}' \left[1 \ \boldsymbol{\mu}'_{\beta} \right]',$$

where $\mathcal{O}_{\triangle}$, $\mathbf{U}^*$ and $\mathbf{V}_{f\triangle}$ are given by (A.40), (A.35) *and* (A.2), *respectively.*

Proof Recall the expression of (A.20) using $\mathbf{F}_{\triangle,[b]} = \mathbf{F}^*_{\triangle,[b]} \mathcal{O}'_{\triangle} + o_p(1)$ as

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]} \mathbf{\Lambda} \mathcal{O}_{\triangle} \mathbf{F}^{*\prime}_{\triangle,[b]} + \left(\mathbf{B}_{\triangle,[b]} \mathbf{\Lambda} \right) o_p(1) + \mathbf{E}_{[b]}. \tag{A.45}$$

Note that

$$\begin{aligned} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]} \mathbf{R}'_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}\tau} &= \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]} \mathbf{F}^*_{\triangle,[b]}}{\tau} \left(\mathbf{\Lambda} \mathcal{O}_{\triangle} \right)' \right) \left(\frac{\mathbf{B}_{\triangle,[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}} \right) + \frac{\mathbf{F}^{*\prime}_{\triangle,[b]} \mathbf{E}_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}\tau} + o_p(1) \\ &= \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]} \mathbf{F}^*_{\triangle,[b]}}{\tau} \right) \left(\mathbf{\Lambda} \mathcal{O}_{\triangle} \right)' \boldsymbol{\mu}_{\beta} + \mathcal{E}^*_{U,[b]} + o_p(1), \end{aligned} \tag{A.46}$$

where

$$\mathcal{E}^*_{U,[b]} = \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]} \mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \left(\mathbf{\Lambda} \mathcal{O}_{\triangle} \right)' \right) \left(\frac{\mathbf{B}_{\triangle,[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}} - \boldsymbol{\mu}_{\beta} \right) + \left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau} \right) \left(\frac{\mathbf{E}_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}} \right).$$

From Lemma A.2,

$$\frac{1}{B} \sum_{b=1}^{B} \mathcal{E}^*_{U,[b]} \xrightarrow{p} \mathbf{0}_{K+1}. \tag{A.47}$$

Hence, from (A.46) and (A.47),

$$\mathbf{U}^* = \frac{1}{B} \sum_{b=1}^{B} \left(\frac{\mathbf{F}'_{\triangle,[b]} \mathbf{R}'_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}\tau} \right) \xrightarrow{p} \mathcal{O}_{\triangle} \mathbf{V}_{f\triangle} \mathcal{O}'_{\triangle} \left(\mathbf{\Lambda} \mathcal{O}_{\triangle} \right)' \left[1 \boldsymbol{\mu}'_{\beta} \right]' = \mathcal{O}_{\triangle} \mathbf{V}_{f\triangle} \mathbf{\Lambda}' \left[1 \boldsymbol{\mu}'_{\beta} \right]',$$

completing the proof of the lemma. □

Lemma A.12. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{D}^{*e} \xrightarrow{p} \mathcal{O}' \left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array} \right] \mathbf{\Lambda}^{e\prime} \mathbf{V}_{\beta} \mathbf{\Lambda}^{e} \left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array} \right]' \mathcal{O},$$

where $\mathbf{D}^{*e}$ and $\mathbf{\Lambda}^e$ are given in (A.36) *and* (A.4), *respectively.*

Proof We decompose $\mathbf{D}^{*e}$ in as

$$\mathbf{D}^{*e} = \mathbf{D}^{*e}_1 \mathbf{D}^{*e}_2,$$

where

$$\mathbf{D}_1^{*e} = \frac{\mathbf{F}^{*\prime}\mathbf{F}^*_{\triangle}}{T},$$

$$\mathbf{D}_2^{*e} = \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{F}^*_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau^2}\right)\right].$$

From Assumption 1,

$$\mathbf{D}_1^{*e} \xrightarrow{p} \mathcal{O}'\begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{bmatrix}\mathcal{O}_{\triangle}. \tag{A.48}$$

Hence, it suffices to show that $\mathbf{D}_2^{*e} \xrightarrow{p} \mathcal{O}'_{\triangle}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^{e}\begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{bmatrix}'\mathcal{O}$. Rewrite $\mathbf{R}_{[b]}$ in (2.23) as

$$\mathbf{R}^e_{[b]} = \mathbf{B}_{[b]}\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\mathbf{F}^{*\prime}_{\triangle,[b]} + \left(\mathbf{B}_{[b]}\boldsymbol{\Lambda}^e\right)o_p\left(1\right) + \mathbf{E}_{[b]}. \tag{A.49}$$

Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}_{[b]}\mathbf{F}^*_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{F}^*_{[b]}}{\tau^2}\\
=&\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\mathcal{O}_{\triangle}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau} + \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}^*_{e,[b]}\right)\frac{\mathbf{F}^*_{[b]}}{\tau}\\
&+\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau} + \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\mathcal{O}_{\triangle}\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}^*_{[b]}}{\tau} + o_p\left(1\right),
\end{aligned}$$

yielding

$$\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}_{[b]}\mathbf{F}^*_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{F}^*_{[b]}}{\tau^2}\right) \tag{A.50}$$

$$=\mathcal{O}_{\triangle}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau} + \mathcal{E}^*_{D^e,[b]}, \tag{A.51}$$

where

$$\begin{aligned}
\mathcal{E}^*_{D^e,[b]} =& \mathcal{O}_{\triangle}\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}} - \mathbf{V}_\beta\right)\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau}\\
&+\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}^*_{e,[b]}\right)\frac{\mathbf{F}^*_{[b]}}{\tau}\\
&+\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}^e\mathcal{O}'_{\triangle}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau} + \mathcal{O}_{\triangle}\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{F}^*_{[b]}}{\tau}.
\end{aligned}$$

