Inference for Nonparametric High-Frequency Estimators with an Application to Time Variation in Betas*

Ilze Kalnina[†]
Université de Montréal

July 24, 2015

Abstract

We consider the problem of conducting inference on nonparametric high-frequency estimators without knowing their asymptotic variances. We prove that a multivariate subsampling method achieves this goal under general conditions that were not previously available in the literature. We suggest a procedure for a data-driven choice of the bandwidth parameters. Our simulation study indicates that the subsampling method is much more robust than the plug-in method based on the asymptotic expression for the variance. Importantly, the subsampling method reliably estimates the variability of the Two Scale estimator even when its parameters are chosen to minimize the finite sample Mean Squared Error; in contrast, the plug-in estimator substantially underestimates the sampling uncertainty. By construction, the subsampling method delivers estimates of the variance-covariance matrices that are always positive semi-definite.

We use the subsampling method to study the dynamics of financial betas of six stocks on the NYSE. We document significant variation in betas within year 2006, and find that tick data captures more variation in betas than the data sampled at moderate frequencies such as every five or twenty minutes. To capture this variation we estimate a simple dynamic model for betas. The variance estimation is also important for the correction of the errors-in-variables bias in such models. We find that the bias corrections are substantial, and that betas are more persistent than the naive estimators would lead one to believe.

*I would like to thank an associate editor and two referees for many helpful suggestions. I would also like to thank Kirill Evdokimov, Vasyl Golosnoy, Silvia Gonçalves, Christian Gourieroux, Benoit Perron, Myung Hwan Seo, and Kevin Sheppard for helpful comments, as well as the participants of the Second Annual CIRPEE Applied Financial Time Series Workshop, 2011 Mathematical Finance Days of IFM2, 2011 Canadian Economics Association Meetings, 2011 North American meetings of the Econometric Society, 2011 European summer meetings of the Econometric Society, Second International Panel Data Conference, the 2011 Meeting of the Canadian Econometrics Study Group, and the seminar participants at the HEC Montréal and the University of Western Ontario. I gratefully acknowledge financial support from *Institut de finance mathématique de Montréal*, SSHRC, and FQRSC. Earlier drafts of this paper were previously circulated under the title "Nonparametric Tests of Time Variation in Betas".

[†]Département de sciences économiques and CIREQ, Université de Montréal. Address: C.P.6128, succ. Centre-Ville, Montréal, QC, H3C 3J7, Canada. E-mail address: `ilze.kalnina@umontreal.ca`

1 Introduction

Financial econometrics has proposed many multivariate estimators for high frequency data over relatively short intervals of time such as a day or a quarter. Inference in this setting requires estimation of the variance-covariance matrices of the estimators, which can be of complicated form, and difficult to derive or estimate. The current paper proposes to estimate the asymptotic variance-covariance matrices of high-frequency estimators using a subsampling method, which does not rely on the expression of these matrices.

We apply the multivariate subsampling method to study time variation in financial betas. The failure of the unconditional, constant beta Capital Asset Pricing Model of Lintner (1965) and Sharpe (1964) is widely accepted, but there is no consensus on the nature of time variation in betas. To alleviate this problem, in practice betas are often estimated on rolling windows of say 5 years of monthly data, see, e.g., Fama and MacBeth (1973) and Fama and French (1992). It is not clear however what length of window is short enough to warrant the assumption of constant betas within the window. The dynamics of betas has also been modelled with parametric auto-regressive models, see, e.g., Braun, Nelson, and Sunier (1995), Bekaert and Wu (2000), and Jostova and Philipov (2005). More recently, Andersen, Bollerslev, Diebold, and Wu (2005, 2006), Hansen, Lunde, and Voev (2010), and Patton and Verardo (2012) use high-frequency estimators in auto-regressive models of lower-frequency beta. We contribute to this literature by developing robust tools for econometric analysis of time-variation in betas under general conditions.

We prove the validity of the multivariate inference method for a general class of estimators including many estimators of the integrated covariance matrices. This result holds under general conditions allowing for a rich dynamics of stochastic volatility such as a Brownian semimartingale or long memory structure, leverage effects, and for autocorrelated market microstructure noise. It is therefore a nontrivial extension of the previously available univariate subsampling result in Kalnina (2011). The multivariate method has the appealing property of always giving estimates of the variance-covariances matrices that are positive semi-definite.

Importantly, our simulation study indicates that the multivariate subsampling method is much more robust than the plug-in method based on the asymptotic variance-covariance matrix. In particular, when we consider the Two Scales (TS) estimator of Aït-Sahalia, Zhang, and Mykland (2005) and Zhang (2011) calculated with parameters that minimise its finite sample Mean Squared Error (MSE) as suggested by Bandi and Russell (2011), the subsampling method delivers estimates that are much closer to the finite sample variance of the TS estimator than those of the plug-in method. The result is not driven by the failure of the plug-in estimator; the plug-in estimator is quite close to the asymptotic variance, but the latter can differ significantly from the finite sample variance. As a result, the tests based on the subsampling estimator have better size properties than the tests based on the plug-in estimator across a wide range of tuning parameters of the TS estimator.

We use the following data-driven method to choose the smoothing parameters of our procedure. We simulate many paths from a Heston model fitted to daily option prices and high-frequency stock prices. We then choose those bandwidths that deliver optimal performance. The researcher might want to use different bandwidths for inference and testing applications, and this method allows to choose the desired criterion. We find that the subsampling performance is very stable across a wide range of parameters, and very similar for the two criteria. See Sections 2.3 and 4 for further details.

We apply the multivariate subsampling method to high-frequency data on six stocks on the NYSE with the ETF for S&P500 (SPDR) as the market factor. We consider two types of financial beta estimators: those based on realized variances calculated with moderate frequency data (5, 10, or 20 minutes), and the TS with tick data. We consider two complementary approaches. First, we consider a powerful approach for nonparametric detection of breaks in betas. We find that for each stock, the TS estimator can find at least one break in weekly betas in 2006, while moderate frequency estimators find at least one break in two or four stocks depending on the frequency of data used. Second, we consider the estimation of dynamic models of time variation in betas. Variance estimation is important to correct for the errors-in-variables bias in such models; after bias-correction our results suggest that betas are more persistent than the naive estimators would lead one to believe.

Several papers are related to the multivariate subsampling method. Kalnina (2011) shows the validity of the univariate subsampling method for a general estimator in the absence of the leverage effect; she also assumes the volatility is a Brownian semimartingale, which excludes, for example, the long memory property. The method is partly related to the classical subsampling in the statistics literature for stationary data, see, e.g., Politis and Romano (1994) and Lahiri, Kaiser, Cressie, and Hsu (1999). All the above methods and the current method rely on the squared differences of the estimator calculated on nested subsamples of various forms. In contrast, Mykland and Zhang (2014) recently develop an estimator of the asymptotic variance that relies on the squared differences of adjacent estimators together with a bias correction.

The remainder of this paper is organized as follows. Section 2 introduces the framework and the multivariate subsampling method, and presents the main theoretical results. Section 3 describes methods of analysis of time-variation in betas using the subsampled variances. Section 4 studies the finite sample properties of the proposed methods and investigates the choice of the tuning parameters. Section 5 contains an empirical illustration. Section 6 concludes. All proofs are collected in the appendix.

2 High-Frequency Estimators and their Asymptotic Variances

2.1 Theoretical Framework

A large literature in financial econometrics is concerned with using high frequency data to estimate various integrated functionals of the variance-covariance matrices of asset returns. They often assume the data are discrete observations from some vector-valued process X, which follows a Brownian semimartingale,

$$X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s dW_s, \tag{1}$$

where W is a d-dimensional Brownian motion and σ_s is a $d \times d$ stochastic volatility process. The the spot covariance of X is the $d \times d$ matrix $c_s = \sigma_s \sigma_s^\top$. Often, the parameter of interest θ is an integral of some vector-valued function of the spot covariance

$$\theta = \int_0^T f(c_t)dt, \tag{2}$$

such as the integrated covariance $\int_0^T c_t dt$. We consider fixed T, so set $T = 1$ without loss of generality. We may also be interested in a estimating a nonlinear function $\gamma = \gamma(\theta)$ of θ.

An important application of this multivariate problem is estimation of the financial betas, i.e., the exposures of assets to risk factors. In the simplest case, suppose $X_t = \left(X_t^S, X_t^F\right)'$ contains the log-price of one stock X_t^S and one factor X_t^F, so $d = 2$. The analysis with more stocks and/or factors is not conceptually more difficult, but involves more notation. A series of papers has considered the following measure of beta of the stock,[1]

$$\beta := \beta(\theta) = \frac{\theta_2}{\theta_1}, \text{ where } \theta = \begin{pmatrix} \langle X^F, X^F \rangle \\ \langle X^F, X^S \rangle \end{pmatrix}. \tag{3}$$

In the above, $\langle X^S, X^F \rangle$ is the quadratic covariation between X^S and X^F, which is a natural measure of the co-variability of two processes in continuous time. See Bollerslev and Zhang (2003) and Todorov and Bollerslev (2010) for a discussion of how this beta is related to a discrete-time regression model.

At the highest frequencies such as one second, the assumption that data is drawn from a Brownian semimartingale is not reasonable due to different market microstructure effects such as the bid-ask bounce. To model data at such frequencies one usually assumes that data is generated by a contaminated process $Y = X + \varepsilon$ where ε represents the noise. Suppose we have n equi-distant observations on Y over $[0, 1]$ at times $0 = t_0 < t_1 < \ldots < t_n = 1$. Then, β can be estimated using for example the Two Scales estimator of Zhang (2011) (TS henceforth),[2]

$$\widehat{\langle X^S, X^F \rangle}^{TS} = \left[Y^S, Y^F\right]^{(G_1)} - \frac{\overline{n}_{G_1}}{\overline{n}_{G_2}} \left[Y^S, Y^F\right]^{(G_2)} \tag{4}$$

where

$$\left[Y^S, Y^F\right]^{(G_j)} = \frac{1}{G_j} \sum_{i=G_j}^{n} \left(Y_{t_i}^S - Y_{t_{i-G_j}}^S\right) \left(Y_{t_i}^F - Y_{t_{i-G_j}}^F\right), \; j = 1, 2.$$

The quantities $\widehat{\langle X^F, X^F \rangle}^{TS}$ and $\left[Y^F, Y^F\right]^{(G_l)}$ are again defined analogously. In the above, $\overline{n}_{G_j} = \frac{n - G_j - 1}{G_j}$ for $j = 1, 2$, $G_1 = \lfloor \varphi^{TS} n^{2/3} \rfloor$ for some tuning parameter φ^{TS}, and $G_2 / G_1 \to 0$.

[1] See, e.g., Andersen et al. (2005), Andersen et al. (2006), Barndorff-Nielsen and Shephard (2004), and Bandi and Russell (2005).

[2] Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011) and Christensen, Kinnebrock, and Podolskij (2010) propose alternative multivariate volatility estimators that are robust to market microstructure noise.

2.2 Inference for Multivariate Estimators

The asymptotic distribution of the beta estimator is obtained by applying the Delta method to the asymptotic distribution of the estimator of the bivariate vector θ in (3). Therefore, inference on beta requires estimation of the multivariate covariance matrix of $\widehat{\theta}$. Then, inference on $\widehat{\gamma} = \gamma(\widehat{\theta})$ can be performed by the Delta method.

The asymptotic variances are often of complicated form and are therefore hard to estimate. In the univariate case, Kalnina (2011) proposes a subsampling method that estimates the asymptotic variance of a general estimator $\widehat{\theta}$ without using the analytic expression of the asymptotic variance.

We extend the univariate subsampling method for variance estimation to the multivariate case and prove its validity for a general estimator $\widehat{\theta}$ under much weaker conditions than those previously available. In particular, we allow for leverage effect in X as well as for much richer dynamics in the volatility c_t; the volatility may, for example, follow a long-memory process.

