Large Sample Estimators of the Stochastic Discount Factor*

Soohun Kim† Robert A. Korajczyk‡

November 9, 2017

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 chosen by a specific asset pricing model but a set of agnostic factors estimated by 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 panel of individual stock returns. We apply our estimators to large cross section of real financial data and provide novel evidence on the premia of frequently used factors.

JEL classification codes: G12

Keywords: asset pricing, factor structure, stochastic discount factor

*We thank Christian Brownlees, Ravi Jagannathan, Andreas Nuehierl, Kuntara Pukthuanthong, Richard Roll, and participants at the 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 arbitrages, there exists a valid stochastic discount factor such that the price of any security is obtained as the expected value of the discounted future payoff. This equilibrium relationship provides theoretical backgrounds for empirical researchers to study asset pricing models with the convenient tools of generalized methods of moments. It is not surprising that the stochastic discount factor (SDF) representation of asset pricing models has been widely used in the empirical finance literature. What is somewhat unnatural is that the most empirical studies using SDF representation are applied to the equilibrium relations of only a handful of assets. This practice can be traced back to the technical reason that the conventional generalized methods of moments are developed for a finite number of moment conditions or to the historical reason that empirical asset pricing studies had been commonly applied to the portfolio levels to mitigate the beta measurement errors since the early literature (Miller and Scholes (1972), Black, Jensen, and Scholes (1973)).

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 balanced or unbalanced financial data. The choice of test subjects is essential in the design of any empirical studies including ones on asset prices. Lewellen, Nagel, and Shanken (2010) and Daniel and Titman (2012) convincingly demonstrate that how well an asset pricing model fits varies significantly across different sets of test portfolios. A particular set of portfolios can be too kind to a specific model if, in the process of forming portfolios, important information against the model is averaged out within portfolios (Rolls (1977)). Another set of portfolios can be too harsh to the model if stocks are grouped into portfolios with respect to some attributes adverse to the model (Lo and MacKinlay (1990)). Both biases can misguide empirical researchers to wrong conclusions. We offer novel methodologies not subject to these biases.

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 a mild assumption that individual assets have an approximate factor structure as in Chamberlain and Rothschild (1983). For the case of balanced panel data, the minimization problem is reduced to a linear regression problem and the solution can be easily obtained. In handling 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 short time bias correction similar to

Litzenberger and Ramaswamy (1992) and Kim and Skoulakis (2017), our estimators can consistently estimate the true SDF. Besides, we find that our estimators are robust to the latency of factors, being partially agnostic in the spirit of Pukthuanthong and Roll (2016), 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 SDF can be summarized by the cross-sectional average of log returns with the assumption of a single factor structure. We allow multiple pervasive risks in an economy and let our estimators find the priced factors among those by observing prices dynamics of the large panel. In particular, our paper is closely related to Pukthuanthong and Roll (2016). 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 representations. However, 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 simulated 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 model economy leads to significant improvement in the performance of the estimated SDF with empirically relevant panel sizes. Kozak, Nagel, and Santosh (2017) propose a novel method of estimating SDF using individual stocks through a vast array of characteristics. However, they do not use the pricing restriction of large panel directly but take the detour of maximizing Sharpe ratio of a characteristic based portfolios resorting to Hansen and Jagannathan (1991).

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) provide a framework to dichotomize a large cross section of returns into pervasive factors and diversifiable risks. A long literature initiated by Connor and Korajczyk (1986) propose to methodology to extract pervasive factors from a large cross sectional data (Connor and Korajczyk (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 from a large cross sectional data. As pointed out in Merton (1973), Jagannathan and Wang (1996), Campbell and Vuolteenaho (2004), Kelly and Pruitt (2013) and Jagannathan and Marakani (2015), all pervasive factors need not be important for explaining the cross section of asset prices. Alternatively, whether a pervasive factor is priced or not can be examined with the beta pricing form. Exploiting a large cross section of beta pricing equation, a series of papers have proposed risk premia estimators using large cross sectional data (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 (2017a, 2017b) obtain the large panel asymptotic distribution of the risk premia estimator along with an estimator of its variance-covariance matrix. This paper is differentiated in that we are using the SDF representation, not beta pricing representation. This difference is particularly important when we use large panel data where the measurement errors in the beta can severely distort the empirical results. Although the equivalence between SDF form and beta form is well known in the small N and large T setup (Jagannathan and Wang (2002)), more studies need to be done to understand the differences between two approaches in the large panel data.

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. In Section 4, we apply our SDF estimators into large cross section of financial data and provide an evidence of premia of candidate factors. 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 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 set $\mathbb{E}[e_{i,t}] = 0$ and $\mathbb{E}[\mathbf{f}_t e_{i,t}] = \mathbf{0}_K$, a the $(K \times 1)$ 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 exist a scalar λ_0, the gross return on the riskless asset, and a $(K \times 1)$ vector of $\boldsymbol{\lambda}_f$ such that

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

We assume that exact factor pricing holds, so that the equation (2.2) holds as an equality. Let the $(K \times 1)$ vector of $\boldsymbol{\mu}_f$ is defined by $\boldsymbol{\mu}_f = \mathbb{E}\left[\mathbf{f}_t\right]$. By combining the return generating process of (2.1) and 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 $\boldsymbol{\mu}_{\boldsymbol{f}} = \mathbf{0}$. In the literature, there are a number of papers which use a combination of traded excess returns and pre-whitened macroeconomic variables, as in 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 $\boldsymbol{\mu}_{\boldsymbol{f}} = \boldsymbol{\lambda}_{\boldsymbol{f}}$, but the procedure requires a 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 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 show 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, of 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 balance panel of asset returns and then extent 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 all individual assets $i = 1, \cdots, N$ and 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. It needs to be noted that 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 assume that the residual terms of $e_{i,t}$ are not too correlated in both cross-sectional and time-series dimensions. 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 (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 residual and the product of residual 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* $\widetilde{m}^e_t = 1 + \mathbf{f}'_t \widetilde{\boldsymbol{\delta}}^e$ *converge to* m_t *and* m^e_t *given in* (2.5) *and* (2.9), *respectively, when the* $((K+1)\times 1)$ *vector of* $\widetilde{\boldsymbol{\delta}} = \left[\widetilde{\delta}_0 \ \widetilde{\boldsymbol{\delta}}'_f\right]'$ *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 $\mathbf{F}_{\triangle} = [\mathbf{1}_T \ \mathbf{F}]$.