Using the property of $\mathrm{vec}\,(ABC) = (C' \otimes A)\,\mathrm{vec}\,(B)$ and Lemma A.3, we have

$$\begin{aligned}
\mathrm{vec}\left(\mathcal{E}_{D^e,[b]}\right) &= \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau}\left(\mathcal{O}_\triangle\mathbf{\Lambda}^{e\prime}\right)\right)\otimes\left(\mathcal{O}_\triangle\mathbf{\Lambda}^{e\prime}\right)\right)\mathrm{vec}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}-\mathbf{V}_\beta\right)\\
&+\left(\frac{\mathbf{F}^{*\prime}_{[b]}}{\tau}\otimes\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\right)\mathbf{K}_{[b]}\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}}-\mathbf{V}_{e,[b]}\right)\\
&+\left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau}\left(\mathcal{O}_\triangle\mathbf{\Lambda}^{e\prime}\right)\right)\otimes\left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\right)\right)\mathrm{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)\\
&+\left(\frac{\mathbf{F}^{*\prime}_{[b]}}{\tau}\otimes\left(\mathcal{O}_\triangle\mathbf{\Lambda}^{e\prime}\right)\right)\mathrm{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\right)+o_p\left(1\right),
\end{aligned}$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}^*_{D^e,[b]}\xrightarrow{p}\mathbf{0}_{(K+1)\times K}$$

from Lemma A.2. Hence, from the expression of (A.50), we have

$$\begin{aligned}
\mathbf{D}^{*e}_2 &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}_{[b]}\mathbf{F}^*_{[b]}}{N_{[b]}\tau^2}-\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{F}^*_{[b]}}{\tau^2}\right)\right].\\
&= \mathcal{O}_\triangle\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\mathcal{O}'_\triangle\frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{[b]}}{\tau}\right)+\frac{1}{B}\sum_{b=1}^{B}\mathcal{E}^*_{D^e,[b]}\\
&\xrightarrow{p}\boldsymbol{\mathcal{O}}_\triangle\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\left[\begin{array}{cc}\boldsymbol{\mu_f} & \mathbf{V}_f\end{array}\right]'\mathcal{O}'. \qquad\qquad (\text{A.52})
\end{aligned}$$

Combining (A.48) and (A.52), we have

$$\mathbf{D}^{*e}=\mathbf{D}^{*e}_1\mathbf{D}^{*e}_2\xrightarrow{p}\mathcal{O}'\left[\begin{array}{cc}\boldsymbol{\mu_f} & \mathbf{V}_f\end{array}\right]\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\left[\begin{array}{cc}\boldsymbol{\mu_f} & \mathbf{V}_f\end{array}\right]'\mathcal{O},$$

completing the proof of the lemma. □

Lemma A.13. *Let Assumptions 1-4 be in effect. As $N, T \to \infty$,*

$$\mathbf{U}^{*e}\xrightarrow{p}\mathcal{O}\left[\begin{array}{cc}\boldsymbol{\mu}_f & \mathbf{V}_f\end{array}\right]\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\left[\begin{array}{cc}1 & \boldsymbol{\mu}'_f\end{array}\right]',$$

where $\mathbf{U}^{*e}$ and $\mathbf{\Lambda}^e$ are given in (A.37) *and* (A.4), *respectively.*

Proof We decompose $\mathbf{U}^{*e}$ in as

$$\mathbf{U}^{*e}=\mathbf{U}^{*e}_1\mathbf{U}^{*e}_2,$$

where

$$\mathbf{U}_1^{*e} = \frac{\mathbf{F}^{*\prime}\mathbf{F}_\triangle^*}{T},$$
$$\mathbf{U}_2^{*e} = \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{1}_\tau}{\tau^2}\right)\right].$$

From Assumption 1,

$$\mathbf{U}_1^{*e} \xrightarrow{p} \mathcal{O}'\left[\begin{array}{cc} \boldsymbol{\mu}_f & \mathbf{V}_f \end{array}\right]\mathcal{O}_\triangle. \tag{A.53}$$

Hence, it suffices to show that $\mathbf{U}_2^{*e} \xrightarrow{p} \mathcal{O}_\triangle'\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\left[\begin{array}{cc} 1 & \boldsymbol{\mu}_f' \end{array}\right]'$. Rewrite $\mathbf{R}_{[b]}$ in (2.23) using $\mathbf{F}_{\triangle,[b]} = \mathbf{F}_{\triangle,[b]}^*\mathcal{O}_\triangle' + o_p(1)$ as

$$\mathbf{R}_{[b]}^e = \mathbf{B}_{[b]}\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\mathbf{F}_{\triangle,[b]}^{*\prime} + \left(\mathbf{B}_{\triangle,[b]}\boldsymbol{\Lambda}^e\right)o_p(1) + \mathbf{E}_{[b]}. \tag{A.54}$$

Plugging the expression of (2.24), we have

$$\begin{aligned}
&\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\\
=&\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\left(\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\right)'\frac{\mathbf{B}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{1}_\tau}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}\right)\frac{\mathbf{1}_\tau}{\tau}\\
+&\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{1}_\tau}{\tau} + \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\left(\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\right)'\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{1}_\tau}{\tau} + o_p(1),
\end{aligned}$$

yielding

$$\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{1}_\tau}{\tau^2}\right) \tag{A.55}$$
$$=\mathcal{O}_\triangle\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{1}_\tau}{\tau} + \mathcal{E}_{U^e,[b]}^* + o_p(1), \tag{A.56}$$

where

$$\begin{aligned}
\mathcal{E}_{U^e,[b]}^* =& \mathcal{O}_\triangle\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{B}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}} - \mathbf{V}_\beta\right)\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{1}_\tau}{\tau}\\
&+\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{E}_{[b]}}{N_{[b]}} - \widehat{\mathbf{V}}_{e,[b]}^*\right)\frac{\mathbf{1}_\tau}{\tau}\\
&+\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}}{\tau}\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)\boldsymbol{\Lambda}^e\mathcal{O}_\triangle'\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{1}_\tau}{\tau} + \mathcal{O}_\triangle\boldsymbol{\Lambda}^{e\prime}\left(\frac{\mathbf{E}_{[b]}'\mathbf{B}_{[b]}}{N_{[b]}}\right)'\frac{\mathbf{1}_\tau}{\tau} + o_p(1).
\end{aligned}$$