An estimator of the asymptotic variance of $\widehat{\theta}$ can be constructed as follows. Recall that n is the total number of observations. Form a series of longer blocks of observations, m consecutive returns in each block, as well as a series of shorter blocks of observations, J returns in each block, $J < m < n$. For any time interval $[a, b]$, denote by $\widehat{\theta}\left([a, b]\right)$ the estimator $\widehat{\theta}$ calculated using all price observations in the interval $[a, b]$. Using this notation, the subsampling estimator of the asymptotic variance-covariance matrix V is

$$\widehat{V}^{sub} = \left(1 - \frac{J}{m}\right)^{-1} \frac{J}{n} \frac{1}{K} \sum_{k=0}^{K-1} \tau_n^2 \left(\frac{n}{J}\widehat{\theta}_k^{short} - \frac{n}{m}\widehat{\theta}_k^{long}\right) \left(\frac{n}{J}\widehat{\theta}_k^{short} - \frac{n}{m}\widehat{\theta}_k^{long}\right)' \tag{5}$$

where

$$\begin{aligned} \widehat{\theta}_k^{long} &= \widehat{\theta}\left([t_{km}, t_{km+m}]\right) \text{ and} \\ \widehat{\theta}_k^{short} &= \widehat{\theta}\left(\left[t_{km+\lfloor (m-J)/2 \rfloor}, t_{km+\lfloor (m-J)/2 \rfloor + J}\right]\right). \end{aligned}$$

In the above, $K = \lfloor n/m \rfloor$ is the number of subsamples, τ_n is the rate of convergence of $\widehat{\theta}$ when n observations are used, and the term $(1 - J/m)^{-1}$ is a finite sample adjustment factor. This adjustment is negligible asymptotically, but improves finite sample behaviour; it has the usual motivation in variance estimation.

Note that $\widehat{V}^{sub}$, by construction, is positive semi-definite, which is an obviously key property for variance-covariance matrix estimators.

A few comments on the intuition of the estimator $\widehat{V}^{sub}$ are in order. Notice that $\frac{n}{m}\widehat{\theta}_k^{long}$ and $\frac{n}{J}\widehat{\theta}_k^{short}$ estimate the same object, but one uses more observations than the other. Therefore, $\frac{n}{m}\widehat{\theta}_k^{long}$ can be used to center $\frac{n}{J}\widehat{\theta}_k^{short}$. The outer product of the differences is multiplied by the rate of convergence $\frac{n}{J}\tau_n^2$ of $\frac{n}{J}\widehat{\theta}_k^{short}$, and then averaged over subsamples. Hence, the estimator approximates the sum of the variances of the local estimators. The latter sum equals V if V is additive over time. Hence, $\widehat{V}^{sub}$ estimates V under the regularity assumptions described below.

The estimator $\widehat{V}^{sub}$ in (5) uses non-overlapping blocks, but can be modified to use overlapping blocks. The latter estimator is more efficient, but can be more computationally demanding. To describe the definition of the modified estimator, denote by s (for "shift") the number of observations to roll the window to obtain the next subsample, $s \in \{1, ..., m\}$. Then, the number of subsamples is $K = \left\lfloor \frac{n-m}{s} + 1 \right\rfloor$, and the first observation time in the l^{th} long subsample is t_{ls}. The expression in (5) is obtained by setting $s = m$.

We use the following assumptions to prove the consistency of $\widehat{V}^{sub}$ in (5). For any $k_1 \times k_2$ matrix Q let $\|Q\| = \sqrt{\sum_{j=1}^{k_1} \sum_{l=1}^{k_2} Q_{jl}^2}$.

Assumption A1. *θ and V are the following functions of the spot covariance path $\{c_s, s \in [0,1]\}$, $\theta = \int_0^1 f(c_s)\, ds$, $V = \int_0^1 g(c_s)\, ds$ where functions f and g are continuously differentiable.*

Assumption A2. *There exists a constant B_c, and $\alpha > 0$ such that $E\left[\|c_{t_2} - c_{t_1}\|^2\right] \leq B_c |t_2 - t_1|^\alpha$ for all t_1 and t_2. Also, $\{c_s, s \in [0,1]\}$ is tight.*

Assumption A3. *$J \to \infty$, $m \to \infty$, $J/n \to 0$, $m/n \to 0$, $J/m \to 0$, and $\tau_n^2\, Jm^\alpha / n^{1+\alpha} \to 0$.*

Assumption A4. *Let $\mathcal{I}_{k,s} = \left[t_{km+\lfloor (m-s)/2 \rfloor}, t_{km+\lfloor (m-s)/2 \rfloor + s}\right]$. For both $s = J$ and $s = m$,*

$$\frac{1}{K} \sum_{k=1}^{K} \frac{n}{s} \left(\tau_n^2 \left(\widehat{\theta}(\mathcal{I}_{k,s}) - \int_{\mathcal{I}_{k,s}} f(c_u)\, du \right) \left(\widehat{\theta}(\mathcal{I}_{k,s}) - \int_{\mathcal{I}_{k,s}} f(c_u)\, du \right)' - \int_{\mathcal{I}_{k,s}} g(c_u)\, du \right) \xrightarrow{p} 0.$$

We now discuss the above assumptions. Examples of the parameters of interest θ satisfying Assumption A1 include integrated covariance, integrated quarticity, integrated betas in high-frequency regression (see section 4.2 in Mykland and Zhang (2006) and Zhang (2012)), and principal components (see Ait-Sahalia and Xiu (2015)). The asymptotic variances V of the corresponding estimators of such θ typically also satisfy Assumption A1. Assumption A2 is well known to be satisfied with $\alpha = 1$ if c_t is assumed to be a Brownian semimartingale. Lemma 3 below shows that if one assumes a long-memory process with parameter $\overline{\alpha}$, then Assumption A2 is satisfied with $\alpha = \overline{\alpha}$. Assumption A3 requires that there are many observations in each subsample and many subsamples. It also requires J/m to be small so that the long subsample can approximate the true value for centering the estimator on the short subsample. The last requirement in Assumption A3 arises due to the "discretization bias" in the volatility, i.e., from us implicitly approximating c_t by integrals of c_t on short intervals. The less smooth is the volatility, the more restrictive the last condition of Assumption A3 is. Assumption A4 is relatively high level, but it is simple. The term n/s is the inverse of the length of the subsample; it ensures that each term in the sum is of order one. Notice that when we choose the longer subsamples ($s = m$), we have

$$\frac{1}{K} \frac{n}{m} \sum_{k=1}^{K} \int_{\mathcal{I}_{k,m}} g(c_u)\, du = \sum_{k=1}^{K} \int_{\mathcal{I}_{k,m}} g(c_u)\, du = \int_0^1 g(c_u)\, du = V.$$

In this case, Assumption A4 requires that the sample second moment matrix, centered at the true (unknown) parameter, converges in probability to V. Moreover, similar property is assumed to hold when we choose the shorter blocks ($s = J$) and scale up the estimator and the integrals by the larger factor n/J. This scaling compensates for the fact that the union of the smaller blocks does not cover the whole interval $[0, 1]$ (the union has "holes"). Note that the smoothness condition on c_s was already imposed in Assumption A2. The substance of Assumption A4 is therefore a restriction on the dependence in the returns and on the tails of the conditional distribution of the returns. Since the returns are usually modeled as a sum of a Brownian semimartingale and short-memory market microstructure noise, these restrictions are usually satisfied.

With the above assumptions, the subsampling estimator of the asymptotic variance of a general estimator is consistent:

Theorem 1. *Suppose assumptions A1, A2, A3, and A5 hold, let $\widehat{V}^{sub}$ be defined by (5). Then, as $n \to \infty$,*

$$\widehat{V}^{sub} \xrightarrow{p} V.$$

We illustrate the application of Theorem 1 in the TS example with $d = 2$ and the following dynamics of the market microstructure noise $\epsilon = \left(\epsilon^S, \epsilon^F\right)'$:

Assumption N. *The noise ϵ_{t_i} is independent of the efficient price X, it is stationary, exponentially α-mixing, and has finite $(4+\delta)^{th}$ moments for some $\delta > 0$.*

The asymptotic distribution of the TS estimator of the θ vector in (3) is

$$n^{1/6}\left(\begin{pmatrix} \widehat{\langle X^F, X^F\rangle}^{TS} \\ \widehat{\langle X^F, X^S\rangle}^{TS} \end{pmatrix} - \begin{pmatrix} \langle X^F, X^F\rangle \\ \langle X^F, X^S\rangle \end{pmatrix}\right) \Rightarrow MN\left(0, V^{TS}\right), \tag{6}$$

where

$$V_{11}^{TS} = \varphi^{TS}\frac{4}{3}\int_0^1 (c_{u,22})^2 du + 8(\varphi^{TS})^{-2} Var\left(\epsilon^F\right)^2 + 16(\varphi^{TS})^{-2}\lim_{n\to\infty}\sum_{i=1}^{n} Cov\left(\epsilon_0^F, \epsilon_{i/n}^F\right)^2 \tag{7}$$

was first derived in Aï-Sahalia et al. (2011).[3] In (7), the first part is clearly a smooth function of c_u. The remaining parts do not change across time. Therefore, the whole expression V_{11}^{TS} is an integral of a smooth function of c_u and hence satisfies Assumption A1. The same argument applies to other elements of V^{TS}.

The following corollary is proved in the appendix by verifying the assumptions of Theorem 1.

Corollary 2. *Suppose log-price X satisfies equation (1) with $d = 2$, where b_s and σ_s are adapted and càdlàg. Let $\widehat{\theta}_n$ be the TS estimator defined by (4), with sequences of parameters G_1 and G_2 satisfying $G_1 = \lfloor \varphi^{TS} n^{2/3} \rfloor$ for some tuning parameter φ^{TS}, G_2 is such that $\mathrm{Cov}\left(\epsilon_1, \epsilon_{G_2}\right) = o\left(n^{-1/2}\right)$, $G_2 \to \infty$, and $G_2/G_1 \to 0$. Let V be defined by V^{TS} in (6), and $\widehat{V}^{sub}$ be defined by (5). Suppose assumptions A2, A3 and N hold. Then,*

$$\widehat{V}^{sub} \xrightarrow{p} V.$$

[3] The expression for $V_{2,2}^{TS}$ is rather complicated and can be found in Zhang (2011), while the expression for $V_{1,2}^{TS}$ does not seem to be available in the published literature, although it can be derived following the arguments of Zhang (2011).

Since $\tau_n = n^{1/6}$ for the TS estimator, if volatility is a Brownian semimartingale, the last condition of Assumption A3 is satisfied if $Jm/n^{5/3} \to 0$. For general $\alpha > 0$, the last condition of Assumption A3 is $Jm^{\alpha}/n^{2/3+\alpha} \to 0$.

As discussed above, Assumption A2 is satisfied with $\alpha = 1$ if the volatility is Brownian seminartingale. Lemma 3 below presents an alternative sufficient condition of Assumption A2 where volatility is a long-memory process.

Lemma 3. *Let $x(t) = \frac{1}{2}\ln c_t$ be a scalar log-volatility process. Assume it follows dynamics*

$$dx(t) = -\kappa x(t)dt + \gamma dB_{\overline{\alpha}}(t), \qquad t \in [0, T]$$

where

$$B_{\overline{\alpha}}(t) = \int_0^t \frac{(t-s)^{\overline{\alpha}}}{\Gamma(1+\overline{\alpha})} d\widetilde{W}(s),$$

where $\widetilde{W}(t)$ is a standard Brownian motion, and where γ, κ, and α are constants such that $\kappa > 0, 0 < \overline{\alpha} < \frac{1}{2}$. Then, Assumption A2 is satisfied with $\alpha = \overline{\alpha}$.

2.3 The Choice of Parameters for Subsampling

For practical application, one needs to choose specific values of m and J. Here we describe a recommended data-driven method.

First, one fits a parametric model of the stock price dynamics to the real data, on which the subsampling method is to be used. Note that one can use not only the price data, but also any additional data such as data on options if it is available. For example, the model of Heston (1993) is particularly convenient for this purpose. It is a parsimonious model with mean-reverting stochastic volatility, where the prices of options are straightforward to calculate. It is well known how to estimate the parameters of this model. Second, one draws many sample paths from this model, and calculates the subsampled variances for a range of values of m and J. Then, one chooses the values that provide the best performance of the subsampling estimator.

This method has several advantages. First, one is typically concerned whether the analytic expressions of the higher-order approximations with estimated parameters would match the finite sample performance well. Second, while computationally intensive, the method is straightforward to implement. Third, this method allows to easily specify and alter the criterion of interest.