–

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 $\mathbf{F}^* = \mathbf{F}\mathcal{O} + o_p(1)$. 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* $N, T \to \infty$, $\widetilde{m}^*_t = \widetilde{\delta}^*_0 + \mathbf{f}^{*\prime}_t \widetilde{\boldsymbol{\delta}}^*_f$ *and* $\widetilde{m}^{*e}_t = 1 + \mathbf{f}^{*\prime}_t \widetilde{\boldsymbol{\delta}}^{*e}$ *converge to* m_t *and* m^e_t *given in* (2.5) *and* (2.9), *respectively, when the*

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

$$\widetilde{\boldsymbol{\delta}}^* = \left(\frac{\mathbf{F}_\triangle^{*\prime}\mathbf{R}'\mathbf{R}\mathbf{F}_\triangle^*}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}_\triangle^{*\prime}\mathbf{R}'\mathbf{1}_N}{NT}\right) \tag{2.15}$$

$$\widetilde{\boldsymbol{\delta}}^{*e} = -\left(\frac{\mathbf{F}^{*\prime}\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{F}^*}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}^{*\prime}\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{1}_T}{NT^2}\right), \tag{2.16}$$

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

Furthermore, even for the case that we do not have a consistent estimator of factor realization, 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 (2016).

Proposition 2.1. *With Assumptions 1 and 2, the homoscedasticity 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*

$$\breve{m}_t = \iota_t'\left(\frac{\mathbf{R}'\mathbf{R}}{NT^2}\right)^{-1}\left(\frac{\mathbf{R}'\mathbf{1}_N}{NT}\right), \tag{2.17}$$

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

The estimator proposed in Theorem 2.1 can be intuitively understood as follows. By specifying the $(T \times 1)$ vector of 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:

$$\widetilde{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}}\left(\mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_\triangle}{T}\boldsymbol{\delta}\right)'\left(\mathbf{1}_N - \frac{\mathbf{R}\mathbf{F}_\triangle}{T}\boldsymbol{\delta}\right). \tag{2.18}$$

Similarly, given the $(T \times 1)$ vector of 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:

$$\widetilde{\boldsymbol{\delta}}^e = \arg\min_{\boldsymbol{\delta}^e} \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T} \boldsymbol{\delta}^e \right)' \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T} \boldsymbol{\delta}^e \right). \tag{2.19}$$

The formation of $\widetilde{\boldsymbol{\delta}}$ and $\widetilde{\boldsymbol{\delta}}^e$ as in (2.18) and (2.19) 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 to select important factors which are useful in minimizing pricing errors. For example, we can increase the factors into L factors where $L \geq K$ and then let the following minimizer to select the true K factors:

$$\widetilde{\boldsymbol{\delta}}_{lasso} = \arg\min_{\delta_0, \mathbf{d}} \left(\mathbf{1}_N - \frac{\mathbf{RF}_\triangle}{T} \boldsymbol{\delta} \right)' \left(\mathbf{1}_N - \frac{\mathbf{RF}_\triangle}{T} \boldsymbol{\delta} \right) + \gamma \|\mathbf{d}\|_1, \tag{2.20}$$

where $\gamma > 0$, $\mathbf{d} = [d_1 \cdots d_L]'$, $\|\mathbf{d}\|_1 = |d_1| + \cdots + |d_L|$, $\mathcal{F}$ is the $(T \times L)$ matrix such that $\mathbf{F} = \mathcal{F}\mathcal{S}$, $\boldsymbol{\delta}_f = \mathcal{S}'\mathbf{d}$, $\boldsymbol{\delta} = [\delta_0 \ \boldsymbol{\delta}_f]'$, and the $(L \times K)$ selection matrix of $\mathcal{S}$ is defined by as follows. The (l, k) element of $\mathcal{S}$ is 1 if d_l is the k-th non zero elements among $d_1, \cdots, d_l$ 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

$$\widetilde{\boldsymbol{\delta}}^e_{lasso} = \arg\min_{\mathbf{d}} \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T} \boldsymbol{\delta}^e \right)' \left(\frac{\mathbf{R}^e \mathbf{1}_N}{T} - \frac{\mathbf{R}^e \mathbf{F}}{T} \boldsymbol{\delta}^e \right) + \gamma \|\mathbf{d}\|_1, \tag{2.21}$$

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

2.2 Unbalanced Panel Estimator

Since the estimators proposed in the previous subsection utilize the full panel data of large N, T, they are appropriate for the case using a large number of portfolios over a long horizon. However, if empirical researchers are interested in using individual stocks as test assets, the estimators are not applicable due to the limitation of the unbalanced return data. In this section, we propose estimators which can deal with unbalanced panel data. First, 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. Then, relabel those stocks 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 $i_{[b]}$ can be different from $i_{[b']}$ when $b \neq b'$.

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

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

and

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

where the $\left(i_{[b]}, s\right)$ element of the $\left(N_{[b]} \times \tau\right)$ matrices of $\mathbf{R}_{[b]}$ and $\mathbf{R}_{[b]}^e$ are $R_{i_{[b]},(b-1)\tau+s}$ and $R_{i_{[b]},(b-1)\tau+s}^e$, 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 $N \to \infty$, $N_{[b]} \to \infty$ for each $b = 1, \cdots, B$. 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. Furthermore, noting that factor loadings and factors can be always rescaled and rotated so that the changes can be cancelled out, this assumption is not practically restrictive as long as the systematic factors of an asset pricing model are properly constructed. Also, note that from the last limit in the assumption, we allow that the variance of residuals can vary within a block as well as across blocks, similar to Jones (2001).

Next, we modify Assumption 2 to accommodate our short-time block structure.

Assumption 4. *Consider any function of* $f_{(1)} : \mathbb{R}^{\tau \times K} \to \mathbb{R}^K$, $f_{(2)} : \mathbb{R}^{\tau \times K} \to \mathbb{R}^{K^2}$, $f_{(3)} : \mathbb{R}^{\tau \times K} \to \mathbb{R}^{\tau}$, $f_{(4)} : \mathbb{R}^{\tau \times K} \to \mathbb{R}^{K\tau}$, *and* $f_{(4)} : \mathbb{R}^{\tau \times K} \to \mathbb{R}^{\tau^2}$. *As* $N, T \to \infty$,

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

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

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

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

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

Assumption 4 states that various kinds of cross-sectional errors are independent of the factors so that the time series average of the product of any arbitrary functions of factors and the cross-sectional errors converges to zero over time.