Using the property of $\text{vec}\,(ABC) = (C' \otimes A)\,\text{vec}\,(B)$ and Lemma A.9, we have

$$\begin{aligned}
\mathcal{E}^*_{U^e,[b]} &= \left(\left(\mathbf{\Lambda}^e \mathcal{O}'_\triangle \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{1}_\tau}{\tau}\right)' \otimes \left(\mathcal{O}_\triangle \mathbf{\Lambda}^{e\prime}\right)\right) \text{vec}\left(\frac{\mathbf{B}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}} - \mathbf{V}_\beta\right) \\
&+ \left(\frac{\mathbf{1}'_\tau}{\tau} \otimes \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\right)\right) \mathbf{K}_{[b]} \text{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{E}_{[b]}}{N_{[b]}} - \mathbf{V}_{e,[b]}\right) \\
&+ \left(\left(\mathbf{\Lambda}^e \mathcal{O}'_\triangle \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{1}_\tau}{\tau}\right)' \otimes \left(\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1} \frac{\mathbf{F}^{*\prime}_{\triangle,[b]}}{\tau}\right)\right) \text{vec}\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right) \\
&+ \left(\frac{\mathbf{1}'_\tau}{\tau} \otimes \left(\mathcal{O}_\triangle \mathbf{\Lambda}^{e\prime}\right)'\right) \text{vec}\left(\left(\frac{\mathbf{E}'_{[b]}\mathbf{B}_{[b]}}{N_{[b]}}\right)'\right) + o_p(1),
\end{aligned}$$

implying that

$$\frac{1}{B}\sum_{b=1}^{B} \mathcal{E}_{D^e,[b]} \xrightarrow{p} \mathbf{0}_K$$

from Lemma A.2. Hence, from the expression of (A.55), we have

$$\begin{aligned}
\mathbf{U}_2^{*e} &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{F}^*_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^e_{[b]}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}^*_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right]. \\
&= \mathcal{O}_\triangle \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta \mathbf{\Lambda}^e \mathcal{O}'_\triangle \frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}^{*\prime}_{\triangle,[b]}\mathbf{1}_\tau}{\tau}\right) + \frac{1}{B}\sum_{b=1}^{B}\mathcal{E}^*_{U^e,[b]} + o_p(1) \\
&\xrightarrow{p} \mathcal{O}_\triangle \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta \mathbf{\Lambda}^e \begin{bmatrix} 1 & \boldsymbol{\mu}'_{\boldsymbol{f}} \end{bmatrix}'. \qquad\qquad \text{(A.57)}
\end{aligned}$$

Combining (A.53) and (A.57), we have

$$\mathbf{U}^{*e} = \mathbf{U}_1^{*e}\mathbf{U}_2^{*e} \xrightarrow{p} \mathcal{O}\begin{bmatrix} \boldsymbol{\mu}_f & \mathbf{V}_f \end{bmatrix}\mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^e\begin{bmatrix} 1 & \boldsymbol{\mu}'_f \end{bmatrix}',$$

completing the proof of the lemma. □

Proof of Corollary 2.2 Note that from Lemmas A.10 and A.11

$$(\mathbf{D}^*)^{-1}\mathbf{U}^* \xrightarrow{p} \mathcal{O}'_\triangle\left(\mathbf{V}_{f_\triangle}\mathbf{\Lambda}'\mathbf{V}_{\beta_\triangle}\mathbf{\Lambda}\mathbf{V}_{f_\triangle}\right)^{-1}\mathbf{V}_{f_\triangle}\mathbf{\Lambda}'\left[1\boldsymbol{\mu}'_\beta\right] = \mathcal{O}'_\triangle\boldsymbol{\delta}$$

and that from Lemmas A.12 and A.13

$$
\begin{aligned}
&\left(\mathbf{D}_e^*\right)^{-1}\mathbf{U}_e^* \\
&\xrightarrow{p} -\mathcal{O}'\left(\begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{bmatrix}'\right)^{-1}\begin{bmatrix} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{bmatrix}\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix} 1 & \boldsymbol{\mu}_f' \end{bmatrix}' \\
&= -\mathcal{O}'\left(\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\mathbf{V}_{\beta_\triangle}\boldsymbol{\Lambda}\mathbf{V}_{f_\triangle}\right)^{-1}\mathbf{V}_{f_\triangle}\boldsymbol{\Lambda}'\left[1\boldsymbol{\mu}_\beta'\right] = -\mathcal{O}'\boldsymbol{\delta}^e.
\end{aligned}
$$

Hence, the claim of the corollary is proved. $\square$

Table 1: SDF Estimator Performance when Gross Returns follow CAPM

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500					0.63	0.39	0.23	0.12	0.38	0.61	0.77	0.88
1000	N.A.				0.60	0.39	0.22	0.12	0.40	0.61	0.78	0.88
2000					0.60	0.38	0.22	0.12	0.41	0.63	0.78	0.88
4000					0.65	0.42	0.25	0.13	0.35	0.58	0.75	0.87
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.96	0.97	0.97	0.97	0.62	0.41	0.26	0.16	0.38	0.59	0.74	0.84
1000	0.98	0.98	0.98	0.98	0.60	0.40	0.24	0.14	0.40	0.60	0.76	0.86
2000	0.99	0.99	0.99	0.99	0.60	0.38	0.23	0.13	0.41	0.62	0.77	0.87
4000	0.99	0.99	0.99	1.00	0.65	0.42	0.25	0.14	0.35	0.58	0.75	0.86
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500					-0.05	-0.04	-0.01	0.01	1.07	1.05	1.01	1.00
1000	N.A.				-0.06	-0.02	0.00	-0.01	1.07	1.03	1.00	1.01
2000					-0.05	-0.02	0.00	-0.01	1.07	1.03	1.01	1.01
4000					-0.03	-0.02	0.00	0.00	1.05	1.03	1.00	1.00
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.96	0.97	0.97	0.97	0.04	0.06	0.05	0.05	0.98	0.95	0.96	0.95
1000	0.98	0.98	0.98	0.98	0.02	0.02	0.01	0.02	0.99	0.99	0.99	0.98
2000	0.99	0.99	0.99	0.99	-0.03	0.03	0.01	0.00	1.05	0.98	0.99	1.00
4000	0.99	0.99	0.99	1.00	-0.01	-0.01	-0.01	0.00	1.03	1.02	1.01	1.00
Panel C: Other Estimators												
C-1: Pukthuanthong and Roll's (2017) Estimator												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.04	0.02	0.01	0.00	0.60	0.36	0.19	0.08	0.40	0.64	0.81	0.92
1000	0.06	0.04	0.02	0.01	0.59	0.37	0.20	0.09	0.41	0.63	0.80	0.91
2000	0.09	0.06	0.04	0.02	0.59	0.36	0.21	0.10	0.42	0.64	0.79	0.90
4000	0.11	0.09	0.06	0.04	0.64	0.40	0.23	0.12	0.36	0.60	0.77	0.88
C-2: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta					-0.01	0.00	0.00	0.00	1.03	1.01	1.00	1.01
25 S&B	N.A.				0.31	0.18	0.09	0.04	0.70	0.83	0.91	0.96
25 I&P					0.26	0.15	0.06	0.03	0.75	0.86	0.94	0.97

This table summarizes the performance of various SDF estimators for gross returns when the true return generating process of each individual asset follows CAPM. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 2: SDF Estimator Performance when Excess Returns follow CAPM