3 Econometric Methods of Analysis of the Dynamic Properties of the Betas

3.1 Tests of Parameter Constancy

Structural asset pricing models such as the conditional CAPM are often difficult to estimate, for example, because they require specifying the information sets of investors. One solution to this problem has been to assume that betas are constant over moderately long periods of time such quarters or years (see, e.g., Lewellen and Nagel, 2006). The question remains whether a given selected interval is short enough for this assumption to be reasonable.

We begin by introducing the necessary notation and describe simple tests of parameter constancy of some scalar parameter β across k time periods (such as days, weeks, months, or quarters). We use the notation β because we later apply this test to financial beta, but the test applies to any parameter of interest. Denote by $\widehat{\beta}$ some generic estimator of β, where

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 & \ldots & \widehat{\beta}_k \end{pmatrix}' \text{ and } \beta = \begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_k \end{pmatrix}',$$

and let n_i be the number of observations in period i. The estimation errors $\widehat{\beta}_i - \beta_i$ of most high-frequency estimators are independent across time periods $i = 1, 2, \ldots, k$ (see, e.g., Mykland and Zhang, 2006), with the asymptotic distribution

$$\tau_{n_1} \Phi \Sigma^{-1/2} \left(\widehat{\beta} - \beta \right) \Rightarrow N\left(0, I_k\right), \tag{8}$$

$$\Sigma = diag\left(V_1, V_2, ..., V_k\right), \text{ and } \Phi = diag\left(\phi_1, \phi_2, \ldots, \phi_k\right),$$

where Σ has the asymptotic variances of $\widehat{\beta}_i$ for each period i on the diagonal, and I_k denotes the $k \times k$ identity matrix. The rate of convergence of the estimator in period 1 is denoted as τ_{n_1}. In most applications it is natural to assume that the number of observations n_i across periods are of the same order of magnitude, so that $\tau_{n_i} = \tau_{n_1}\left(\phi_i + o(1)\right)$ for some positive finite constants ϕ_i, $i = 2, \ldots, k$.

The hypothesis of constancy of β is

$$H_0 : \beta_1 = \ldots = \beta_k, \text{ versus } H_1 : \beta_i \neq \beta_j \text{ for some } i \text{ and } j. \tag{9}$$

Rewriting the null hypothesis as $H_0 : \beta_i = \beta_1$ for $i = 2, \ldots, k$ as usual we have $\tau_{n_1}^2 \widehat{\beta}' \underline{\Delta}' \left(\underline{\Delta} \Phi^{-1} \Sigma \Phi^{-1} \underline{\Delta}' \right)^{-1} \underline{\Delta} \widehat{\beta} \Rightarrow \chi_{k-1}^2$, where matrix $\underline{\Delta}$ is $(k-1) \times k$ and is defined as $\underline{\Delta} = (-\mathrm{i}_{k-1}, \mathrm{I}_{k-1})$, where i_{k-1} is a $(k-1) \times 1$ vector of ones, so $\underline{\Delta}\beta = 0_{k-1}$ under H_0. The convergence to χ_{k-1}^2 follows from (8). The subsampling method described in the previous section can be used to estimate each of the elements of the diagonal matrix Σ; denote this estimator by $\widehat{\Sigma}$. Since we have established the consistency of $\widehat{\Sigma}$ in the previous section, we have the test statistic

$$\tau_{n_1}^2 \widehat{\beta}' \underline{\Delta}' \left(\underline{\Delta} \Phi^{-1} \widehat{\Sigma} \Phi^{-1} \underline{\Delta}' \right)^{-1} \underline{\Delta} \widehat{\beta} \Rightarrow \chi_{k-1}^2. \tag{10}$$

We can similarly test the null hypothesis of constant betas jointly across stocks, i.e.,

$$H_0 : \beta_1^{(j)} = \beta_2^{(j)} = \ldots = \beta_k^{(j)} \text{ for each } j,$$

where $\beta_i^{(j)}$ denotes the beta of the j^{th} stock in the i^{th} time period. In this case, the matrix $\widehat{\Sigma}$ will be block diagonal, with each block corresponding to a separate time period, and can also be estimated by the subsampling method of the previous section.

3.2 Nonparametric Search for Breaks in Betas

If the test for constant betas rejects, we do not have any information as to the number or timing of breaks. This might be information we are interested in. We propose to search for breaks in betas nonparametrically by considering each possible break time as a hypothesis and accounting for multiple testing.

A popular method for controlling for multiple testing is Bonferroni correction. It controls the family-wise error (FWE, see (11) below), but is in general conservative. In a sense, Bonferroni correction achieves the control of FWE by assuming the worst-case dependence structure across the test statistics. White (2000) proposes a bootstrap-based solution that implicitly accounts for the dependence structure across the test statistics. Romano and Wolf (2005) propose a stepwise procedure that accounts for the dependence structure across the test statistics, controls FWE, and has higher power due to additional steps that can reject more null hypotheses. The main example in Romano and Wolf (2005) is the search for outperforming trading strategies.

We apply the ideas of Romano and Wolf (2005) to search for breaks in betas instead. Our multiple hypotheses correspond to breaks in different periods. Since we know the dependence structure of estimation errors in betas or their differences across periods, we can fully account for that and do not need to use conservative testing methods.

Our setting is conceptually different from the classical literature on structural breaks in parametric models, see, e.g., Andrews (1993). We do not need many betas to be the same to achieve their identification. Instead, our setting reflects a situation where a lot of high frequency data is available for the purpose of estimating each quantity of interest at lower frequency (such as one weekly beta).

We now describe the procedure. We consider the case of one asset, but the procedure can be extended easily to search for simultaneous breaks across several assets. Suppose we have k time intervals, and we are interested in finding the breaks in beta across these time intervals. Instead of one null hypothesis, we have $k-1$ null hypotheses. Label the null hypotheses by the intervals,

$$H_s : \beta_s = \beta_{s+1} \text{ vs. } H_s' : \beta_s \neq \beta_{s+1}$$

for $s = 1, \ldots, k-1$. We aim to control the family-wise error rate,

$$FWE = P\left\{\text{Reject at least one true null hypothesis}\right\}, \tag{11}$$

i.e., we aim to construct a test that has $\limsup FWE \leq \alpha$ when all null $k-1$ hypotheses are true.

Define the following $k-1$ statistics $\widetilde{w}_1, \widetilde{w}_2, \ldots, \widetilde{w}_{k-1}$ based on the differences between betas on the subsequent intervals,

$$\widetilde{w} = \begin{pmatrix} \widetilde{w}_1 & \widetilde{w}_2 & \ldots & \widetilde{w}_{k-1} \end{pmatrix}' = \tau_{n_1} \Delta \widehat{\beta},$$

where Δ is a $(k-1)\times k$ differencing matrix with $\Delta_{i,j} = 1_{\{i=j\}} - 1_{\{i=j-1\}}$. If all null hypotheses are true,

$$\widetilde{w} = \tau_{n_1}\Delta\widehat{\beta} = \tau_{n_1}\Delta\left(\widehat{\beta}-\beta\right) \Rightarrow MN\left(0, \Delta\Phi^{-1}\Sigma\Phi^{-1}\Delta'\right).$$

Let $\widehat{\Sigma}$ be an estimator of Σ, and let $\hat{s}_i^2$ be the the i'th element on the diagonal of $\left(\Delta\Phi^{-1}\widehat{\Sigma}\Delta'\right)$. We use the standardized test statistics $w = \left(\left|\hat{s}_1^{-1}\widetilde{w}_1\right|, \ldots, \left|\hat{s}_{k-1}^{-1}\widetilde{w}_{k-1}\right|\right)'$. The testing procedure is implemented in a stepwise manner. Order elements of w according to their absolute values, from largest to smallest, $|w_{r_1}| \geq |w_{r_2}| \geq \ldots \geq |w_{r_{k-1}}|$ (r_1 is the index of the largest test statistic and so on). For the first step, the ideal critical value is the $1-\alpha/2$ quantile of the sampling distribution of $\max_j |w_j|$,

$$c_1 = c_1(\alpha) = \inf\left\{x : P\left\{\max_{1\leq s\leq k-1} |w_{r_s}| \leq x\right\} \geq 1-\frac{\alpha}{2}\right\}.$$

Since we can estimate the joint distribution of $w_1, \ldots, w_{k-1}$, the estimate of the above, $\hat{c}_1$, can be obtained by simulation. The test procedure is then to reject those null hypotheses, for which $|w_{r_s}| > \hat{c}_1$.

The first step is sufficient to control FWE. However, adding further steps increases the power of the procedure. The choice of critical values for subsequent steps is analogous. Suppose R_1 hypotheses were rejected in the first step. The ideal critical value c_2 is the $1-\alpha/2$ quantile of the sampling distribution of $\max_{R_1+1\leq s\leq k-1} |w_j|$ defined as

$$c_2 = c_2(\alpha) = \inf\left\{x : P\left\{\max_{R_1+1\leq s\leq k-1} |w_{r_s}| \leq x\right\} \geq 1-\frac{\alpha}{2}\right\},$$

and it can be estimated by simulation as before. For the j^{th} step, the ideal critical value is

$$c_j = c_j(\alpha) = \inf\left\{x : P\left\{\max_{R_{j-1}+1\leq s\leq k-1} |w_{r_s}| \leq x\right\} \geq 1-\frac{\alpha}{2}\right\},$$

where R_{j-1} is the number of hypotheses rejected in the first $j-1$ steps ($R_0 = 0$). The procedure is continued until no new hypotheses can be rejected.

3.3 Estimation of Models of the Dynamics

When the hypothesis of constancy of betas is rejected, as an alternative to searching for breaks, researchers are often interested in estimating models of the dynamics of the beta. Such models often also include low-frequency macroeconomic and financial variables. For example, different asset pricing models imply different dynamics for factor betas, motivating the researcher to estimate the dynamics in betas. If the model includes lagged betas, the errors-in-variables leads to inconsistent parameter estimates. This problem can however be addressed if we have access to the variance of the estimation error. We discuss this in more detail in Section 5.3.

4 Simulation Studies

The current section has two objectives. First, we investigate the suggested procedure for choosing the subsampling parameters m and J. Second, we verify the performance of the subsampling variance estimators in finite samples. We also discuss the choice of the parameters of the Two Scale estimator.

Monte Carlo Design

To make the Monte Carlo design realistic, we use a Heston (1993) model with parameters calibrated from the data, and simulate the data to be irregular and asynchronous. We simulate the efficient log-price for six stocks $X^{(1)}, \ldots, X^{(6)}$ and the market portfolio $X^{(7)}$ over one week using the Heston (1993) model:

$$\begin{aligned} dX_t^{(j)} &= \left(\alpha_1^{(j)} - c_t^{(j)}/2\right) dt + \sigma_t^{(j)} dW_t^{(j)} \\ dc_t^{(j)} &= \alpha_2^{(j)} \left(\alpha_3^{(j)} - c_t^{(j)}\right) dt + \alpha_4^{(j)} \left(c_t^{(j)}\right)^{1/2} dB_t^{(j)}, \; j = 1, \ldots, 7, \end{aligned}$$

where $c_t^{(j)} = \left(\sigma_t^{(j)}\right)^2$, and $W_t^{(j)}$ and $B_t^{(j)}$ are Brownian Motions with $Corr\left(W_t^{(j)}, B_t^{(j)}\right) = \rho^{(j)}$. The latter correlation induces the classical leverage effect for each of the stocks and the market portfolio.