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 $\mathbf{V}_{e,[b]}$ so that the bias due to the small τ is corrected. 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. *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* (A.16) *in the appendix.*

Lastly, we state our main theorem which show that a consistent estimator of the SDF can be constructed even 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* $\widehat{m}_t^e = 1 + \mathbf{f}_t'\widehat{\boldsymbol{\delta}}^e$ *converge to* m_t *and* m_t^e *given in* (2.5) *and* (2.9), *respectively, when the*

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

$$\widehat{\boldsymbol{\delta}} = \mathbf{D}^{-1}\mathbf{U} \tag{2.24}$$

$$\widehat{\boldsymbol{\delta}}^e = -\left(\mathbf{D}^e\right)^{-1}\mathbf{U}^e, \tag{2.25}$$

where

$$\mathbf{D} = \left(\frac{\mathbf{F}_\triangle'\mathbf{F}_\triangle}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}'\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right)\right], \tag{2.26}$$

$$\mathbf{U} = \frac{1}{B}\sum_{b=1}^{B}\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}'\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau}, \tag{2.27}$$

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

$$\mathbf{U}^e = \left(\frac{\mathbf{F}'\mathbf{F}_\triangle}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right], \tag{2.29}$$

$$\mathbf{F}_{\triangle,[b]} = \left[\mathbf{1}_\tau \;\; \mathbf{F}_{[b]}\right],$$

and $\widehat{\mathbf{V}}_{e,[b]}$ is given in lemma 2.1.

The intuition behind Theorem 2.2 follows. We focus on the case of $\widehat{\boldsymbol{\delta}}$ given by (2.24) because the underlying intuition can be applied to $\widehat{\boldsymbol{\delta}}^e$ given by (2.25) in a similar manner. For expositional simplicity, we consider the traded factor case, i.e., $\boldsymbol{\lambda}_f = \boldsymbol{\mu}_{\boldsymbol{f}}$. 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 the $\frac{\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}$ to estimate $\boldsymbol{\delta}$, a bias will be induced by the non-neglible term of $\frac{\mathbf{E}_{[b]}\mathbf{F}_{\triangle,[b]}}{\tau}$ with the finite τ in the regressor. This is why we need to deduct $\frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}$ in (2.26). Furthermore, because the true value of $\boldsymbol{\delta} = [\delta_0 \ \boldsymbol{\delta}_f]$ in (2.6) and (2.10) involves the population moment of $\mathbf{f}_t$, 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 so that the estimator $\widehat{\boldsymbol{\delta}}$ in (2.24) converges to the true $\boldsymbol{\delta}$ in a consistent manner.

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

$$\widehat{\boldsymbol{\delta}}^* = (\mathbf{D}^*)^{-1}\mathbf{U}^* \tag{2.30}$$

$$\widehat{\boldsymbol{\delta}}^{*e} = -(\mathbf{D}^{*e})^{-1}\mathbf{U}^{*e}, \tag{2.31}$$

where

$$\mathbf{D}^* = \left(\frac{\mathbf{F}_\triangle^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}'\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}^*}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{F}_{\triangle,[b]}^*}{\tau^2}\right)\right], \tag{2.32}$$

$$\mathbf{U}^* = \frac{1}{B}\sum_{b=1}^{B}\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}'\mathbf{1}_{N_{[b]}}}{N_{[b]}\tau}, \tag{2.33}$$

$$\mathbf{D}^{*e} = \left(\frac{\mathbf{F}^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{F}_{[b]}^*}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{F}_{[b]}^*}{\tau^2}\right)\right], \tag{2.34}$$

$$\mathbf{U}^{*e} = \left(\frac{\mathbf{F}^{*\prime}\mathbf{F}_\triangle^*}{T}\right)\frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{F}_{\triangle,[b]}^*}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}^{*\prime}\widehat{\mathbf{V}}_{e,[b]}^*\mathbf{1}_\tau}{\tau^2}\right)\right], \tag{2.35}$$

$$\mathbf{F}_{\triangle,[b]}^* = \left[\mathbf{1}_\tau \ \mathbf{F}_{[b]}^*\right], \ \mathbf{F}_{[b]}^* = \mathbf{F}_{[b]}\mathcal{O} + o_p(1),$$

and $\widehat{\mathbf{V}}_{e,[b]}^*$ *is given in lemma* **??**.

Lastly, we show how to incorporate regularization/selection tools such as Ridge regression or Lasso to the estimator defined in Theorem 2.2 so that we can handle a large number of factors. Note that $\widehat{\boldsymbol{\delta}}$ and $\widehat{\boldsymbol{\delta}}^e$, given by (2.24) and (2.25), respectively, can be equivalently expressed as the solutions of the following minimization problems: $\widehat{\boldsymbol{\delta}} = \arg\min_{\boldsymbol{\delta}}(\boldsymbol{\delta}'\mathbf{D}\boldsymbol{\delta} - 2\boldsymbol{\delta}'\mathbf{U})$ and $\widehat{\boldsymbol{\delta}}^e = \arg\min_{\boldsymbol{\delta}^e}(\boldsymbol{\delta}^{e\prime}\mathbf{D}^e\boldsymbol{\delta}^e + 2\boldsymbol{\delta}^{e\prime}\mathbf{U}^e)$, where $\mathbf{D}$, $\mathbf{U}$, $\mathbf{D}^e$

and $\mathbf{U}^e$ are given by (2.26)-(2.29). Hence, the estimators in Theorem 2.2 can be easily extended by adding proper penalty terms to the minimization problems. In particular, consider a problem where we need to select true K factors among a relatively large number of L factors with Lasso penality as in the last part of Section 2.1. Then, the estimator

$$\widehat{\boldsymbol{\delta}}_{lasso} = \arg\min_{\delta_0,\mathbf{d}} \left(\boldsymbol{\delta}'\mathbf{D}\boldsymbol{\delta} - 2\boldsymbol{\delta}'\mathbf{U} + \gamma\|\mathbf{d}\|_1\right), \tag{2.36}$$