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500					0.01	0.00	0.01	0.00	0.98	1.00	0.98	0.99
1000		N.A.			0.01	0.01	0.00	0.00	0.98	0.98	0.99	0.99
2000					-0.02	0.02	0.00	0.00	1.01	0.97	0.99	0.99
4000					0.00	-0.01	-0.01	0.00	0.99	1.00	1.00	0.99
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.96	0.97	0.97	0.97	0.05	0.03	0.05	0.03	0.94	0.96	0.95	0.96
1000	0.98	0.98	0.98	0.98	0.03	0.02	0.02	0.02	0.96	0.97	0.97	0.98
2000	0.99	0.99	0.99	0.99	-0.01	0.03	0.01	0.01	1.00	0.96	0.98	0.99
4000	0.99	0.99	0.99	1.00	0.01	0.00	0.00	0.00	0.98	0.99	0.99	0.99
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500					0.01	0.00	0.01	0.00	0.98	0.99	0.98	0.99
1000		N.A.			0.01	0.01	0.00	0.00	0.98	0.98	0.99	0.99
2000					-0.02	0.02	0.00	0.00	1.01	0.97	0.99	0.99
4000					0.00	-0.01	-0.01	0.00	0.99	1.00	1.00	0.99
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.96	0.97	0.97	0.97	0.05	0.03	0.04	0.03	0.94	0.96	0.95	0.96
1000	0.98	0.98	0.98	0.98	0.03	0.02	0.02	0.02	0.96	0.97	0.97	0.98
2000	0.99	0.99	0.99	0.99	-0.01	0.03	0.01	0.00	1.00	0.96	0.98	0.99
4000	0.99	0.99	0.99	1.00	0.01	0.00	0.00	0.00	0.98	0.99	0.99	0.99
Panel C: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta					-0.02	0.00	-0.01	-0.01	1.02	0.99	1.00	1.00
25 S&B		N.A.			-0.08	-0.02	-0.01	-0.01	1.07	1.02	1.00	1.00
25 I&P					-0.09	-0.03	-0.01	0.00	1.08	1.02	1.00	0.99

This table summarizes the performance of various SDF estimators for excess returns when the true return generating process of each individual asset follows CAPM. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 3: SDF Estimator Performance when Gross Returns follow FF3

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.45	0.61	0.77	0.88	0.57	0.36	0.21	0.11	0.43	0.64	0.79	0.89
1000	0.45	0.62	0.78	0.89	0.58	0.38	0.22	0.12	0.42	0.63	0.78	0.88
2000	0.49	0.65	0.80	0.90	0.54	0.35	0.20	0.11	0.46	0.66	0.80	0.89
4000	0.52	0.67	0.81	0.90	0.50	0.31	0.18	0.09	0.50	0.69	0.82	0.91
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.36	0.52	0.67	0.79	0.62	0.44	0.31	0.22	0.38	0.56	0.69	0.78
1000	0.36	0.53	0.70	0.82	0.63	0.44	0.30	0.19	0.37	0.56	0.70	0.81
2000	0.44	0.60	0.75	0.85	0.57	0.38	0.25	0.16	0.44	0.62	0.75	0.84
4000	0.45	0.62	0.76	0.87	0.54	0.36	0.22	0.13	0.46	0.64	0.78	0.87
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.52	0.45	0.46	0.52	-0.07	-0.16	0.00	-0.04	1.09	1.17	1.00	1.04
1000	0.53	0.49	0.53	0.60	-0.05	-0.03	-0.01	0.00	1.07	1.04	1.02	1.01
2000	0.55	0.54	0.60	0.70	-0.05	-0.04	-0.02	0.00	1.07	1.05	1.03	1.00
4000	0.55	0.58	0.66	0.77	-0.04	-0.03	-0.01	-0.01	1.06	1.04	1.02	1.01
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.42	0.46	0.53	0.64	0.22	0.15	0.12	0.12	0.79	0.85	0.88	0.89
1000	0.43	0.50	0.59	0.72	0.19	0.14	0.12	0.08	0.82	0.87	0.89	0.92
2000	0.50	0.56	0.66	0.77	0.08	0.06	0.06	0.06	0.93	0.95	0.95	0.94
4000	0.49	0.57	0.70	0.82	0.09	0.06	0.05	0.04	0.92	0.95	0.96	0.97
Panel C: Other Estimators												
C-1: Pukthuanthong and Roll's (2017) Estimator												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.05	0.04	0.02	0.00	0.56	0.35	0.19	0.09	0.44	0.65	0.81	0.92
1000	0.08	0.07	0.05	0.03	0.57	0.36	0.21	0.10	0.43	0.64	0.79	0.90
2000	0.13	0.13	0.10	0.06	0.54	0.34	0.19	0.11	0.46	0.66	0.81	0.89
4000	0.18	0.19	0.16	0.12	0.50	0.32	0.18	0.09	0.50	0.69	0.82	0.91
C-2: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta	0.39	0.42	0.46	0.52	0.68	0.55	0.41	0.26	0.33	0.45	0.59	0.74
25 S&B	0.47	0.61	0.75	0.86	0.20	0.11	0.06	0.03	0.81	0.90	0.94	0.97
25 I&P	0.38	0.47	0.58	0.73	0.45	0.32	0.20	0.11	0.55	0.68	0.81	0.89

This table summarizes the performance of various SDF estimators for gross returns when the true return generating process of each individual asset follows FF3. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 4: SDF Estimator Performance when Excess Returns follow FF3