To obtain realistic values of the dynamics of the efficient log-price, we calibrate them to the data as follows. The parameters of processes $X^{(1)}, \ldots, X^{(7)}$ are matched to data from the seven assets we use in the empirical application: AIG, GE, IBM, $INTC$, MMM, and $MSFT$ stock prices, and the SP500 index, see Section 5.1 for further details on the data. For each of the six stocks, we collect full record transaction prices as well as daily option data over the year 2006. For the market portfolio, we use the full-record transaction prices of the S&P500 index ETF (ticker SPY) as well as the daily S&P500 index option data over the year 2006 (ticker SPX). The parameter $\alpha_4^{(j)}$ is estimated using the following identity:

$$\alpha_4^{(j)} = \frac{[c^{(j)}, c^{(j)}]_t}{[X^{(j)}, X^{(j)}]_t}. \tag{12}$$

The numerator is the quadratic variation of the spot variance of the j^{th} asset. To estimated it, we use moderate frequency price returns and the estimator of Vetter (2011), which has been extended to include jump truncation by Jacod and Rosenbaum (2012). We estimate the denominator in the above with truncated realized variance, see Mancini (2009). The initial value of the variance process is constrained to equal the value of the parameter $\alpha_3^{(j)}$. Parameters $\alpha_2^{(j)}$, $\alpha_3^{(j)}$, and $\rho^{(j)}$ are chosen to minimise the sum of squared weighted differences between the model-implied option prices and the observed option data of the asset j, with weights being smaller when the bid-ask spread is larger. We set $\alpha_1^{(j)}$ to 0.05 as in Zhang, Mykland, and Aït-Sahalia (2005). Finally, we set the correlation of the individual stock and the market $Corr\left(W_t^{(j)}, W_t^{(7)}\right) = \varrho^{(j)}$, $j = 1, \ldots, 6$, to the value of the realized beta with 50-tick observations of the j^{th}

stock.[4] In this model, the beta over [0,1] is

$$\beta^{(j)} = \varrho^{(j)} \int_0^1 \sigma_u^{(j)} \sigma_u^{(7)} du \Bigg/ \int_0^1 \left(\sigma_u^{(7)}\right)^2 du \ , \ j = 1, \ldots, 6. \tag{13}$$

Hence, we obtain six sets of parameters for the bivariate model with one stock and the market factor; we denote them as scenarios $(1), \ldots, (6)$.

Microstructure noise is simulated as a normally distributed white noise with variance $\xi^{(j)} IQ^{(j)}$, where $\xi^{(j)}$ is a noise-to-signal ratio that equals either 0 or 0.001 (all estimated values of the noise-to-signal ratio are between these two values, see Table E.1), and $IQ^{(j)}$ is the weekly integrated quarticity of the j^{th} stock. We simulate the noise to be i.i.d. to minimise the number of total scenarios and parameters, and concentrate instead on the choice of key smoothing parameters; note that the properties of the univariate subsampling method with autocorrelated and heteroscedastic noise have been documented before. Observed prices are efficient log-prices plus noise.

We match the number of asynchronous observations of each asset, n_j, $j = 1, \ldots, 7$, to the data (see the first column in Table E.1). To do so, we first simulate one week of 1 second synchronous observations. From these, we take n_i irregular and asynchronous observations by drawing a random permutation of all 1-second observation times in a week, taking the first n_i of them, and sorting them. Observations are then synchronized using the Refresh Time method,[5] resulting in a random number of observations to be used for estimation.

The Choice of Parameters of the Two Scale Estimator (Table M.1)

The Two Scale estimator is calculated according to equation (4) with $G_2 = 1$ (for the i.i.d. noise) and the finite sample adjustment suggested in Zhang, Mykland, and Aït-Sahalia (2005). For implementation, we need to select the G_1 parameter. There are two principles of data-based selection of this parameter available in the literature: by minimisation of the asymptotic variance (see Zhang, Mykland, and Aït-Sahalia (2005)) and by minimisation of the finite sample MSE (see Corollary 3 of Bandi and Russell (2011), henceforth BR). We denote the theoretical values of these rules by $G_1^\star$ and G_1^{BR}, respectively. Both depend on the unobservable values of the integrated quarticity and the variance of the noise. The former can be estimated by, e.g., 5-minute realized quarticity (see Barndorff-Nielsen and Shephard (2004)) and the latter can be estimated by, e.g., the realized variance divided by twice the number of observations (see Bandi and Russell (2008)). We denote the estimated values by $\widehat{G}_1^\star$ and $\widehat{G}_1^{BR}$, respectively. The last four rows of Table M.1 show the average over simulations of the four resulting values of G_1 for the TS-beta estimator. The table also shows the finite sample MSE across simulations of the TS-beta estimator for a range of G_1. In all

[4] The realized beta estimator is defined as $\widehat{\beta}^{RV} = \widehat{\theta}_2^{RV} / \widehat{\theta}_1^{RV}$ where $\widehat{\theta}_1^{RV} = [X^F, X^F] = \sum_{j=1}^n (X_{t_j}^F - X_{t_{j-1}}^F)(X_{t_j}^F - X_{t_{j-1}}^F)$ is the realized variance of the market returns, $\widehat{\theta}_2^{RV} = [X^S, X^F] = \sum_{j=1}^n (X_{t_j}^S - X_{t_{j-1}}^S)(X_{t_j}^F - X_{t_{j-1}}^F)$ is the realized covariance of the stock and the market returns, and where θ is defined in (3).

[5] See Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011).

cases considered, the BR rule results in a smaller finite sample MSE than the rule based on the asymptotic variance.

The Choice of Subsampling Parameters (Tables M.2-M.5)

To implement the subsampling estimator, we require a choice of J for the short subsample and m for the long subsample. To obtain a data-driven method for choosing m and J, we calibrate a parametric model to the data we use and choose the values of m and J that minimise the distance to the true finite sample variability in the estimator. This strategy of choosing parameters is another reason for matching the Heston model parameters to the data in the current section.

Tables M.2-M.3 show the average over simulations of the subsampled variances $\widehat{V}^{sub}$ of the TS-beta estimator. (The quantities V^{FS}, V^{Asy}, and $\hat{V}^{pl}$ are discussed below.) We find in that the values of subsampled variances are relatively flat over wide ranges of m and J for all G_1 considered indicating the method is not very sensitive to the choice of the smoothing parameters. This is a very desirable property, but it does imply that pinning down the exact values of J and m is difficult, as they lead to very similar finite sample performances. In light of these results we choose $J = 500$ and $m = 3000$ for the applications.

We also consider the coverage of the nominal 0.95 confidence interval based on the TS-beta estimator. The results are given in Tables M.4 and M.5. We again find that the results are relatively flat over wide ranges of m and J. To control the number of total tables, Tables M.2-M.5 consider scenario (3) (with the design calibrated to IBM data), but the same conclusions hold for other scenarios.

Performance of the Subsampling Estimator (Tables M.6-M.13)

We now consider the problem of assessing the performance of the subsampled variances, first, in matching the true variability of the TS estimator, second, in terms of the coverage of the confidence intervals of the TS estimator. Due to stochastic volatility, the asymptotic and finite sample variances vary across simulations. In this setting, we can obtain the finite sample variability as the scaled mean squared estimation error, e.g., in the case of TS-betas, it is the average over simulations of $n^{1/3}\big(\widehat{\beta}^{TS} - \beta\big)^2$. We denote it by V^{FS}. We also compare the subsampled variances with two benchmarks that rely on the analytic expression of the asymptotic variances: the (unobserved) asymptotic variance V^{Asy}, as well as its estimated counterpart $\hat{V}^{pl}$ using the same estimation approach as for the choice of G_1.

Tables M.6 and M.7 contain the results for the TS-beta estimator (see also Tables M.2 and M.3 for these results in scenario (3)). The subsampling method appears to be much more robust than the plug-in method based on the expression for the asymptotic variance. While the plug-in estimator estimates the asymptotic variance well, the asymptotic variance itself is not very close to the finite sample variability V^{FS} for relatively small values of G_1. The subsampling method, on the other hand, delivers good estimates of the finite sample variability for the whole range of G_1 considered. As can be seen from Table M.1, these smaller values of G_1 where subsampling method performs particularly well compared to $\hat{V}^{pl}$ (and V^{Asy}), are often close to the values of G_1 that minimise the finite sample MSE of the TS-beta.

We next present the results in terms of the coverage of the confidence intervals for the TS-beta estimator. The results for scenarios (1)-(6) are collected in Tables M.8 and M.9. We confirm the conclusions above: the subsampling method performa very well and is more robust than the plug-in method based on the asymptotic variance with respect to the TS parameter G_1.

Finally, we repeat the above analysis for the TS estimator of the variance of the stock, as this estimator does not contain the nonlinearities of beta. The results are collected in Tables M.10-M.13. Again, the subsampling method is more robust than the plug-in method $\widehat{V}^{pl}$ with respect to the TS parameter.

5 Empirical Analysis

This section implements the above methods with real data using both moderate and high frequencies. We use RV-based estimators for moderate frequencies (5, 10, and 20 minutes), and TS-based estimators for high frequencies (tick data).

Some key implementation choices are as follows. To obtain the asymptotic variance of RV-based estimators, we use Barndorff-Nielsen and Shephard (2004) estimator of the variance-covariance matrix of $\widehat{\theta}^{RV}$ (recall notation in (3)). The estimator of variance-covariance matrix of $\widehat{\theta}^{TS}$ is obtained by subsampling. The Two Scale estimator is implemented with the finite sample correction suggested in Aït-Sahalia, Mykland, and Zhang (2011). In all that follows, length of intervals $[i-1, i)$ is taken to be one week.

5.1 Data and Preliminary Analysis

We use high frequency transactions data on six individual stocks. They are American International Group, Inc. (listed under the ticker symbol AIG), General Electric Co. (GE), International Business Machines Co. (IBM), Intel Co. (INTC), Minnesota Mining and Manufacturing Co. (MMM), and Microsoft Co. (MSFT). To proxy for the market portfolio, we use Standard and Poor's Depositary Receipts (SPIDERS for short, ticker symbol SPY), which are an Exchange Traded Fund set up to mimic the movements of the S&P 500 index. Our data covers the whole year 2006 and is obtained from the NYSE TAQ database. We clean the data according to the recommendations of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) and remove jumps with the thresholding methodology of Mancini (2009).

We start by analyzing the high frequency data. Table E.1 contains some summary statistics of the data before synchronization: transactions per week, estimates of the noise variance, noise-to-signal ratio, and autocorrelations of returns at first three lags. First autocorrelations are all large and negative, which is typical of noisy data and unlikely to arise from Brownian Semimartingale. Second autocorrelations are all positive, some are large. Alternating signs of autocorrelations indicate that the main component of the noise is the bid-ask bounce. In fact, if we removed all zero returns, the remaining data would display very persistent autocorrelation with alternating signs, see, e.g., Griffin and Oomen (2005). In the full data set with zero returns, this effect is attenuated because the switching times of bid and ask are random. Third autocorrelations are of different signs and small. The estimates of the noise variance (columns 2 and 3 in

Table E.1) are very small, and in fact several orders of magnitude smaller than Hansen and Lunde (2006) estimates for year 2004. For example, the simplest estimator of the noise variance is

$$\widehat{\omega}^2 = [X, X] / 2n.$$

Our estimate for INTC in 2006 is $0.518 \cdot 10^{-7}$, while Hansen and Lunde (2006) report this number for 2004 to be $0.46 \cdot 10^{-3}$. Apart from the obvious fact that years are different, there are also important differences in methodology. We calculate $\widehat{\omega}^2$ using the whole year, they calculate it every day and report the annual average. Also, data cleaning can also be an important source of differences.

Table E.2 contains the same summary statistics for the data after synchronization. The number of synchronized observations is smaller, especially after joint synchronization across assets. Noise variances are larger as measured by $\widehat{\omega}^2$, but we can easily verify this is purely due to a larger finite-sample bias caused by the smaller number of observations. In particular, the bias-adjusted estimators

$$\widehat{\omega}^2 = \left([X, X] - \widehat{\langle X, X \rangle}^{TSRV} \right) / 2n$$

are the same with and without synchronization.

Figure E.3 contains volatility signature plots for each individual stock (plots of realized variance against the frequency used in its calculation), as well as the correlation signature plots (plots of realized correlation against the frequency).[6] Volatility signature plots show a large increase for highest frequencies, consistent with the additive noise model where bias explodes as we sample more and more frequently. On the other hand, realized covariances display the so-called Epps effect due to Epps (1979), i.e., they tend to zero as the frequency increases, so that the realized correlations are also driven to zero.[7] Not surprisingly, realized beta signature plots in Figure E.4 show a clear bias towards zero for highest frequencies. Therefore, neither realized variance, nor realized covariance should be calculated using the highest frequencies. On the other hand, the Two Scale estimator, while using all the synchronized data, cancels both the effect of noise and asynchronous observations and is consistent (see Zhang (2011)).