where $\mathbf{D}$ and $\mathbf{U}$ are given by (2.26) and (2.27), $\gamma > 0$, $\mathbf{d} = [d_1 \cdots d_L]'$, $\|\mathbf{d}\|_1 = |d_1| + \cdots + |d_L|$, $\mathcal{F}$ is the $(T \times L)$ 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 $(L \times K)$ factor selection matrix of $\mathcal{S}_f$ is defined as before. The (l, k) element of $\mathcal{S}_f$ is 1 if d_l is the k-th non zero elements among $d_1, \cdots, d_l$ and 0 otherwise. Similarly, the lasso-type extension of $\widehat{\boldsymbol{\delta}}^e$ can be formulated by

$$\widehat{\boldsymbol{\delta}}^e_{lasso} = \arg\min_{\mathbf{d}} \left(\boldsymbol{\delta}^{e\prime}\mathbf{D}^e\boldsymbol{\delta}^e + 2\boldsymbol{\delta}^{e\prime}\mathbf{U}^e + \gamma\|\mathbf{d}\|_1\right), \tag{2.37}$$

where $\mathbf{D}^e$ and $\mathbf{U}^e$ are given by (2.28) and (2.29), and $\boldsymbol{\delta}^e = \mathcal{S}_f'\mathbf{d}$.

3 Performance of the SDF estimators in a simulated economy

We provide the simulation evidence on the desired 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 (2017).

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 process implied by 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 the data from Ken French's database.[1] 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

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

realizations over 600 months (1967:01-2016:12), 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: $\lambda_0 = \frac{1}{600}\sum_{t=1}^{600} R_{f,t}$.

Next, we explain how we obtain the parameters for a large number of assets. We obtain all available individual stock returns over 600 months (1967:01-2016:12) from CRSP monthly database. We calibrate the betas ($\boldsymbol{\beta}_i$) and the variances of residual returns $\left(\sigma_{i,\varepsilon} = \mathbb{E}\left[\varepsilon_{i,t}^2\right]\right)$ 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 $\left(\sigma_{i,e} = \mathbb{E}\left[e_{i,t}^2\right]\right)$ for each 14,277 individual stocks which have more than 60 observations over our sample period 1967:01-2016:12.

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 i-th asset in the simulation is chosen to be asset i from CRSP then it is assigned betas($\boldsymbol{\beta}_i$) and the variances 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 $i = 1, \cdots, N$ 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 in our 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, 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 reports the performance in 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 in Theorem 2.1. The results for the unbalanced panel estimators in Theorem 2.2 follow in Panel B of each table. Furthemore, 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 APC method by Connor and Korajczyk (1986) from 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 consider the Pukthuanthong and Roll's (2017) estimator and the conventional GMM estimator for small number of assets.[2] Because Pukthuanthong and Roll's (2017) estimator is not applicable to excess returns, we compare only to GMM estimator for excess returns. 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. The beta and residual variance of each portfolio are calibrated for each of the three asset pricing models with the identical methods used for individual stocks.

We start with 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 report a poor performance of Pukthuanthong and Roll's (2017) estimator in terms of R^2. For example, for $N = 4000$ and $T = 120$, the average R^2 is less than 10%.[3] 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 in the performance of the estimated SDF, comparing to the

[2] We use two-step optimally weighted GMM.

[3] Some readers may find this result contradicting to Proposition 2.1. However, simulation shows that if we increase N upto 10^6, R^2 from Pukthuanthong and Roll's (2017) estimator gets close to 1.

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 it decreases as T increases. The unbalanced panel estimator in Panel B which explicitly incorporates the bias correction clearly eliminates the bias even with short T. We find this pattern for both cases of using the observed 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 test 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.

We proceed to the results for FF3 and FF5. Because the results are qualitatively similar across the two models, we focus on FF3 (Tables 3 and 4). In terms of R^2, the performance of using true factors or estimated factors becomes similar as N and T increases. 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 observed true factors and 0.78 for the case of using the 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 N or T is not large enough, the unbalanced panel estimator still suffers from the bias due to the estimation errors in the estimated factors (Panel B-2). The performance of the small N GMM estimator changes with test portfolios. Obviously, it does not work well with the 10 portfolios formed on Market Beta, from which econometricians would not find sufficient information on the cross sectional prices of some factors in FF3 or FF5 model. Although it 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.

In sum, we show that our SDF estimators have some desired properties from simulation exercise. As N and T increases to the realistic size of financial panel data, R^2 in the regression of the estimated SDF on a constant and the true SDF gets closer to 1. Furthermore, the intercept (slope) converge to 0 (1) with a large but empirically relevant size of N and T. Also, we find that with a small number of estimated factors, our estimators performs similarly to the case with true factors when N is large enough. This is contrasted to the disappointing performance by Pukthuanthong and Roll's (2017) estimator. Interestingly, the SDF estimators for the excess returns suffer less from the small T bias and show faster convergence to the true SDF. Resorting to the superior performances of SDF estimators using excess returns relative to those using gross returns, we mainly focus on the estimators utilizing excess returns in our empirics.

4 Empirical Application

In this section, we apply our SDF estimators to real 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 Korjczyk's (1986) APC). Hence, we consider both cases.

In particular, 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)). 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 conjuction with the momentum (MOM factor). Barillas and Shanken (2017) show 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 1967:01 to 2016:12. Second, we utilize a recent technique developed by Pelger and Lettau (2017). They propose to apply PCA to a matrix strengthened by a signal on the average returns. We apply their method to 209 portfolio returns, which are also used as our test assets in subsection 5.1.2, 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 the large cross section of portflios, 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 expected returns over the sample period of 293 months from 1990:01-2014:05.[4] The rational behind this way of portfolio construction is that the wide cross section of expected return is supposed to be aligned with the wide cross section of true factor loadings, which will help our estimators to search for the right SDF. The expected returns for individual stocks are computed by using a non-linear function of twenty one characteristics of individual stocks. Freyberger, Neuhierl and Weber (2017) provide the details on how to extract the expected returns of an individual stock from various characteristics of the stock.[5] The second set of cross section is a collection of decile portfolios sorted on various characteristics and industry portfolios. In particular, we consider sixteen sets of decile portfolios sorted on various characteristics: accruals (Sloan (1996)), book-to-market ratio (Chan et al (1991), Fama and French (1992, 1993)), cashflow-to-price ratio, 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 (), 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 except the 10 quality minus junk portfolio returns, which are from AQR data library.

4.1.1 1200 portfolios sorted by expected returns

Table 7 reports the estimation results of 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.[6] The sample periods are 293 months over 1990:01-2014:05.