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.56	0.69	0.81	0.90	0.01	0.00	0.00	0.00	0.92	0.94	0.96	0.96
1000	0.56	0.69	0.82	0.90	0.01	0.00	0.01	0.00	0.92	0.94	0.95	0.96
2000	0.57	0.70	0.82	0.90	0.00	0.00	-0.01	0.00	0.93	0.95	0.96	0.95
4000	0.56	0.70	0.82	0.90	0.00	0.00	0.00	0.00	0.93	0.95	0.95	0.96
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.44	0.58	0.71	0.80	0.18	0.13	0.11	0.10	0.75	0.82	0.85	0.86
1000	0.45	0.59	0.73	0.83	0.17	0.12	0.10	0.07	0.77	0.83	0.86	0.89
2000	0.51	0.65	0.77	0.86	0.08	0.06	0.05	0.05	0.85	0.89	0.91	0.91
4000	0.50	0.65	0.78	0.87	0.10	0.06	0.04	0.03	0.84	0.89	0.91	0.93
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.52	0.65	0.78	0.88	0.01	0.00	0.00	0.00	0.92	0.95	0.96	0.96
1000	0.53	0.67	0.80	0.89	0.01	0.00	0.01	0.00	0.92	0.94	0.95	0.96
2000	0.56	0.69	0.81	0.90	0.00	0.00	-0.01	0.00	0.93	0.95	0.96	0.95
4000	0.55	0.68	0.81	0.90	0.00	0.00	0.00	0.00	0.93	0.95	0.95	0.96
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.44	0.57	0.70	0.79	0.17	0.13	0.11	0.11	0.76	0.82	0.85	0.86
1000	0.46	0.59	0.73	0.83	0.15	0.12	0.09	0.07	0.78	0.83	0.87	0.89
2000	0.51	0.64	0.77	0.86	0.08	0.07	0.05	0.04	0.85	0.88	0.90	0.92
4000	0.50	0.64	0.78	0.87	0.09	0.06	0.04	0.04	0.84	0.89	0.91	0.92
Panel C: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta	0.50	0.56	0.62	0.69	-0.05	-0.02	0.00	-0.01	0.97	0.96	0.95	0.97
25 S&B	0.55	0.69	0.81	0.90	-0.15	-0.05	-0.01	-0.01	1.07	0.99	0.97	0.97
25 I&P	0.51	0.64	0.77	0.87	-0.14	-0.04	-0.02	-0.01	1.06	0.99	0.97	0.97

This table summarizes the performance of various SDF estimators for excess returns when the true return generating process of each individual asset follows FF3. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 5: SDF Estimator Performance when Gross Returns follow FF5

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.51	0.69	0.82	0.90	0.45	0.28	0.16	0.09	0.55	0.72	0.84	0.92
1000	0.52	0.69	0.83	0.91	0.49	0.32	0.19	0.10	0.52	0.68	0.81	0.90
2000	0.55	0.72	0.84	0.92	0.47	0.30	0.17	0.09	0.53	0.70	0.83	0.91
4000	0.59	0.74	0.85	0.92	0.40	0.24	0.14	0.07	0.60	0.76	0.86	0.93
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.38	0.55	0.68	0.77	0.54	0.40	0.29	0.22	0.46	0.60	0.71	0.78
1000	0.41	0.56	0.67	0.75	0.54	0.41	0.31	0.26	0.46	0.59	0.69	0.74
2000	0.47	0.64	0.77	0.85	0.51	0.35	0.23	0.16	0.49	0.65	0.77	0.84
4000	0.51	0.66	0.77	0.84	0.45	0.31	0.21	0.15	0.56	0.70	0.79	0.84
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.47	0.26	0.29	0.37	-0.09	-0.09	-0.01	-0.03	1.11	1.10	1.02	1.03
1000	0.51	0.34	0.39	0.49	-0.07	-0.01	-0.03	-0.01	1.09	1.02	1.03	1.02
2000	0.54	0.44	0.51	0.64	-0.05	-0.02	-0.02	-0.01	1.07	1.03	1.02	1.01
4000	0.54	0.53	0.63	0.75	-0.05	-0.02	-0.01	-0.01	1.06	1.03	1.02	1.01
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.39	0.35	0.43	0.55	0.23	0.19	0.17	0.16	0.78	0.82	0.83	0.85
1000	0.43	0.41	0.50	0.61	0.18	0.18	0.17	0.18	0.84	0.83	0.83	0.82
2000	0.47	0.49	0.61	0.73	0.11	0.09	0.08	0.08	0.91	0.92	0.92	0.92
4000	0.48	0.52	0.64	0.75	0.09	0.09	0.09	0.09	0.92	0.92	0.91	0.91
Panel C: Other Estimators												
C-1: Pukthuanthong and Roll's (2017) Estimator												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.12	0.09	0.05	0.00	0.45	0.28	0.16	0.08	0.55	0.72	0.84	0.92
1000	0.16	0.15	0.11	0.06	0.48	0.31	0.18	0.09	0.52	0.69	0.82	0.91
2000	0.24	0.24	0.20	0.13	0.47	0.30	0.17	0.09	0.53	0.70	0.83	0.91
4000	0.32	0.34	0.30	0.23	0.42	0.26	0.16	0.09	0.58	0.74	0.84	0.91
C-2: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta	0.24	0.25	0.29	0.35	0.68	0.55	0.35	0.24	0.32	0.45	0.65	0.76
25 S&B	0.33	0.47	0.63	0.78	0.38	0.25	0.14	0.08	0.62	0.76	0.86	0.93
25 I&P	0.31	0.43	0.59	0.74	0.41	0.27	0.16	0.09	0.60	0.73	0.84	0.91

This table summarizes the performance of various SDF estimators for gross returns when the true return generating process of each individual asset follows FF5. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 6: SDF Estimator Performance when Excess Returns follow FF5

	R^2				intercept(a)				slope(b)			
Panel A: Balanced Panel Estimator												
A-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.55	0.71	0.83	0.91	0.00	0.00	0.00	0.00	0.84	0.88	0.89	0.90
1000	0.56	0.72	0.84	0.91	0.00	0.00	0.00	0.00	0.85	0.87	0.89	0.90
2000	0.57	0.73	0.84	0.92	0.00	0.00	0.00	0.00	0.85	0.88	0.89	0.90
4000	0.57	0.73	0.84	0.91	0.00	0.00	0.00	0.00	0.85	0.88	0.89	0.90
A-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.40	0.56	0.69	0.77	0.20	0.17	0.15	0.13	0.66	0.72	0.76	0.78
1000	0.44	0.58	0.68	0.75	0.16	0.16	0.16	0.16	0.70	0.73	0.75	0.76
2000	0.48	0.64	0.77	0.85	0.11	0.08	0.07	0.06	0.75	0.80	0.83	0.84
4000	0.49	0.65	0.76	0.84	0.09	0.08	0.08	0.08	0.76	0.80	0.82	0.83
Panel B: Unbalanced Panel Estimator												
B-1: With Observed Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.48	0.65	0.79	0.88	0.00	0.00	0.00	0.00	0.84	0.87	0.89	0.90
1000	0.52	0.68	0.81	0.90	-0.01	0.00	0.00	0.00	0.85	0.88	0.89	0.90
2000	0.54	0.70	0.83	0.91	0.00	0.00	-0.01	0.00	0.85	0.88	0.90	0.90
4000	0.55	0.71	0.83	0.91	-0.01	0.00	0.00	0.00	0.85	0.88	0.89	0.90
B-2: With Estimated Factors												
$N\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
500	0.40	0.55	0.68	0.76	0.20	0.17	0.15	0.14	0.66	0.72	0.76	0.78
1000	0.44	0.57	0.68	0.74	0.16	0.15	0.15	0.16	0.70	0.74	0.75	0.76
2000	0.48	0.63	0.76	0.84	0.10	0.09	0.07	0.07	0.75	0.80	0.83	0.84
4000	0.49	0.64	0.76	0.83	0.09	0.08	0.08	0.07	0.76	0.80	0.82	0.83
Panel C: GMM Estimator												
Pfo$\backslash T$	60	120	240	480	60	120	240	480	60	120	240	480
10 Beta	0.31	0.36	0.44	0.56	-0.05	-0.01	-0.01	-0.01	0.89	0.88	0.90	0.91
25 S&B	0.47	0.61	0.74	0.83	-0.18	-0.07	-0.03	-0.01	1.00	0.94	0.92	0.91
25 I&P	0.45	0.59	0.73	0.83	-0.18	-0.07	-0.03	-0.01	1.00	0.93	0.92	0.91