5.2 Testing the Constant Beta Hypothesis and the Search for Breaks

Figures E.5-E.7 show plots of estimated betas using $\widehat{\beta}^{RV}_{10min}$ and $\widehat{\beta}^{TS}$ together with 95% confidence intervals, which are based on the estimator of Barndorff-Nielsen and Shephard (2004) and subsampling, respectively. In fact, similar series of confidence intervals for $\widehat{\beta}^{RV}_{10min}$ was also graphed by Andersen et al. (2004) in their Figures 13-15, except they used 10 minute and daily data to calculate estimated betas over intervals of one quarter. In figures E.5-E.7, we see that beta is estimated more precisely using full record transaction prices.

[6] Realized correlation is defined as

$$\frac{[X, Y]}{\sqrt{[X, X]\,[Y, Y]}},$$

and for correlation signature plots the interval is taken to be the whole year 2006.

[7] Zhang (2011) analytically characterizes this bias for realized covariance based on previous-tick interpolated prices (Refresh Time synchronization method is a special case since it also uses previous-tick interpolation).

The two parameters in $\widehat{\beta}^{TS}$ were chosen as follows. We set $G_2 = 3$ as no stocks (after synchronisation) display autocorrelated returns beyond the second lag. G_1 was chosen as $5G_2$. The two parameters of the subsampling scheme were set to $J = 500$ and $m = 3000$ guided by the simulation results calibrated to the data we use, see Section 4.

Table E.8 contains the results of the test for constant betas for individual stocks. The null hypothesis is that the true beta is constant over some time period. We implement the test for five different time periods: the whole year 2006 and each quarter separately. This means using $k = 52$ and $k = 13$ respectively in equation (9). Four different tests are implemented based on four estimators: $\widehat{\beta}^{RV}_{5min}, \widehat{\beta}^{RV}_{10min}, \widehat{\beta}^{RV}_{20min}$ and $\widehat{\beta}^{TS}$. The reader should be careful when interpreting the p-values since at this stage they are not adjusted to reflect multiple testing. The null hypothesis of beta being constant over the whole year can be rejected using a test based on any of the four estimators/frequencies. For shorter periods, answer varies depending on the stock and the exact time period. The test based on $\widehat{\beta}^{TS}$ can reject the null, at 5% level of significance, for all quarters with three exceptions (AIG Q1, IBM Q1, MMM Q3). The tests based on moderate frequencies show similar results with generally smaller number of rejections.

Table E.9 contains the results of the joint test for constant betas. The null hypothesis tested is that the betas of all 6 stocks are constant across some time interval. We implement the test for the same five time periods and the same estimators as in the univariate case. One would in general expect that it is easier to detect beta variation jointly across stocks (partly because more data is used, partly because the null is different and is less likely to be true). Indeed, we see that with only one exception, even the moderate frequency estimators now reject all null hypotheses.

We next implement the procedure for searching for breaks. Table E.10 shows the number of breaks found among 52 weeks in 2006. The result of the procedure is also the timing of every break that is found (we suppress it for brevity). Clearly, higher frequencies contain more information about breaks. For example, the number of breaks detected by 5-min RV-beta is larger than by the 20-min based RV-beta. (INTC is somewhat of an exception, which seems to be driven by a very large value of beta in the second week, so the breaks before and after the second week are easy to identify). From Table E.10, the TS-beta reveals at least one significant break in betas for every stock, while moderate frequency data does not find any for GE and IBM. The number of identified breaks with high frequency data is higher for every stock and every level of significance considered.

Overall, it seems that tick data contain important additional information about the variability in betas over time. This finding does not appear to be specific to the highest frequency because RV-based beta at 5 minutes similarly contains more information than RV-based beta at 20 minutes. It is well known that RV-based beta cannot be used at frequencies much higher than 5 minutes without accounting for the market microstructure noise. Our results suggest that the noise-robust estimators such as the TS-beta is able to extract additional information on time variation in betas while being robust to the noise contaminations.

5.3 Consistent Estimation of the Dynamics of Betas

Given our empirical evidence suggesting multiple breaks in the betas, the question arises what drives these changes. One approach used in financial econometrics is to model their joint behaviour with (observable) macroeconomic fundamentals. For example, in the discrete-time (low-frequency) setting, Connor, Hagmann, and Linton (2012) estimate a Fama-French model where betas are functions of observable variables. Using daily data and the RV-beta estimator, Andersen et al. (2005) set up a state-space framework for betas and macroeconomic fundamentals, where the beta dynamics depends on the value of beta in the previous period as well as on the additional regressors. Authors estimate their model with the Kalman filter.

Following Andersen et al. (2005), suppose β follows a simple AR(p) model

$$\beta_i = \rho_1 \beta_{i-1} + \ldots + \rho_p \beta_{i-p} + \gamma' X_i + U_i, \quad i = p+1, \ldots, k, \tag{14}$$

where β_i is the value of financial beta in i^{th} time period (such as a week, a month, or a quarter), and the variables X_i could include low frequency macroeconomic or financial variables, as well as the intercept (we follow the notation of Section 3.1).

Since the true β_i is unobserved, in the empirical analysis it is replaced with its estimate $\widehat{\beta}_i$ for each i. This substitution leads to the problem of measurement errors in covariates due to the $\widehat{\beta}_{i-1}$ on the right-hand side of equation (14). Let $\varepsilon_i = \widehat{\beta}_i - \beta_i$ denote the difference between the estimated and the true β_i, then ε_i can be seen as the measurement error.

The presence of measurement errors in covariates $\widehat{\beta}_i$ means that we cannot use the standard OLS estimators of $\rho_1, \ldots, \rho_p$, and γ as they are biased and inconsistent. However, it is possible to account for the measurement errors and to provide consistent estimators of the parameters of interest, see, e.g., Andersen, Bollerslev, and Meddahi (2005) who adjust the forecasting loss functions.

We first introduce some additional notation. Let $\beta_{i,L} = (\beta_{i-1}, \ldots, \beta_{i-p})'$, $\widehat{\beta}_{i,L} = \left(\widehat{\beta}_{i-1}, \ldots, \widehat{\beta}_{i-p}\right)'$, $Z_i \equiv \left(\beta'_{i,L}, X'_i\right)'$, $\widehat{Z}_i \equiv \left(\widehat{\beta}'_{i,L}, X'_i\right)'$, $\rho = (\rho_1, \ldots, \rho_p)'$, $\theta = (\rho', \gamma')'$. Consider the infeasible OLS estimator

$$\widehat{\theta}^{Infeasible} = \left(\frac{1}{k-p-1} \sum_{i=p+1}^{k} Z_i Z'_i \right)^{-1} \frac{1}{k-p-1} \sum_{i=p+1}^{k} Z_i \beta_i.$$

This estimator is infeasible because β_i and $\beta_{i,L}$ are not observable. Replacing β_i and $\beta_{i,L}$ with $\widehat{\beta}_i$ and $\widehat{\beta}_{i,L}$ we obtain the feasible OLS estimator

$$\widehat{\theta}^{OLS} = \left(\frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{Z}'_i \right)^{-1} \frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{\beta}_i.$$

Let us consider the effect of replacing β_i with $\widehat{\beta}_i$. Remember that $\varepsilon_i = \widehat{\beta}_i - \beta_i \sim_a N\left(0, V_i/\tau^2_{n_i}\right)$, and that

$E[\varepsilon_i \varepsilon_j] = 0$ for $i \neq j$. Usually ε_j are also uncorrelated with X_i and hence we have

$$\frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{\beta}_i - \frac{1}{k-p-1} \sum_{i=p+1}^{k} Z_i \beta_i \xrightarrow{p} 0 \text{ as } k \quad \rightarrow \quad \infty,$$

$$\frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{Z}_i' - \frac{1}{k-p-1} \sum_{i=p+1}^{k} Z_i Z_i' - \Xi_k \xrightarrow{p} 0 \text{ as } k \quad \rightarrow \quad \infty,$$

$$\text{where } \Xi_k = \frac{1}{k-p-1} \sum_{i=p+1}^{k} diag \left(V_{i-1}/\tau^2_{n_{i-1}}, \dots, V_{i-p}/\tau^2_{n_{i-p}}, 0, \dots, 0 \right). \tag{15}$$

The feasible and infeasible OLS estimators differ because of the term Ξ_k. When this term is not negligible, we should correct the OLS estimator. The term Ξ_k can be estimated by $\widehat{\Xi}_k$, which is obtained by replacing V_i with their estimators $\widehat{V}_i$ in equation (15). Hence, in for the empirical analysis we can use the Measurement Error Corrected (MEC) estimator

$$\widehat{\theta}^{MEC} = \left(\frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{Z}_i' - \widehat{\Xi}_k \right)^{-1} \frac{1}{k-p-1} \sum_{i=p+1}^{k} \widehat{Z}_i \widehat{\beta}_i.$$

It is worth noting that in contrast to Kalman filter, this estimator does not need to assume that U_i (or ε_i) are homoscedastic, which is important in finance applications.

To illustrate the described estimators, for each of the six stocks we estimate model (14) with $p = 1$ and $p = 2$, where the betas are estimated by the TS-betas. We first consider the case with only the intercept ($X_i = 1$). Then, for illustration we consider $X_i = (1, VIX_{i-1})'$, where VIX_{i-1} is the (low frequency) CBOE VIX at the end of week $i-1$. For comparison, we include both the results for the inconsistent OLS estimator $\widehat{\theta}^{OLS}$ and for the corrected estimator $\widehat{\theta}^{MEC}$. Table E.11 presents the results.

Several comments are in order. The betas of all stocks appear to follow stationary mean-reverting processes, which is in line with the findings of Andersen et al. (2006). The measurement error corrections are substantial; the MEC estimates indicate that betas are more persistent than the OLS estimates (suffering from the attenuation bias) would suggest. This is important in practice, for example because well-known anomalies in finance may be sensitive to the specification of the dynamics of betas, see Avramov and Chordia (2006). The VIX does not appear to be an important determinant of the betas.[8] It is worth noting that the estimates for all of the considered stocks are qualitatively similar, which could be interpreted as an illustration of the robustness of the methodology.

Considering the linear model (14) allows us to obtain explicit expression for the estimator $\widehat{\theta}^{MEC}$. Note that we could also use the estimated $\widehat{V}_i$ to bias-correct the estimators in nonlinear models, although the estimators may not have simple analytic expressions.

[8] For the OLS estimator, VIX is not statistically significant for any of the stocks and specifications (using Newey and West (1987) estimator of standard errors). Also, ρ_2 is not significantly different from zero for all stocks except INTC. We do not present the standard errors for the OLS and MEC estimators since in the presence of leverage these would require an additional analysis that is beyond the scope of this paper (see Andersen, Bollerslev, and Meddahi (2005) for a discussion).

6 Conclusion

The current paper proves the validity of a multivariate subsampling method for inference on nonparametric estimators with high-frequency data. The subsampling method estimates the asymptotic variance-covariance matrices of multivariate high-frequency estimators without relying on the expression of these matrices. Our theoretical result for the validity of the multivariate subsampling is a nontrivial extension of the existing univariate results because we allow for the leverage effect in prices, as well as more general volatility dynamics. We study a data-driven bandwidth choice in a Monte Carlo study, and find that the performance of the subsampling method is very stable across wide ranges of the bandwidths. We compare the finite sample performance of the subsampling method with that of the plug-in method based on the expression of the asymptotic variance. We find that the performance of the subsampling method is much more stable across the TS parameters. Importantly, if the TS estimator is evaluated using the rule of Bandi and Russell (2011), which minimises its finite sample MSE, the subsampling estimator works much better than the plug-in rule. The underlying reason is the substantial difference between the asymptotic variance and the finite sample variance of the TS estimator for certain ranges of the TS parameter.

We apply the multivariate subsampling method to the analysis of time variation in equity betas for six stocks in 2006. We find strong evidence for variation in betas in 2006, and we identify multiple breaks across weekly betas for each of the six stocks considered. Our results suggest that tick data contains substantially more information on the time-variation in betas than data sampled at moderate frequencies such as five, ten, or twenty minutes. We use a simple dynamic model to capture this time-variation in betas. We then use the subsampled variances to correct for the errors-in-variables bias in this dynamic model. After the bias-correction, the betas appear to be substantially more persistent than the naive estimators suggest.