[4]The number of portfolios is determined by the smallest number of individual stocks over the sample period.

[5]We give thanks to Andreas Neuhierl for providing the data on the expected returns of individual stocks.

[6]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 $s = 1, \cdots, 1000$.

The direction of discounting in the estimated stochastic discount factor is pretty well aligned with the intention of each model. In CAPM, the coefficient for market excess returns is significantly negative. For FF3 and HXZ4, every single factor appears to discount a large cross section of test assets. In FF5, the direction of discount for HML is estimated as the opposite. 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). However, the monthly updated value factor of HML (devil) still appears to be significantly priced in BS6. 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 discounting factor.

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 over 1967:01-2016:12. The estimated coefficients along with their standard errors in parenthesis are reported in Table 8.[7]

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 to the case of using 1200 portfolios. One explanation would be that the most of the decile portfolios are constructed on the basis that the return differences are not explained by CAPM. For FF3 and FF5, SMB does not appear to be important in the stochastic discount factor. Given the choice of 209 test portfolios, HXZ4 is the only model such that every single factor affect the discount factor to the direction supported by the model. Although I/A is very important in the construction of SDF in HXZ4, I/A becomes almost redundant in BS6. Regarding statistical factors in Panel B, all of the four principal components from individual stocks affect stochastic discount factor significantly. However, although RP-PCA extracts principal components from the 209 portfolios, only the first principal components seems to matter for the discount factor.

4.2 Application to individual stock returns

In this subsection, we apply our unbalanced panel estimators to individual stocks returns in CRSP. We consider all individual stocks which were traded in the three main

[7]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 $s = 1, \cdots, 1000$.

exchanges of NYSE, AMEX, and NASDAQ over our sample period of 50 years over 1967:01-2016:12. 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. Also, we drop individual stocks with a lifespan of less than 5 years. After applying the three filters, we obtain 10112 individual stocks.

4.2.1 Individual Stocks in NYSE, AMEX, NASDAQ

Table 9 reports the results of unbalanced panel data from all individual stocks available in CRSP over our sample period 600 months from 1967:01 to 2016:12. Standard errors are computed by bootstrap method and reported in parenthesis.[8]

In Panel A, the behavior of the estimated stochastic discount factor tend to go along with the intuition. Across all models, MKT is significantly factored into the estimated SDF with the right direction. For FF3 and HXZ4, every factor appears to matter in the estimated SDF. Consistent with the findings from a large cross section of portfolios, the direction of discount for HML is estimated as the opposite. The most of 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 priced.

4.2.2 Individual Stocks in NYSE

We repeat the analysis for a smaller set of individual stocks traded in NYSE and report results in Table 10. Over our sample period of the 600 months from 1967:01-2016:01, we have 3373 NYSE common stocks after applying five dollar filter.

Since the results from NYSE individual stocks are consistent with the overall results from all individual stocks, we point out only the differences. For the cross section of NYSE stocks, the importance of SMB disappears in FF3. One reason might be that the size of NYSE stocks tend to be larger. However, noting that SMB revives in other models of HXZ4, FF5, and BS6, the weak behavior of SMB in FF3 might be caused by missing factors in FF3. In Panel B, we find that the fourth principal component of APC is not priced significantly anymore. Recalling that the factor was significantly priced in the cross section of all individual stocks, it hints the possibility that the factor is important only for the non-NYSE stocks.

[8] 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 $s = 1, \cdots, 1000$.

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 other methods in an economy with risk matching that of the U.S. equity market.

A common thread through most of the asset pricing models is that a small number of pervasive risks about which investors care determine the price of a large cross section of individual assets. To reflect this aspect of an economy, we assume that return generating process follows a linear factor structure in this paper and restrict our attention to the linear SDF estimator. Relaxing this linearity assumption would be an interesting venue in the future research.

A Proofs

Proof of Theorem 2.1 Define $\mathbf{B}_{\triangle} = [\mathbf{1}_N \mathbf{B}]$. Assumption 1 implies that

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

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

In addition, from Assumption 2, we have that

$$\frac{\mathbf{B}_{\triangle}'\mathbf{E}\mathbf{F}_{\triangle}}{NT} = \begin{bmatrix} \frac{\mathbf{1}_N'\mathbf{E}\mathbf{1}_T}{NT} & \frac{\mathbf{1}_N'\mathbf{E}\mathbf{F}}{NT} \\ \frac{\mathbf{B}'\mathbf{E}\mathbf{1}_T}{NT} & \frac{\mathbf{B}'\mathbf{E}\mathbf{F}}{NT} \end{bmatrix} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1)} \tag{A.3}$$

$$\frac{\mathbf{F}_{\triangle}'\mathbf{E}'\mathbf{E}\mathbf{F}_{\triangle}}{NT^2} = \begin{bmatrix} \frac{\mathbf{1}_T'\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2} & \frac{\mathbf{1}_T'\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2} \\ \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2} & \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2} \end{bmatrix} \xrightarrow{p} \mathbf{0}_{(K+1)\times(K+1).} \tag{A.4}$$

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

$$\widetilde{\boldsymbol{\delta}} = \left(\frac{\mathbf{F}_{\triangle}'\mathbf{R}'\mathbf{R}\mathbf{F}_{\triangle}}{NT^2}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle}'\mathbf{R}'\mathbf{1}_N}{NT}\right). \tag{A.5}$$

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

$$\begin{aligned} \mathbf{R} &= \mathbf{1}_N\lambda_0\mathbf{1}_T' + \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E} \\ &= \mathbf{B}_{\triangle}\boldsymbol{\Lambda}\mathbf{F}_{\triangle}' + \mathbf{E}, \end{aligned} \tag{A.6}$$

where

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_0 & \mathbf{0}_K' \\ (\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f) & \mathbf{I}_K \end{bmatrix}. \tag{A.7}$$

From Assumptions 1, 2, and the limits of (A.1)-(A.4), 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

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

Hence, from (A.5), it follows that

$$\widetilde{\boldsymbol{\delta}} \xrightarrow{p} \left(\mathbf{V}_{f_{\triangle}}\boldsymbol{\Lambda}'\mathbf{V}_{\beta_{\triangle}}\boldsymbol{\Lambda}\mathbf{V}_{f_{\triangle}}\right)^{-1}\mathbf{V}_{f_{\triangle}}\boldsymbol{\Lambda}'\left[1\boldsymbol{\mu}_{\beta}'\right]' = \mathbf{V}_{f_{\triangle}}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}_{\beta_{\triangle}}^{-1}\left[1 \quad \boldsymbol{\mu}_{\beta}'\right]'. \tag{A.8}$$