This table summarizes the performance of various SDF estimators for excess returns when the true return generating process of each individual asset follows FF5. We consider different levels of $N = 500,\ 1000,\ 2000$ and 4000 and $T = 60,\ 120,\ 240$ and 480. After obtaining a time series of estimates $\widehat{m}_t$ for $t = 1, \cdots, T$, we regress the estimated SDF $\widehat{m}$ on a constant and the true SDF m: $\widehat{m}_t = a + b \cdot m_t + error_t$. If the fit to the true SDF is perfect, R^2 is 1, the intercept (a) is zero and the coefficient on the true SDF (b) is 1. We report the mean of the estimated R^2, a, and b across 10,000 repetitions.

Table 7: Factor Selection Performance using Lasso

Panel A: RGP follows CAPM					
	MKT	SMB	HML		
Gross return case	99.01	8.47	7.93		
Excess return case	99.20	0.07	0.59		
Panel B: RGP follows FF3					
	MKT	SMB	HML	RMA	CMW
Gross return case	98.95	94.79	99.54	12.85	12.42
Excess return case	99.44	93.33	98.33	0.01	0

This table reports the probability of a given factor to be selected in the estimated stochastic discount factor. In Panel A (B), we simulate the returns of assets in the economy using the calibrated parameters of CAPM (FF3) and allow the stochastic discount factors to be a linear function of the three factors of MKT, SMB, and HML (the five factors of MKT, SMB, HML, RMA, and CMW). We set $N = 2000$ and $T = 480$. Reported figures are the empirical frequencies (in percentage) of a given factor to have a non-zero loading in the estimated SDF with Lasso penalty, computed from 10,000 repetitions.

Table 8: SDF Estimates for the Balanced Panel of 1200 Portfolios sorted on Expected Returns

Model/Methods	Factors								
	Panel A: Specific Asset Pricing Models								
	MKT	SMB	HML	I/A	ROE	CMA	RMW	MOM	HML (devil)
CAPM	-8.07								
	(0.08)								
FF3	-5.81	-5.27	-1.79						
	(0.25)	(0.58)	(0.82)						
Lasso	-6.08	-4.49	0.00						
	(0.24)	(0.51)							
HXZ4	-6.63	-7.48		-4.00	-5.54				
	(0.28)	(0.59)		(1.17)	(0.91)				
Lasso	-6.52	-7.12		0.00	-6.04				
	(0.28)	(0.60)			(0.90)				
FF5	-7.94	-7.41	7.66			-6.97	-12.17		
	(0.37)	(0.85)	(1.27)			(1.23)	(1.62)		
Lasso	-6.19	-7.72	0.00			-3.49	-4.18		
	(0.31)	(0.92)				(1.01)	(1.27)		
BS6	-6.53	-9.79		-4.23	-8.60			-0.36	-2.34
	(0.33)	(0.80)		(1.75)	(1.20)			(0.88)	(1.26)
Lasso	-6.44	-9.08		0.00	-9.63			1.61	0.00
	(0.31)	(0.82)			(1.25)			(0.49)	
	Panel B: Statistical Factor Models								
	PC1	PC2	PC3	PC4					
APC (EM)	-7.61	-0.73	-3.37	-0.87					
	(0.82)	(0.54)	(0.21)	(0.39)					
RP-PCA	-21.54	-17.01	-3.26	-0.32					
	(3.21)	(3.43)	(0.46)	(0.44)					
Lasso	-21.36	-16.85	-3.28	0.00					
	(3.05)	(3.27)	(0.45)						

This table reports the estimated values of $\boldsymbol{\delta}^e$, when we apply our balanced panel estimator using 1200 portfolios sorted by expected returns. We estimate $\boldsymbol{\delta}^e$ with $\widetilde{\boldsymbol{\delta}}^e$ in (2.14) and $\widetilde{\boldsymbol{\delta}}^e_{lasso}$ in (2.18). In panel A, we consider six asset pricing models: CAPM (Sharpe (1964) and Lintner (1965)), FF3 (Fama and French (1992)), HXZ4 (Hou, Xue and Zhang (2015)), FF5 (Fama and French (2015)), BS6 (Barillas and Shanken (2017)). In Panel B, we examine two versions of statistical factors. First, we apply APC by Connor and Korajczyk (1986) to excess returns of a large cross section of individual stocks from CRSP. Second, we utilize the method by Pelger and Lettau (2017) to 209 portfolios (16 sets of characteristic-based decile portfolios plus 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. We report the Lasso factor selection results only when a smaller set of factors are selected. The sample periods are 293 months over the period January 1990 to May 2014.

Table 9: SDF Estimates for the Balanced Panel of 209 Portfolios

Model/Methods	Factors								
	Panel A: Specific Asset Pricing Models								
	MKT	SMB	HML	I/A	ROE	CMA	RMW	MOM	HML (devil)
CAPM	-2.40								
	(0.13)								
FF3	-3.50	1.04	-6.44						
	(0.33)	(1.92)	(2.19)						
Lasso	-3.27	0.00	-6.76						
	(0.16)		(1.78)						
HXZ4	-4.11	-6.00		-13.34	-11.61				
	(0.22)	(3.13)		(3.14)	(3.26)				
Lasso	-4.82	0.00		-12.31	-6.22				
	(0.32)			(2.85)	(1.02)				
FF5	-4.90	-0.60	3.52			-2.66	-18.21		
	(0.54)	(1.36)	(5.18)			(5.64)	(5.70)		
Lasso	-3.49	0.00	-6.05			-2.87	0.00		
	(0.25)		(2.78)			(4.55)	(0.00)		
BS6	-3.85	-6.31		1.62	-13.43			-7.91	-11.48
	(0.54)	(3.77)		(7.88)	(5.60)			(4.26)	(6.84)
Lasso	-4.84	0.00		0.00	-4.59			-9.30	-9.57
	(0.50)				(1.81)			(3.81)	(2.77)
	Panel B: Statistical Factor Models								
	PC1	PC2	PC3	PC4					
APC (EM)	-3.04	-3.59	-4.15	-11.57					
	(0.28)	(1.63)	(1.04)	(6.97)					
RP-PCA	-10.21	-8.43	-2.08	-0.37					
	(5.24)	(5.79)	(0.95)	(0.26)					