References

Aït-Sahalia, Y., P. Mykland, and L. Zhang (2011). Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise. Journal of Econometrics 160, 190-203.

Aldous, D. G. and G. K. Eagleson (1978). On Mixing and Stability of Limit Theorems. The Annals of Probability 6, 325-331.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42-55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. Econometrica 71, 579-625.

Andersen, T.G., T. Bollerslev, F. X. Diebold, and J. Wu (2005). A Framework for Exploring the Macroeconomic Determinants of Systematic Risk. American Economic Review, 95, 398-404.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and J. Wu (2006). Realized beta: persistence and predictability, in T. Fomby and D. Terrell (eds.) Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger , Volume B, 1-40

Andersen, T.G., T. Bollerslev, and N. Meddahi (2005). Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities. Econometrica 73, 275-29.

Andrews, D.W.K. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. Econometrica 61, 821-856.

Avramov, D. and T. Chordia (2006). Asset Pricing Models and Financial Market Anomalies. Review of Financial Studies 19, 1001-1040.

Bajgrowicz, P. and O. Scaillet (2011). Jumps in high-frequency data: spurious detections, dynamics, and news. Working Paper.

Bandi, F. and B. Perron (2006). Long memory and the relation between implied and realized volatility. Journal of Financial Econometrics 4, 636-670.

Bandi, F. and B. Perron (2008). Long-run risk-return trade-offs. Journal of Econometrics. 143, 349-374.

Bandi, F. M. and J. R. Russell (2005). Realized covariation, realized beta, and microstructure noise. Working paper, University of Chicago.

Bandi, F. M. and J. R. Russell (2008). Microstructure noise, realized variance, and optimal sampling. Review of Economic Studies 75, 339–369.

Bandi, F. M. and J. R. Russell (2011). Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Journal of Econometrics 160, 145-159.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and N. Shephard (2008). Designing Realised Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise. Econometrica, Vol. 76, No. 6, 1481-1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and N. Shephard (2009). Realised Kernels in Practice: Trades and Quotes. Econometrics Journal 12, C1-C32.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and N. Shephard (2011). Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-synchronous Trading. Journal of Econometrics 162, 149-169.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics, Vol. 72, No. 3, 885-925.

Bekaert, G. and G. Wu (2010). Asymmetric Volatility and Risk in Equity Markets. Review of Financial Studies 13, 1 - 41.

Benjamini, Y. and Y. Hochberg, (2005). Controlling the False Discovery Rate: a Practical and Powerful

Approach to Multiple Testing. J. R. Statist. Soc. B 57, No. 1, pp. 289-300.

Bollerslev, T. and H. O. Mikkelsen (1996). Modelling and pricing long memory in stock market volatility. Journal of Econometrics, Vol. 73, No. 1, 151-184.

Bollerslev, T. and B. Y. B. Zhang (2003). Measuring and modelling systematic risk in factor pricing models using high-frequency data. Journal of Empirical Finance 10, 533-558.

Braun, P. A., D. B. Nelson, and A. M. Sunier (1995). Good news, bad news, volatility, and betas. Journal of Finance 50, 1575-1603.

Christensen, K., S. Kinnebrock and M. Podolskij (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. Journal of Econometrics 159, 116-133.

Christensen, K., R. Oomen and M. Podolskij (2009). Realised Quantile-Based Estimation of the Integrated Variance. Journal of Econometrics 159, 74-98.

Connor, G., M. Hagmann, and O. Linton (2012). Efficient semiparametric estimation of the Fama-French model and extensions. Economerics 80, 713-756.

Cont, R. (2005). Long range dependence in financial markets. Fractals in engineering, 4, 159-179.

Comte, F. and E. Renault (1998). Long Memory in Continuous-Time Stochastic Volatility Models. Mathematical Finance, Vol. 8, No. 4, 291-323.

Ding, Z., Granger, W. J. and R. Engle (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance, 1, 83-106.

Donovon, P., S. Gonçalves, and N. Meddahi (2008). Bootstrapping realized multivariate volatility measures. Journal of Econometrics 172, 49-65.

Epps, T.W. (1979). Comovements in stock prices in the very short run. Journal of the American Statistical Association, 74, 291-298.

Giraitis, L. P. Kokoszka, R. Leipus, and G. Teyssière (2003). Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics, 112, 265-294.

Griffin, G. E. and R. C. Oomen (2005). Sampling Returns for Realized Variance Calculations: Tick Time or Transaction Time? Econometric Reviews, 27, 230-253.

Hansen, P. R., and A. Lunde (2006). Realized variance and market microstructure noise (with comments and rejoinder). Journal of Business and Economic Statistics 24, 127–218.

Hansen, P. R., A. Lunde, and V. Voev (2010). Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility and CoVolatility. Working paper, CREATES.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scand. J. Statist. 6, 65-70.

Jacod, J. (2008). Asymptotic properties or realized power variations and related functionals of semimartingales. Stochastic Processes and Their Applications 118, 517-559.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure Noise in the Continuous Case: The Pre-Averaging Approach. Stochastic Processes and Their Applications, 119, 2249-2276.

Jacod, J., M. Podolskij, and M. Vetter (2010). Limit Theorems for moving averages of discretized processes plus noise. Annals of Statistics 38, 1478-1545.

Jacod, J. and M. Rosenbaum (2012). Estimation of Volatility Functionals: the Case of $\sqrt{n}$ Window. Discussion paper, Université de Paris-6.

Jostova, G. and A. Philipov (2005). Bayesian Analysis of Stochastic Betas. Journal of Financial and Quantitative Analysis 40, 747-778.

Kalnina (2011). Subsampling High Frequency Data. Journal of Econometrics 161, 262-283.

Kalnina, I. and O. Linton (2008). Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. Journal of Econometrics 147, 47-59.

Kalnina, I. and K. Tewou (2015). Cross-sectional Dependence in Idiosyncratic Volatility. Discussion paper, Université de Montréal.

Kristensen, D. (2010). Nonparametric Filtering of the Realized Spot Volatility: A Kernel-based Approach. Econometric Theory 26, 60-93.

Lahiri, S. N., M. S. Kaiser, N. Cressie, and N. Hsu (1999). Prediction of Spatial Cumulative Distribution Functions Using Subsampling. Journal of the American Statistical Association, 94, 86-97

Lewellen, J. and S. Nagel (2006). The conditional CAPM does not explain asset-pricing anomalies. Journal of Financial Economics 82, 289-314.

Li, Y., Mykland, P.A., Renault, E., Zhang, L., and Zheng, X. (2014). Realized volatility when sampling times are possibly endogenous. Econometric Theory 30, 580-605.

Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps, Scandinavian Actuarial Journal 36, 270-296.

Mykland, P., and L. Zhang (2006). ANOVA for diffusions and Itô Processes. Annals of Statistics, 34, 1931-1963.

Mykland, P., and L. Zhang (2008).Inference for volatility-type objects and implications for hedgning. Statistics and Its Interface 1, 255-278..

Mykland, P., and L. Zhang (2009). Inference for Continuous Semimartingales Observed at High Frequency: A General Approach. Econometrica 77, 1403-1455.

Mykland, P., and L. Zhang (2014). Assessment of Uncertainty in High Frequency Data: The Observed Asymptotic Variance. Working Paper.

Newey, W.K., and K D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

Palandri, A. (2006). Consistent Realized Covariance for Asynchronous Observations Contaminated by Market Microstructure Noise. Working Paper.

Patton, A. and M. Verardo (2012). Does beta move with news? Firm-specific information flows and learning about profitability. Review of Financial Studies 25, 2789-2839.

Politis, D. N. and J. P. Romano (1994), "Large sample confidence regions based on subsamples under minimal assumptions. Annals of Statistics 22, 2031-2050.

Politis, D. N., J. P. Romano and M. Wolf (1999). *Subsampling*, Springer-Verlag, New York.

Ray, B. K. and R. Tsay (2000). Long-range dependence in daily stock volatility. Journal of Business and Economic Statistics 18, 254-262.

Romano, J. P. and M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Snooping. Econometrica 73, 1237-1282.

Rossi, E. and P. S. de Magistris (2014). Estimation of long memory in integrated variance. Econometric Reviews 33, 785-814.

Sizova, N. (2013). Long-Horizon Return Regressions with Historical Volatility. Journal of Business and Economic Statistics 31, 546-559.

Todorov, V and T. Bollserslev (2010). Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks. Journal of Econometrics 157, 220-235.

Todorov, V and G. Tauchen (2011). Limit Theorems for Power Variations of Pure-Jump Processes with Applications to Activity Estimation. Annals of Applied Probability 21, 546-235.

Vetter, M. (2011). Estimation of Integrated Volatility of Volatility with Applications to Goodness-of-fit Testing. Discussion paper, Ruhr-Universität, Bochum.

White, H.L. (2000). A Reality Check for Data Snooping. Econometrica 68, 1097-1126.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Bernoulli, 12 (6), 1019-1043.

Zhang, L. (2011). Estimating Covariation: Epps Effect, Microstructure Noise. Journal of Econometrics

160, 33-47.

Zhang, L. (2012). Implied and realized volatility: empirical model selection. Annals of Finance 8, 259-275.

Zhang, L., P. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association, 100, 1394-1411.

Zu, Y., and P. Boswijk (2015). Estimating Spot Volatility with High-Frequency Financial Data. Journal of Econometrics 181, 117-135.

A Proofs

A.1 Proof of Theorem 1

Let

$$\begin{aligned} \theta_l^{long} &= \int_{(l-1)m/n}^{lm/n} f\left(c_u\right) du, \quad \theta_l^{short} = \int_{(l-1)m/n}^{[(l-1)m+J]/n} f\left(c_u\right) du, \\ V_l^{long} &= \int_{(l-1)m/n}^{lm/n} g\left(c_u\right) du, \quad V_l^{short} = \int_{(l-1)m/n}^{[(l-1)m+J]/n} g\left(c_u\right) du. \end{aligned}$$

The following two equations are proved below in Sections A.1.1 and A.1.2, respectively,

$$V - \sum_{l=1}^{K} \frac{m}{J} V_l^{short} = o_p(1) \tag{16}$$

and

$$\frac{m}{J} \sum_{j=1}^{K} \tau_n^2 \left\| \theta_l^{short} - \frac{J}{m} \theta_l^{long} \right\|^2 = o_p(1). \tag{17}$$

Introduce the following notation,

$$\begin{aligned} \widehat{V}^{\text{infeasible}} &= \frac{1}{K} \sum_{l=1}^{K} \frac{n}{J} \tau_n^2 R_l^{short} \left(R_l^{short}\right)', \text{ where} \\ R_l^{short} &= \frac{n}{J} \widehat{\theta}_l^{short} - \frac{n}{J} \theta_l^{short}. \end{aligned}$$

Assumption A4 with $s = J$ and (16) imply

$$\widehat{V}^{\text{infeasible}} - V = o_p(1).$$

It remains to prove that

$$\widehat{V}^{sub} - \widehat{V}^{\text{infeasible}} = o_p(1). \tag{18}$$

We rewrite this difference in terms of three types of components, d_l, R_l^{short}, and R_l^{long}, where

$$\begin{aligned} d_l &= \frac{n}{m}\theta_l^{long} - \frac{n}{J}\theta_l^{short}, \\ R_l^{long} &= \frac{n}{m}\widehat{\theta}_l^{long} - \frac{n}{m}\theta_l^{long}. \end{aligned}$$

In particular, we can represent the differences in (18) as follows,

$$\begin{aligned} \widehat{V}^{sub} - \widehat{V}^{\text{infeasible}} &= \frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\Big(-R_l^{short}R_l^{long\prime} - R_l^{short}d_l' - R_l^{long}R_l^{short\prime} + R_l^{long}R_l^{long\prime} \\ &\quad + R_l^{long}d_l' - d_l R_l^{short\prime} + d_l R_l^{long\prime} + d_l' d_l'\Big). \end{aligned} \tag{19}$$

If we can show that

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 R_l^{short}R_l^{short\prime} = O_p(1), \tag{20}$$

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 R_l^{long}R_l^{long\prime} = o_p(1), \tag{21}$$