Since

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

it follows that

$$\mathbf{V}_{f_{\triangle}}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{V}_{\beta_{\triangle}}^{-1}\left[1\boldsymbol{\mu}_{\beta}'\right]' = \frac{1}{\lambda_0}\begin{bmatrix} \left(1 + \boldsymbol{\mu_f'}\Sigma_f^{-1}\boldsymbol{\lambda}_f\right) \\ -\boldsymbol{\Sigma_f^{-1}}\boldsymbol{\lambda_f} \end{bmatrix} = \boldsymbol{\delta}. \tag{A.9}$$

Combining (A.8) and (A.9), 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

$$\widetilde{\boldsymbol{\delta}}^e = -\left(\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{F}}{NT^2}\right)^{-1}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^e\mathbf{1}_T}{NT^2}. \tag{A.10}$$

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

$$\begin{aligned}\mathbf{R}^e &= \mathbf{B}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu_f}\right)\mathbf{1}_T' + \mathbf{B}\mathbf{F}' + \mathbf{E} \\ &= \mathbf{B}\boldsymbol{\Lambda}^e\mathbf{F}_{\triangle}' + \mathbf{E},\end{aligned} \tag{A.11}$$

where

$$\boldsymbol{\Lambda}^e = \begin{bmatrix} \left(\boldsymbol{\lambda_f} - \boldsymbol{\mu_f}\right) & \mathbf{I}_K \end{bmatrix}. \tag{A.12}$$

From Assumptions 1 and 2, we have that

$$\begin{aligned}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^{e}\mathbf{F}}{NT^2} &= \frac{\mathbf{F}'\mathbf{F}_{\triangle}}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{B}}{N}\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle}'\mathbf{F}}{T} + \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{F}}{NT^2} \\ &+ \frac{\mathbf{F}'\mathbf{E}'\mathbf{B}}{NT}\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle}'\mathbf{F}}{T} + \frac{\mathbf{F}'\mathbf{F}_{\triangle}}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{E}\mathbf{F}}{NT} \\ &\xrightarrow{p} \left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_{\beta}\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]'\end{aligned}$$

and that

$$\begin{aligned}\frac{\mathbf{F}'\mathbf{R}^{e\prime}\mathbf{R}^{e}\mathbf{1}_T}{NT^2} &= \frac{\mathbf{F}'\mathbf{F}_{\triangle}}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{B}}{N}\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle}'\mathbf{1}_T}{T} + \frac{\mathbf{F}'\mathbf{E}'\mathbf{E}\mathbf{1}_T}{NT^2} \\ &+ \frac{\mathbf{F}'\mathbf{E}'\mathbf{B}}{NT}\boldsymbol{\Lambda}^{e}\frac{\mathbf{F}_{\triangle}'\mathbf{1}_T}{T} + \frac{\mathbf{F}'\mathbf{F}_{\triangle}}{T}\boldsymbol{\Lambda}^{e\prime}\frac{\mathbf{B}'\mathbf{E}\mathbf{1}_T}{NT} \\ &\xrightarrow{p} \left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_{\beta}\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} 1 & \boldsymbol{\mu}_f' \end{array}\right]'.\end{aligned}$$

Hence, from (A.10), it follows that

$$\begin{aligned}\widetilde{\boldsymbol{\delta}}^{e} &\xrightarrow{p} -\left(\left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_{\beta}\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]'\right)^{-1}\left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]\boldsymbol{\Lambda}^{e\prime}\mathbf{V}_{\beta}\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} 1 & \boldsymbol{\mu}_f' \end{array}\right]' \\ &= -\left(\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} \boldsymbol{\mu_f} & \mathbf{V}_f \end{array}\right]'\right)^{-1}\boldsymbol{\Lambda}^{e}\left[\begin{array}{cc} 1 & \boldsymbol{\mu}_f' \end{array}\right]' = -\left(\left(\boldsymbol{\lambda_f}-\boldsymbol{\mu_f}\right)\boldsymbol{\mu_f}' + \mathbf{V}_f\right)^{-1}\left(\boldsymbol{\lambda}_f - \boldsymbol{\mu}_f + \boldsymbol{\mu}_f\right) \\ &= -\left(\boldsymbol{\lambda_f}\boldsymbol{\mu_f}' + \boldsymbol{\Sigma}_f\right)^{-1}\boldsymbol{\lambda}_f = \boldsymbol{\delta}^{e}. \qquad \text{(A.13)}\end{aligned}$$

From (A.9) and (A.13), we complete the proof of the theorem. □

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

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

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 homoscedasticity condition, as $N \to \infty$,

$$\begin{aligned}
\frac{\mathbf{R}'\mathbf{R}}{N} &= \mathbf{F}_\triangle\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{B}_\triangle}{N}\boldsymbol{\Lambda}\mathbf{F}_\triangle' + \frac{\mathbf{E}'\mathbf{E}}{N} + \frac{\mathbf{E}'\mathbf{B}_\triangle}{N}\boldsymbol{\Lambda}\mathbf{F}_\triangle' + \mathbf{F}_\triangle\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{E}}{N} \\
&\xrightarrow{p} \mathbf{F}_\triangle\boldsymbol{\Lambda}'\mathbf{V}_{\beta_\triangle}\boldsymbol{\Lambda}\mathbf{F}_\triangle' + s\mathbf{I}_T \qquad \text{(A.14)} \\
\frac{\mathbf{R}'\mathbf{1}_N}{N} &= \mathbf{F}_\triangle\boldsymbol{\Lambda}'\frac{\mathbf{B}_\triangle'\mathbf{1}_N}{N} + \frac{\mathbf{E}'\mathbf{1}_N}{N} \xrightarrow{p} \mathbf{F}_\triangle\boldsymbol{\Lambda}'\left[1 \quad \boldsymbol{\mu}_\beta'\right]'. \qquad \text{(A.15)}
\end{aligned}$$

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

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

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

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

where the next to the last equality is from (A.9). This completes the proof of the proposition. □