This table reports the estimated values of $\boldsymbol{\delta}^e$, when we apply our balanced panel estimator using 209 portfolios (16 sets of characteristic-based decile portfolios plus 49 industry portfolios). We estimate $\boldsymbol{\delta}^e$ with $\widetilde{\boldsymbol{\delta}}^e$ in (2.14) and $\widetilde{\boldsymbol{\delta}}^e_{lasso}$ in (2.18). In panel A, we consider six asset pricing models: CAPM (Sharpe (1964) and Lintner (1965)), FF3 (Fama and French (1992)), HXZ4 (Hou, Xue and Zhang (2015)), FF5 (Fama and French (2015)), BS6 (Barillas and Shanken (2017)). In Panel B, we examine two versions of statistical factors. First, we apply APC by Connor and Korajczyk (1986) to excess returns of a large cross section of individual stocks from CRSP. Second, we utilize the method by Pelger and Lettau (2017) to 209 portfolios (16 sets of characteristic-based decile portfolios plus 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. We report the Lasso estimation results only when a smaller set of factors are selected. The sample periods are 600 months over the sample period January 1976 to December 2016.

Table 10: SDF Estimates for the Unbalanced Panel of All Individual Stock Returns

Model/Methods	Factors								
	Panel A: Specific Asset Pricing Models								
	MKT	SMB	HML	I/A	ROE	CMA	RMW	MOM	HML (devil)
CAPM	-4.47								
	(0.04)								
FF3	-3.69	-2.51	-1.15						
	(0.11)	(0.26)	(0.27)						
HXZ4	-4.95	-6.35		-7.51	-14.87				
	(0.16)	(0.37)		(0.63)	(0.50)				
FF5	-5.01	-4.28	3.94			-10.65	-8.51		
	(0.22)	(0.37)	(0.66)			(0.99)	(1.26)		
BS6	-5.08	-6.90		-1.58	-14.72			-5.94	-6.75
	(0.18)	(0.43)		(1.36)	(0.96)			(0.68)	(0.81)
	Panel B: Statistical Factor Models								
	PC1	PC2	PC3	PC4					
APC (EM)	-1.82	-3.14	-1.61	-6.39					
	(0.23)	(0.28)	(0.69)	(1.34)					
RP-PCA	-10.61	-7.88	-0.28	-0.44					
	(1.09)	(1.23)	(0.25)	(0.24)					

This table reports the estimated values of $\boldsymbol{\delta}^e$, when we apply our balanced panel estimator using all individual stocks in CRSP. We estimate $\boldsymbol{\delta}^e$ with $\widehat{\boldsymbol{\delta}}^e$ in (2.27). In panel A, we consider six asset pricing models: CAPM (Sharpe (1964) and Lintner (1965)), FF3 (Fama and French (1992)), HXZ4 (Hou, Xue and Zhang (2015)), FF5 (Fama and French (2015)), BS6 (Barillas and Shanken (2017)). In Panel B, we examine two versions of statistical factors. First, we apply APC by Connor and Korajczyk (1986) to excess returns of a large cross section of individual stocks from CRSP. Second, we utilize the method by Pelger and Lettau (2017) to 209 portfolios (16 sets of characteristic-based decile portfolios plus 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 600 months over the sample period January 1976 to December 2016.

Table 11: SDF Estimates for the Unbalanced Panel of NYSE Individual Stocks

Model/Methods	Factors								
	Panel A: Specific Asset Pricing Models								
	MKT	SMB	HML	I/A	ROE	CMA	RMW	MOM	HML (devil)
CAPM	-3.76								
	(0.05)								
FF3	-4.06	0.53	-1.25						
	(0.10)	(0.28)	(0.34)						
HXZ4	-5.08	-5.11		-10.28	-16.13				
	(0.15)	(0.42)		(0.81)	(0.60)				
FF5	-4.68	-1.39	0.93			-9.58	-4.39		
	(0.17)	(0.38)	(0.63)			(0.96)	(1.18)		
BS6	-4.78	-5.85		1.02	-17.05			-7.30	-11.56
	(0.18)	(0.52)		(1.53)	(1.12)			(0.72)	(0.95)
	Panel B: Statistical Factor Models								
	PC1	PC2	PC3	PC4					
APC (EM)	-2.82	-2.55	-3.70	-0.13					
	(0.18)	(0.29)	(1.04)	(0.77)					
RP-PCA	-13.46	-11.28	-0.95	-0.09					
	(1.37)	(1.53)	(0.39)	(0.22)					

This table reports the estimated values of $\boldsymbol{\delta}^e$, when we apply our balanced panel estimator using NYSE individual stocks. We estimate $\boldsymbol{\delta}^e$ with $\widehat{\boldsymbol{\delta}}^e$ in (2.27). In panel A, we consider six asset pricing models: CAPM (Sharpe (1964) and Lintner (1965)), FF3 (Fama and French (1992)), HXZ4 (Hou, Xue and Zhang (2015)), FF5 (Fama and French (2015)), BS6 (Barillas and Shanken (2017)). In Panel B, we examine two versions of statistical factors. First, we apply APC by Connor and Korajczyk (1986) to excess returns of a large cross section of individual stocks from CRSP. Second, we utilize the method by Pelger and Lettau (2017) to 209 portfolios (16 sets of characteristic-based decile portfolios plus 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 600 months over the sample period January 1976 to December 2016.

References

[1] Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X., 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259-299.

[2] Araujo, Fabio and Joao Victor Issler, 2012, A stochastic discount factor approach to asset pricing using panel data asymptotics, Working paper, Princeton University.

[3] Asness, C., Frazzini, A., and Pedersen, L. H., 2014, Quality minus junk, Working paper, AQR Capital Management.

[4] Bai, J., and S. Ng, 2002, Determining the number of factors in approximate factor models, *Econometrica* 70, 191-221.

[5] Barillas, F., and Jay A. Shanken, 2017, Comparing asset pricing models, *Journal of Finance*, forthcoming.

[6] Basu, S., 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129-156.