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 d_l d_l' = o_p(1), \tag{22}$$

then all terms in (19) are negligible by the Cauchy-Schwarz inequality, e.g.,

$$\left|\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\left(R_l^{short}d_l'\right)\right| \leq \sqrt{\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\left(R_l^{short}R_l^{short\prime}\right)}\sqrt{\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\left(d_l d_l'\right)} = o_p(1).$$

Equation (20) follows from $\widehat{V}^{\text{infeasible}} = V + o_p(1) = O_p(1)$, equation (21) follows by Assumption A4 with $s = m$ together with equality $V = \sum_{l=1}^{K} V_l^{long}$, and equation (22) follows by (17). □

A.1.1 Proof of equation (16)

Let $I_{a,b} = \int_a^b \|c_u - c_a\|\, du$ and $I_{a,b,t} = \int_a^b \|c_u - c_t\|\, du$. Consider a general matrix valued function $\eta(c) : \mathbb{R}^{d\times d} \to \mathbb{R}^{q_1\times q_2}$, and let

$$H_{a,b}^{\eta} = \int_a^b \eta(c_u)\, du \text{ for any } 0 \leq a \leq b \leq 1.$$

We prove the following Lemma at the end of this Section,

Lemma 4. *Suppose φ is a functional and is continuously differentiable. Let $A_n \leq a_n < b_n \leq B_n$ be sequences with $B_n - A_n \to 0$ as $n \to \infty$. Let $\overline{\varphi_\nabla} = \sum_{k_1=1}^{d} \sum_{k_2=1}^{d} \sup_{t\in[0,1]} |\partial\varphi(c_t)/\partial c_{t,k_1k_2}|$. Then*

$$\left| \frac{1}{B_n - A_n} H^{\varphi}_{A_n,B_n} - \frac{1}{b_n - a_n} H^{\varphi}_{a_n,b_n} \right| \leq \overline{\varphi_\nabla} \times \left\{ \frac{I_{A_n,a_n,a_n} + I_{a_n,B_n}}{B_n - A_n} + \frac{I_{a_n,b_n}}{b_n - a_n} \right\}.$$

To prove equation (16), consider any k_1, $k_2 \in \{1, \ldots, d\}$ and let $\varphi(\cdot)$ denote function $g_{k_1k_2}(\cdot)$. We have

$$\begin{aligned} V_{k_1k_2} - \sum_{l=1}^{K} \frac{m}{J} V^{short}_{l,k_1k_2} &= \sum_{l=1}^{K} \left(V^{long}_{l,k_1k_2} - \frac{m}{J} V^{short}_{l,k_1k_2} \right) \\ &= \frac{m}{n} \sum_{l=1}^{K} \left(\frac{1}{m/n} H^{\varphi}_{(l-1)m/n,lm/n} - \frac{1}{J/n} H^{\varphi}_{(l-1)m/n,[(l-1)m+J]/n} \right). \end{aligned}$$

Applying Lemma 4 we obtain

$$\begin{aligned} \left| V_{k_1k_2} - \sum_{l=1}^{K} \frac{m}{J} V^{short}_{l,k_1k_2} \right| &\leq \frac{m}{n} \sum_{l=1}^{K} \left| \frac{1}{m/n} H^{\varphi}_{(l-1)m/n,lm/n} - \frac{1}{J/n} H^{\varphi}_{(l-1)m/n,[(l-1)m+J]/n} \right| \\ &\leq \overline{\varphi_\nabla} \frac{m}{n} \sum_{l=1}^{K} \left\{ \frac{I_{(l-1)m/n,lm/n}}{m/n} + \frac{I_{(l-1)m/n,[(l-1)m+J]m/n}}{J/n} \right\}. \end{aligned} \quad (23)$$

Assumptions A1 and A2 imply $\overline{\varphi_\nabla} = O_P(1)$. Using Assumption A2 and Holder inequality we have $E[\|c_{t_1} - c_{t_2}\|] \leq B_c^{1/2} |t_1 - t_2|^{\alpha/2}$, so for all l, m, n we have

$$\frac{1}{m/n} E\left[I_{(l-1)m/n,lm/n}\right] \leq B_c^{1/2} (m/n)^{\alpha/2} \text{ and } \frac{1}{J/n} E\left[I_{(l-1)m/n,[(l-1)m+J]m/n}\right] \leq B_c^{1/2} (J/n)^{\alpha/2}.$$

Hence

$$E\left[\frac{m}{n} \sum_{l=1}^{K} \left\{ \frac{I_{(l-1)m/n,lm/n}}{m/n} + \frac{I_{(l-1)m/n,[(l-1)m+J]m/n}}{J/n} \right\} \right] \leq B_c^{1/2} K \frac{m}{n} \left(\left(\frac{m}{n}\right)^{\alpha/2} + \left(\frac{J}{n}\right)^{\alpha/2} \right) = o(1),$$

as long as $J/n \to 0$ and $m/n \to 0$, since $K\frac{m}{n} \to 1$. Using Markov inequality and equation (23) we obtain $V_{k_1k_2} - \sum_{l=1}^{K} \frac{m}{J} V^{short}_{l,k_1k_2} = o_P(1)$ and hence equation (16) holds. □

Proof of Lemma 4: We omit subscript n and replace A_n, a_n, b_n, B_n with A, a, b, B, respectively.

$$\begin{aligned}
& \left| H_{A,B} - \frac{B-A}{b-a} H_{a,b} \right| \\
= & \left| \int_A^B \varphi(c_u)\, du - \frac{B-A}{b-a} \int_a^b \varphi(c_u)\, du \right| \\
= & \left| \int_A^B (\varphi(c_u) - \varphi(c_a))\, du - \frac{B-A}{b-a} \int_a^b (\varphi(c_u) - \varphi(c_a))\, du \right| \\
\leq & \int_A^B |\varphi(c_u) - \varphi(c_a)|\, du + \frac{B-A}{b-a} \int_a^b |\varphi(c_u) - \varphi(c_a)|\, du \\
\leq & \overline{\varphi_\nabla} \left\{ \int_A^B \|c_u - c_a\|\, du + \frac{B-A}{b-a} \int_a^b \|c_u - c_a\|\, du \right\} \\
\leq & (B-A)\, \overline{\varphi_\nabla} \left\{ \frac{I_{A,a,a} + I_{a,B}}{B-A} + \frac{I_{a,b}}{b-a} \right\}. \ \square
\end{aligned}$$

A.1.2 Proof of equation (17)

Fix any k and let $\varphi(\cdot)$ denote $f_k(\cdot)$. Then, using Lemma 4 in Section A.1.1,

$$\begin{aligned}
\sum_{j=1}^{K} \left(\theta_{l,k}^{short} - \frac{J}{m} \theta_{l,k}^{long} \right)^2 &= \left(\frac{J}{n} \right)^2 \sum_{j=1}^{K} \left(\frac{1}{J/n} \theta_{l,k}^{short} - \frac{1}{m/n} \theta_{l,k}^{long} \right)^2 \\
&\leq \left(\frac{J}{n} \right)^2 \sum_{j=1}^{K} \overline{\varphi_\nabla}^2 \left(\frac{I_{(l-1)m/n, lm/n}}{m/n} + \frac{I_{(l-1)m/n, [(l-1)m+J]m/n}}{J/n} \right)^2 \\
&\leq 2\overline{\varphi_\nabla}^2 \left(\frac{J}{n} \right)^2 \sum_{j=1}^{K} \left(\frac{I^2_{(l-1)m/n, lm/n}}{(m/n)^2} + \frac{I^2_{(l-1)m/n, [(l-1)m+J]m/n}}{(J/n)^2} \right).
\end{aligned}$$

To bound the sum, notice that by Cauchy-Schwarz inequality and Assumption A2 for any $a < b$

$$\frac{1}{(b-a)^2} E\left[I_{a,b}^2 \right] = \frac{1}{(b-a)^2} E\left[\left(\int_a^b \|c_u - c_a\|\, du \right)^2 \right] \leq \frac{1}{(b-a)^2} E\left[(b-a) \int_a^b \|c_u - c_a\|^2\, du \right] \leq B_c (b-a)^\alpha.$$

Thus,

$$E\left[\left(\frac{J}{n}\right)^2\sum_{j=1}^{K}\left(\frac{I^2_{(l-1)m/n,lm/n}}{(m/n)^2}+\frac{I^2_{(l-1)m/n,[(l-1)m+J]m/n}}{(J/n)^2}\right)\right]$$
$$\leq B_c\frac{J^2}{n^2}K\left(\left(\frac{m}{n}\right)^\alpha+\left(\frac{J}{n}\right)^\alpha\right)\leq 2B_c\frac{J^2}{nm}\left(\frac{m}{n}\right)^\alpha.$$

Hence, we can use Markov inequality and $\overline{\varphi_\nabla}=O_P(1)$ to obtain

$$\frac{m}{J}\sum_{j=1}^{K}\tau_n^2\left|\theta^{short}_{l,k}-\frac{J}{m}\theta^{long}_{l,k}\right|^2=O_P\left(\frac{m}{J}\tau_n^2\frac{J^2}{nm}\left(\frac{m}{n}\right)^\alpha\right)=O_P\left(\tau_n^2\frac{Jm^\alpha}{n^{1+\alpha}}\right),$$

so equation (17) follows from Assumption A3. □

B Proof of Lemma 3

The proof follows closely Comte and Renault (1998). The process $x(t)=\frac{1}{2}\ln c_t$ can also be written as $x(t)=\int_0^t a(t-s)dW(s)$ with

$$a(x)=\frac{\gamma}{\Gamma(1+\overline{\alpha})}\left(x^{\overline{\alpha}}-\kappa e^{-\kappa x}\int_0^x e^{\kappa u}u^{\overline{\alpha}}du\right)$$

and $W(s)$ a standard Brownian Motion. Let $t_1\leq t_2$. We have

$$\begin{aligned}
& E\left(c_{t_2}-c_{t_1}\right)^2\\
= & E\left(\exp(2x(t_2))-\exp(2x(t_1))\right)^2\\
= & e^{8\int_0^{t_1}a^2(x)dx}+e^{8\int_0^{t_2}a^2(x)dx}-2e^{2\int_0^{t_1}a^2(x)dx+2\int_0^{t_2}a^2(x)dx+4\int_0^{t_1}a(x)(a(t_2-t_1+x))dx}\\
= & e^{8\int_0^{t_2}a^2(x)dx}\left(1+e^{-8\int_{t_1}^{t_2}a^2(x)dx}-2e^{-6\int_0^{t_1}a^2(x)dx-4\int_0^{t_1}a(x)(a(x)-a(t_2-t_1+x))dx}\right)\\
\leq & 2e^{8\int_0^{t_2}a^2(x)dx}\left(1-e^{-6\int_{t_1}^{t_2}a^2(x)dx-4\int_0^{t_1}a(x)(a(x)-a(t_2-t_1+x))dx}\right).
\end{aligned}$$

The term inside the last parenthesis is necessarily nonnegative, and the term in the last exponential is nonpositive. Moreover $|\int_{t_1}^{t_2}a^2(x)dx|\leq M_1^2|t_2-t_1|$ with $M_1=\sup_{x\in[0,1]}|a(x)|$, and since a is $\overline{\alpha}$-Hölder,

$$\left|\int_0^{t_1}a(x)(a(x)-a(t_2-t_1+x))dx\right|\leq C_{\overline{\alpha}}\left|t_2-t_1\right|^{\overline{\alpha}}\int_0^{t_1}\left|a(x)\right|dx\leq C_{\overline{\alpha}}\left|t_2-t_1\right|^{\overline{\alpha}}M_1,$$

we have

$$\left|\int_{t_1}^{t_2}a^2(x)dx+\int_0^{t_2}a(x)\left(a(x)-a(t_2-t_1+x)\right)dx\right|\leq M_2\left|t_2-t_1\right|^{\overline{\alpha}}.$$

Finally, use $\forall u\leq 0,\ \ 0\leq 1-e^u\leq|u|$, to conclude Assumption A2 with $\alpha=\overline{\alpha}$. □