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 $\mathbf{V}_{e,[b]} = \mathbf{v}_{e,[b]}$. We estimate $\mathbf{v}_{e,[b]}$ with $\widehat{\mathbf{v}}_{e,[b]}$ given by

$$\widehat{\mathbf{v}}_{e,[b]} = \left(\mathbf{H}_{[b]} \odot \mathbf{H}_{[b]}\right)^{-1} \mathcal{S}' \text{vec}\left(\frac{\widehat{\mathbf{E}}'_{[b]}\widehat{\mathbf{E}}_{[b]}}{N_{[b]}}\right), \tag{A.16}$$

where

$$\begin{aligned} \mathbf{H}_{[b]} &= \mathbf{J}_\tau - \mathbf{J}_\tau \mathbf{F}_{[b]} \left(\mathbf{F}'_{[b]} \mathbf{J}_\tau \mathbf{F}_{[b]}\right)^{-1} \mathbf{F}'_{[b]} \mathbf{J}_\tau \\ \mathbf{J}_\tau &= \mathbf{I}_\tau - \frac{1}{\tau} \mathbf{1}_{\tau \times \tau}, \end{aligned} \tag{A.17}$$

and the $(N \times \tau)$ matrix of $\widehat{\mathbf{E}}_{[b]}$ 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 $\widehat{\mathbf{E}}_{[b]} = \mathbf{R}^e_{[b]}\mathbf{H}_{[b]}$. 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

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

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.1. *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* (A.17).

Proof From (A.16),

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

Hence,

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

where the last equality is from

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

This completes the proof of the lemma. □

Lemma A.2. *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* (2.26), (A.2), (A.1), *and* (A.7), *respectively.*

Proof We decompose $\mathbf{D}$ in (2.26) as

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

where

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

From (A.2),

$$\mathbf{D}_1 \xrightarrow{p} \mathbf{V}_{f\triangle}. \tag{A.18}$$

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

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]}\mathbf{\Lambda}\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}, \tag{A.19}$$

where $\mathbf{B}_{\triangle,[b]} = \left[\mathbf{1}_N\mathbf{B}_{[b]}\right]$ and $\mathbf{\Lambda}$ is given in (A.7). Plugging the expression of (A.16), we have

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

yielding

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2}\right) = \mathbf{\Lambda}'\mathbf{V}_{\triangle,\beta}\mathbf{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \mathcal{E}_{D,[b]}, \tag{A.20}$$

where

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

Using the property of $\text{vec}\,(ABC) = (C' \otimes A)\,\text{vec}\,(B)$ and Lemma A.1, we have

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

implying that

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

from Assumption 4. Hence, from the expression of (A.20), we have

$$\mathbf{D}_2 = \frac{1}{B} \sum_{b=1}^{B} \left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} \right)^{-1} \left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}'_{[b]}\mathbf{R}_{[b]}\mathbf{F}_{\triangle,[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{\triangle,[b]}}{\tau^2} \right) \right].$$
$$= \mathbf{\Lambda}'\mathbf{V}_{\beta}\mathbf{\Lambda} \frac{1}{B} \sum_{b=1}^{B} \frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau} + \frac{1}{B} \sum_{b=1}^{B} \mathcal{E}_{[b]}$$
$$\xrightarrow{p} \mathbf{\Lambda}'\mathbf{V}_{\beta}\mathbf{\Lambda}\mathbf{V}_{f\triangle}. \tag{A.21}$$

Combining (A.18) and (A.21), 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.3. *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* (2.27) *and* (A.2), *respectively.*

Proof Recall the expression of (A.19):

$$\mathbf{R}_{[b]} = \mathbf{B}_{\triangle,[b]} \mathbf{\Lambda} \mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}. \tag{A.22}$$

Note that

$$\begin{aligned} \frac{\mathbf{F}'_{\triangle,[b]} \mathbf{R}'_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]} \tau} &= \left(\frac{\mathbf{F}'_{\triangle,[b]} \mathbf{F}_{\triangle,[b]} \mathbf{\Lambda}'}{\tau^2} \right) \left(\frac{\mathbf{B}_{\triangle,[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]}} \right) + \frac{\mathbf{F}'_{\triangle,[b]} \mathbf{E}_{[b]} \mathbf{1}_{N_{[b]}}}{N_{[b]} \tau^2} \\ &= \left(\frac{\mathbf{F}'_{\triangle,[b]} \mathbf{F}_{\triangle,[b]}}{\tau} \right) \mathbf{\Lambda}' \boldsymbol{\mu}_\beta + \mathcal{E}_{U,[b]}, \end{aligned} \tag{A.23}$$

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 Assumption 4,

$$\frac{1}{B} \sum_{b=1}^{B} \mathcal{E}_{U,[b]} \xrightarrow{p} \mathbf{0}_{K+1}. \tag{A.24}$$

Hence, from (A.23) and (A.24),

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

completing the proof of the lemma. □

Lemma A.4. *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* (2.28) *and* (A.12), *respectively.*

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

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

where

$$\mathbf{D}_1^e = \frac{\mathbf{F}'\mathbf{F}_\triangle}{T},$$

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

From Assumption 1,

$$\mathbf{D}_1^e \xrightarrow{p} \begin{bmatrix} \boldsymbol{\mu_f} & \mathbf{V}_f \end{bmatrix}. \tag{A.25}$$

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

$$\mathbf{R}^e_{[b]} = \mathbf{B}_{[b]}\mathbf{\Lambda}^e\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}. \tag{A.26}$$

Plugging the expression of (A.16), 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

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right) = \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau} + \mathcal{E}_{D^e,[b]}, \tag{A.27}$$

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 $\text{vec}\,(ABC) = (C' \otimes A)\,\text{vec}\,(B)$ and Lemma A.1, we have

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

implying that

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

from Assumption 4. Hence, from the expression of (A.27), we have

$$\begin{aligned}\mathbf{D}_2^e &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{F}_{[b]}}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{F}_{[b]}}{\tau^2}\right)\right].\\ &= \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{[b]}}{\tau}\right) + \frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{[b]}\\ &\xrightarrow{p} \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\left[\begin{array}{cc}\boldsymbol{\mu_f} & \mathbf{V}_f\end{array}\right]'.\end{aligned} \tag{A.28}$$

Combining (A.25) and (A.28), we have

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

completing the proof of the lemma. □

Lemma A.5. *As $N, T \to \infty$,*

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

where $\mathbf{U}^e$ and $\boldsymbol{\Lambda}^e$ are given in (2.29) *and* (A.12)*, respectively.*