[7] Black, F., M. C. Jensen, and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, in M. C. Jensen (ed.), Studies in the Theory of Capital Markets, NY, Praeger Publishers.

[8] Campbell, J. Y., and Vuolteenaho, T., 2004, Bad beta, good beta, *The American Economic Review* 94, 1249-1275.

[9] Chamberlain, G., and Rothschild, M., 1983, Arbitrage, factor structure and mean-variance analysis in large asset markets, *Econometrica* 51, 1305-1324.

[10] Chan, L. K. C., Y. Hamao, and J. Lakonishok, 1991, Fundamentals and stock returns in Japan. *Journal of Finance* 46, 1739–1764.

[11] Chen, Z., Connor, G., and Korajczyk, R. A., 2018, A performance comparison of large-n factor estimators, *Review of Asset Pricing Studies* 8, forthcoming, https://ssrn.com/abstract=2380144.

[12] Chen, L., Novy-Marx, R., and Zhang, L., 2011, An alternative three-factor model, Working paper, University of Rochester.

[13] Chen, N.-F., Roll, R., and Ross, S. A., 1986, Economic forces and the stock market, *Journal of Business* 59, 383-403.

[14] Cochrane, J. H., 1996, 6, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572-62

[15] Cochrane, J. H., 2005, Asset Pricing, Revised edition, Princeton, NJ, Princeton University Press.

[16] Connor, G., 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13-31.

[17] Connor, G., and Korajczyk, R. A., 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics* 15, 373-394.

[18] ——, 1987, Estimating pervasive economic factors with missing observations, Working paper, Northwestern University, https://ssrn.com/abstract=1268954.

[19] ——, 1988, Risk and return in an equilibrium APT: Application of a new test methodology, *Journal of Financial Economics* 21, 255-289.

[20] Daniel, K., and S. Titman, 2012, Testing factor-model explanations of market anomalies, *Critical Finance Review*, 1, 103–139.

[21] De Bondt, W. F. M., and Thaler, R. 1985, Does the stock market overreact? *Journal of Finance* 40, 793-805.

[22] Fama, E. F., and French, K. R., 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

[23] ——, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56

[24] ——, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

[25] ——, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 3-56

[26] Frazzini, A., and Pedersen, L. H., 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

[27] Freyberger, J., A. Neuhierl, and M. Weber, 2017, Dissecting characteristics nonparametrically, Working paper.

[28] Gagliardini, P., E. Ossola, and O. Scaillet, 2016, Time-varying risk premium in large cross-sectional equity datasets, *Econometrica* 84, 985-1046.

[29] Hansen, L. P., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, 50, 1029-1054.

[30] Hansen, L. P., and Jagannathan, R., 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225-262.

[31] ——, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance*, 52, 557–590.

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

[33] Huberman, G., and Kandel, S., 1987, Mean-variance spanning, *Journal of Finance* 42, 873-888.

[34] Jagannathan, R., and Marakani, S., 2015, Price-dividend ratio factor proxies for long-run risks, *Review of Asset Pricing Studies* 5, 1-47.

[35] Jagannathan, R., and Wang, Z., 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3-53.

[36] ——, 1998, An asymptotic theory for estimating beta-pricing Models using cross-sectional regression, *Journal of Finance* 53, 1285–1309.

[37] Jagannathan, R., G. Skoulakis, and Z. Wang, 2010, The analysis of the cross section of security returns, in Handbook of Financial Econometrics 2, ed. by Y. Ait-Sahalia, and L. P. Hansen, UK, Elsevier.

[38] Jegadeesh, N., 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881-898.

[39] Jegadeesh, N., and Titman, S., 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

[40] Jones, C. S., 2001, Extracting factors from heteroskedastic asset returns, *Journal of Financial Economics* 62, 293–325.

[41] Joreskog, K., 1967, Some contributions to maximum likelihood factor analysis, *Psychometrica* 34, 183–202.

[42] Ikenberry, D., Lakonishok, J., and Vermaelen, T., 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181-208.

[43] Kelly, B., and Pruitt, S., 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721–1756.

[44] Kim, S., and Skoulakis, G., 2017, Testing ex-post implications of asset pricing models using individual stocks, Working paper.

[45] ——, 2018, Ex-post risk premia estimation and asset pricing tests using large cross sections: The regression-calibration approach, *Journal of Econometrics*, forthcoming

[46] Kozak, S., S. Nagel, and S. Santosh, 2017, Shrinking the cross section, Working paper, University of Michigan.

[47] Lehmann, B. N., and Modest, D. M., 1988, The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics* 21, 213–254.

[48] Lewellen, J., S. Nagel, and J. Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics*, 96, 175–194.

[49] Lintner, J., 1965, Security prices, risk, and maximal gains from diversification, *Journal of Finance*, 20, 587–615.

[50] Litzenberger, R. H., and K. Ramaswamy, 1979, The effect of personal taxes and dividends of capital asset prices: The theory and evidence, *Journal of Financial Economics*, 7, 163–196.

[51] ——, 1982, The effects of dividends on common stock prices: Tax effects or information effects? *Journal of Finance* 37, 429-443.

[52] Lo, A. W., and A. C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies*, 3, 431–467.

[53] Merton, R. C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867-887

[54] Miller, H., and M. Scholes, 1972, Rate of return in relation to risk: A reexamination of some recent findings, in M. C. Jensen (ed.), Studies in the Theory of Capital Markets, NY, Praeger Publishers.

[55] Pelger, M., and Martin Lettau, 2017, Estimating asset pricing factors from large-dimensional panel data, Working paper.

[56] Pukthuanthong, K., and Roll, R., 2017, An agnostic and practically useful estimator of the stochastic discount factor, Working paper, Caltech.

[57] Roll, R., 1977, A critique of the asset pricing theory's tests part I: On past and potential testability of the theory, *Journal of Financial Economics*, 4, 129–176.

[58] Ross, S., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

[59] Shanken, J., 1992, On the estimation of beta-pricing models, *Review of Financial Studies*, 5, 1–33

[60] Shanken, J., and Weinstein, M. I., 2006, Economic forces and the stock market revisited, *Journal of Empirical Finance* 13, 129-144.

[61] Sharpe, W. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance*, 19, 425–442.

[62] Sloan, R. G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71, 289-315.

[63] Stock, J. H., and Watson, M. W., 1998, Diffusion indexes, Working paper, NBER, http://www.nber.org/papers/w6702.

[64] ——, 2002, Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association* 97, 1-13.

[65] Stroyny, Alvin L, 1992, Still more on EM factor analysis, Working paper, University of Wisconsin.