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

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

where

$$\mathbf{U}_1^e = \frac{\mathbf{F}'\mathbf{F}_\triangle}{T},$$

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

From Assumption 1,

$$\mathbf{U}_1^e \xrightarrow{p} \left[\begin{array}{cc} \boldsymbol{\mu}_{\boldsymbol{f}} & \mathbf{V}_f \end{array}\right]. \tag{A.29}$$

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

$$\mathbf{R}^e_{[b]} = \mathbf{B}_{[b]}\mathbf{\Lambda}^e\mathbf{F}'_{\triangle,[b]} + \mathbf{E}_{[b]}. \tag{A.30}$$

Plugging the expression of (A.16), we have

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

yielding

$$\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{R}^{e\prime}_{[b]}\mathbf{R}^{e}_{[b]}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}'_{\triangle,[b]}\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right) = \mathbf{\Lambda}^{e\prime}\mathbf{V}_\beta\mathbf{\Lambda}^{e}\frac{\mathbf{F}'_{\triangle,[b]}\mathbf{1}_\tau}{\tau} + \mathcal{E}_{U^e,[b]}, \tag{A.31}$$

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.1, we have

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

implying that

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

from Assumption 4. Hence, from the expression of (A.31), we have

$$
\begin{aligned}
\mathbf{U}_2^e &= \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{\triangle,[b]}}{\tau}\right)^{-1}\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{R}_{[b]}^{e\prime}\mathbf{R}_{[b]}^{e}\mathbf{1}_\tau}{N_{[b]}\tau^2} - \frac{\mathbf{F}_{\triangle,[b]}'\widehat{\mathbf{V}}_{e,[b]}\mathbf{1}_\tau}{\tau^2}\right)\right]. \\
&= \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\frac{1}{B}\sum_{b=1}^{B}\left(\frac{\mathbf{F}_{\triangle,[b]}'\mathbf{F}_{[b]}}{\tau}\right) + \frac{1}{B}\sum_{b=1}^{B}\mathcal{E}_{U^e,[b]} \\
&\xrightarrow{p} \boldsymbol{\Lambda}^{e\prime}\mathbf{V}_\beta\boldsymbol{\Lambda}^e\begin{bmatrix} 1 & \boldsymbol{\mu}_{\boldsymbol{f}}' \end{bmatrix}'. \qquad \text{(A.32)}
\end{aligned}
$$

Combining (A.29) and (A.32), we have

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

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.2 and A.3, 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.9). From Lemmas A.4 and A.5, we have that

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

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

Proof of Corollary 2.2 Simple algebras confirm that Lemmas 2.1 and A.1-A.5 still hold for the estimated factors. Then, the corollary follows. □

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: 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	CMW	RMW	MOM	HML (devil)
CAPM	-8.07								
	(0.08)								
FF3	-5.81	-5.27	-1.79						
	(0.25)	(0.58)	(0.82)						
HXZ4	-6.63	-7.48		-4.00	-5.54				
	(0.28)	(0.59)		(1.17)	(0.91)				
FF5	-7.94	-7.41	7.66			-6.97	-12.17		
	(0.37)	(0.85)	(1.27)			(1.23)	(1.62)		
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)
	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)					

This reports the estimation results of balanced panel estimator using 1200 portfolios sorted by expected returns. 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 decile portfolios and 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 293 months over 1990:01-2014:05.

Table 8: SDF Estimates for the Balanced Panel of 209 Portfolios

Model/Methods	Factors								
	Panel A: Specific Asset Pricing Models								
	MKT	SMB	HML	I/A	ROE	CMW	RMW	MOM	HML (devil)
CAPM	-2.40								
	(0.13)								
FF3	-3.50	1.04	-6.44						
	(0.33)	(1.92)	(2.19)						
HXZ4	-4.11	-6.00		-13.34	-11.61				
	(0.22)	(3.13)		(3.14)	(3.26)				
FF5	-4.90	-0.60	3.52			-2.66	-18.21		
	(0.54)	(1.36)	(5.18)			(5.64)	(5.70)		
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)
	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 reports the estimation results of balanced panel estimator using 209 portfolios (16 decile portfolios and 49 industry portfolios). 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 decile portfolios and 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 600 months over 1967:01-2016:12.

Table 9: 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	CMW	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 reports the estimation results of unbalanced panel estimator using all individual stocks in CRSP. 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 decile portfolios and 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 600 months over 1967:01-2016:12.

Table 10: 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	CMW	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 reports the estimation results of unbalanced panel estimator using NYSE individual stocks. 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 decile portfolios and 49 industry portfolios). Standard errors are computed by bootstrap method and reported in parenthesis. The sample periods are 600 months over 1967:01-2016:12.

References

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

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

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

[4] Chen, Z., Connor, G., and Korajczyk, R. A., 2017, A performance comparison of large-n factor estimators, *Review of Asset Pricing Studies*, forthcoming, https://ssrn.com/abstract=2380144.

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

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

[7] 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.

[8] ——, 1987, Estimating pervasive economic factors with missing observations, Working paper, Northwestern University, https://ssrn.com/abstract=1268954.

[9] ——, 1988, Risk and return in an equilibrium APT: Application of a new test methodology, *Journal of Financial Economics* 21, 255-289.

[10] ——, 1991, The attributes, behavior, and performance of U.S. mutual funds, *Review of Quantitative Finance and Accounting* 1, 5-26.

[11] ——, 1993, A test for the number of factors in an approximate factor model, *Journal of Finance* 48, 1263-1291.

[12] Fama, E. F., and French, K. R., 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56

[13] ——, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 3-56

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

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

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

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

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

[19] ——, 1998, An asymptotic theory for estimating beta-pricing Models using cross-sectional regression, *Journal of Finance* 53, 1285–1309.

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

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

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

[23] Kim, S., and Skoulakis, G., 2017, Ex-post risk premia estimation and Inference using large cross-sections: The regression-callibration approach, Working paper, Georgia Institute of Technology.

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

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

[26] Pukthuanthong, K., and Roll, R., 2016, An agnostic and practically useful estimator of the stochastic discount factor, Caltech Working Paper #1405

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

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

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

[30] ——, 2002, Forecasting using principal components from a large number of predictors, *Journal of the American Statistical Association* 97, 1-13.