C Proof of Corollary 2

We prove Corollary 2 by verifying the assumptions of Theorem 1. Assumption A1 is clearly satisfied with $f(u) = c_u$ and $g(u)$ such that its $(1,1)$ element is

$$g_{11}(u) = \varphi^{TS}\frac{4}{3}c_{u,11}^2 + 8(\varphi^{TS})^{-2}Var\left(\epsilon^F\right)^2 + 16(\varphi^{TS})^{-2}\lim_{n\to\infty}\sum_{i=1}^{n}Cov\left(\epsilon_0^F, \epsilon_{i/n}^F\right)^2, \tag{24}$$

and other elements of $g(u)$ defined similarly (see references in the main text). Assumption A2 is directly assumed in the statement of Corollary (sufficient conditions are discussed in the main text). Assumption A3 is the restriction on the subsample sizes, that is also directly assumed by the statement of the Corollary. Assumption A4 is verified by the calculations similar to Lemma 7 of Kalnina (2011); it makes use of Assumption N. □

E Figures and Tables for the Empirical Results

	trans./week	$\widehat{\omega}^2 \cdot 10^7$	$\widetilde{\omega}^2 \cdot 10^7$	$\widehat{\xi} \cdot 10^5$	acf(1)	acf(2)	acf(3)
AIG	18,029	0.207	0.136	0.156	-0.320	0.102	-0.014
GE	29,015	0.228	0.188	0.189	-0.582	0.248	-0.118
IBM	20,070	0.162	0.095	0.117	-0.302	0.081	0.008
INTC	35,267	0.518	0.407	0.127	-0.525	0.200	-0.085
MMM	14,005	0.284	0.123	0.121	-0.269	0.092	0.006
MSFT	32,421	0.338	0.282	0.178	-0.555	0.224	-0.100
SPY	39,801	0.037	0.018	0.048	-0.352	0.065	0.006

Table E.1: Summary statistics of data before the synchronization. First column contains average number of transactions per week. Second and third columns contains variance of the noise estimates over the whole year 2006, $\widehat{\omega}^2 = RV/2n$, $\widetilde{\omega}^2 = (RV - \widehat{IV})/2n$ where IV is estimated by the TSRV; n is the total number of transactions in 2006 for the corresponding stock. Fourth column contains estimated noise-to-signal ratio, $\widehat{\xi} = \widehat{\omega}^2/\widehat{IV}$. Last three columns contain autocorrelation functions of returns at first, second, and third lag.

	trans./week	$\widehat{\omega}^2 \cdot 10^7$	$\widetilde{\omega}^2 \cdot 10^7$	$\widehat{\xi} \cdot 10^5$	acf(1)	acf(2)	acf(3)
AIG(SPY)	15,425	0.220	0.138	0.282	-0.15	0.051	0.02
GE(SPY)	21,819	0.229	0.176	0.295	-0.221	0.058	0.015
IBM(SPY)	16,890	0.174	0.095	0.223	-0.166	0.052	0.021
INTC(SPY)	24,601	0.545	0.384	0.700	-0.247	0.060	0.016
MMM(SPY)	12,315	0.303	0.121	0.389	-0.114	0.048	0.014
MSFT(SPY)	23,322	0.347	0.267	0.451	-0.238	0.061	0.017
SPY(AIG)	15,425	0.059	0.011	0.045	-0.276	0.084	-0.006
SPY(GE)	21,819	0.049	0.014	0.040	-0.509	0.173	-0.059
SPY(IBM)	16,890	0.056	0.011	0.040	-0.257	0.069	0.011
SPY(INTC)	24,601	0.045	0.014	0.011	-0.439	0.132	-0.041
SPY(MMM)	12,315	0.071	0.011	0.031	-0.232	0.082	0.013
SPY(MSFT)	23,322	0.046	0.014	0.024	-0.476	0.155	-0.051
AIG(joint)	6,957	0.037	0.018	0.273	-0.111	0.010	-0.007
GE(joint)	6,957	0.032	0.015	0.262	-0.218	0.028	0.005
IBM(joint)	6,957	0.032	0.013	0.228	-0.08	0.010	-0.003
INTC(joint)	6,957	0.094	0.037	0.227	-0.174	0.002	-0.006
MMM(joint)	6,957	0.046	0.014	0.197	-0.103	0.032	0.003
MSFT(joint)	6,957	0.054	0.027	0.277	-0.199	0.013	-0.004
SPY(joint)	6,957	0.011	0.001	0.145	-0.014	0.028	0.010

Table E.2: Summary statistics of the data after the synchronization. The notation "AIG(SPY)" means stock AIG after it has been synchronized with SPY. By construction, number of transactions of AIG(SPY) is the same as that of SPY(AIG). AIG(joint) means stock AIG after it has been synchronised with the other 6 series. See Table E.1 annotation for the meaning of the other column entries.

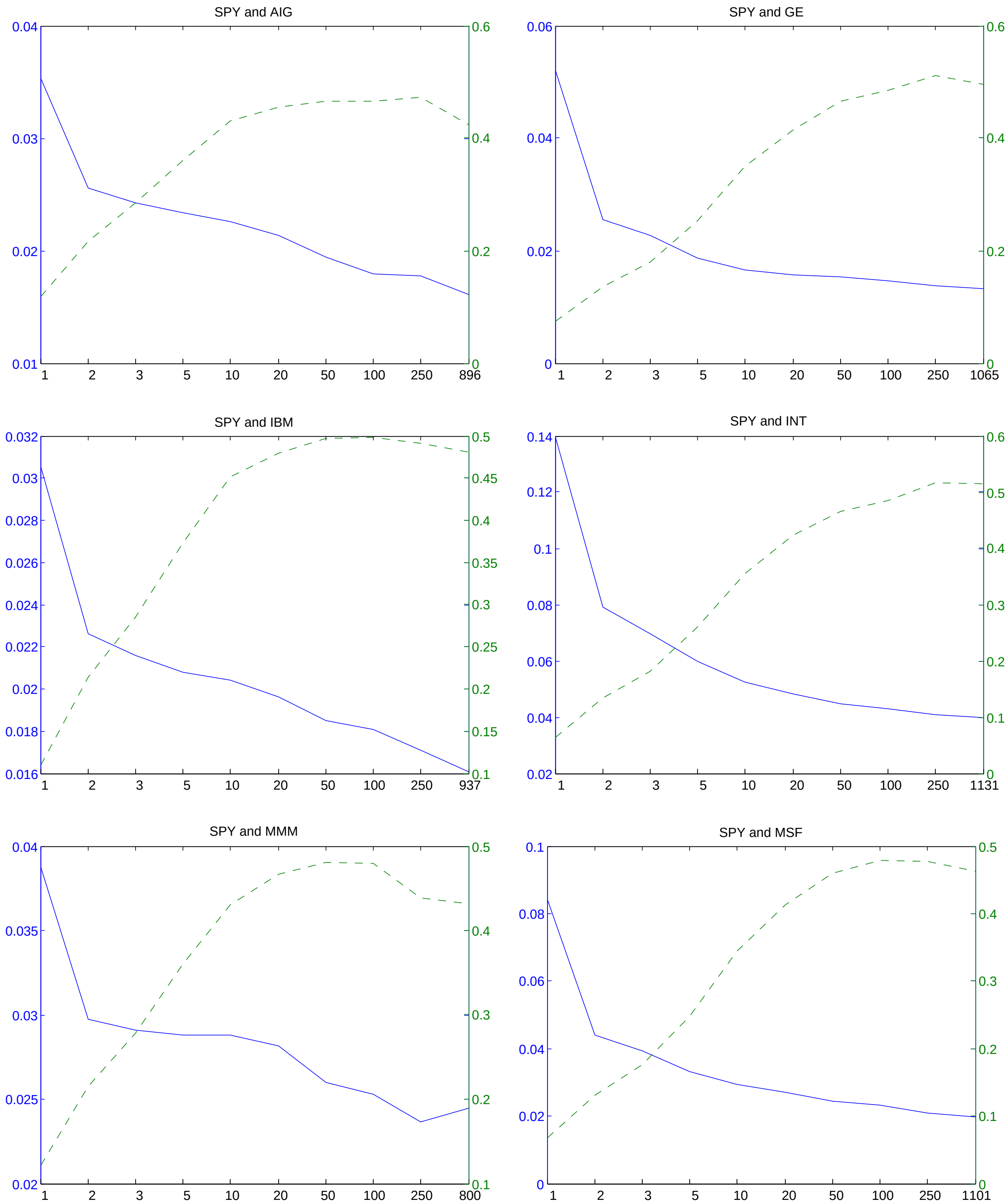


Figure E.3: The solid lines (left y-axis) are the volatility signature plots, i.e., realized variance plotted against the frequency (in ticks) used in its calculation. Dashed lines (right y-axis) are the realized correlation plots against the frequency (in ticks). Data covers the whole year 2006.

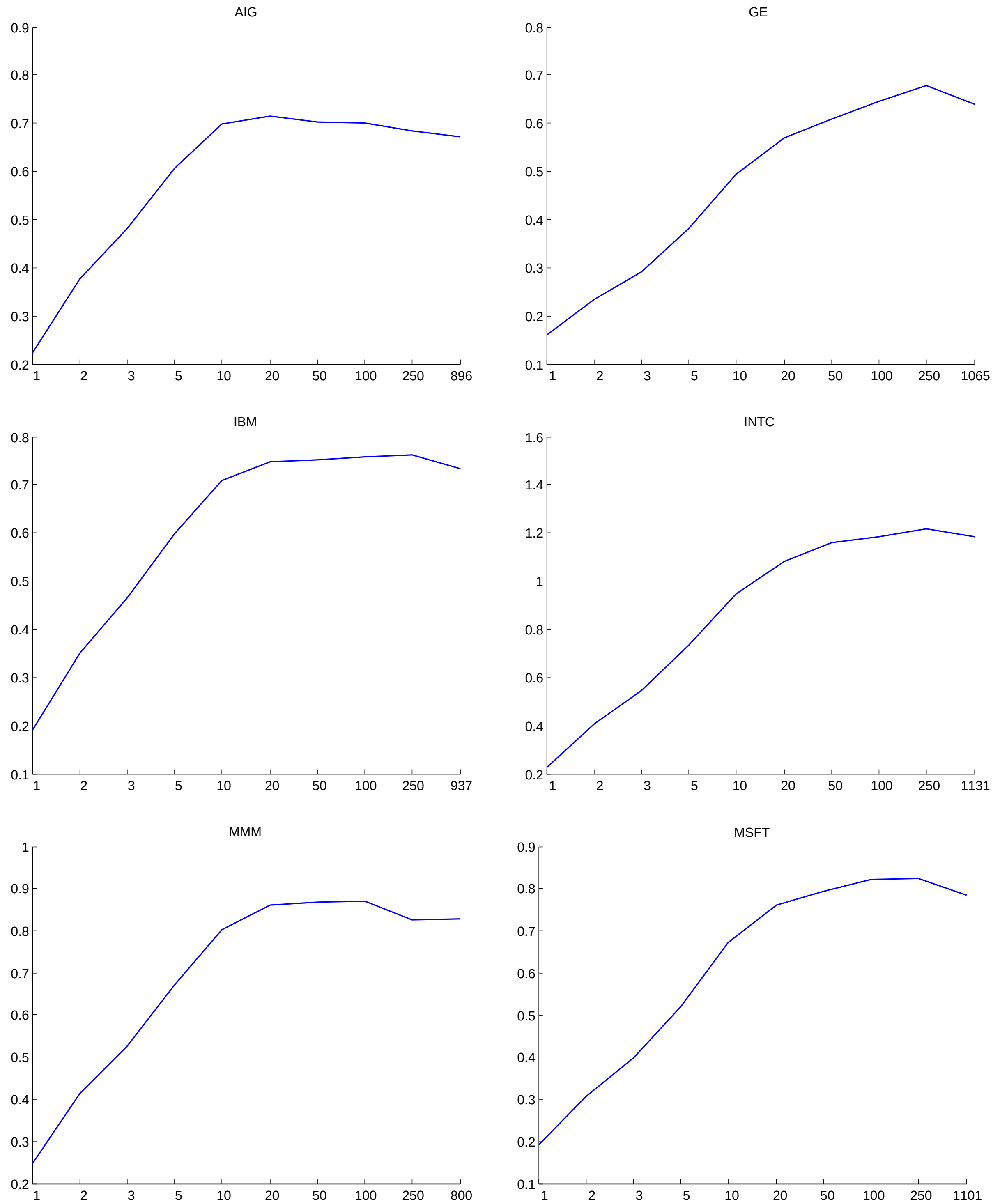
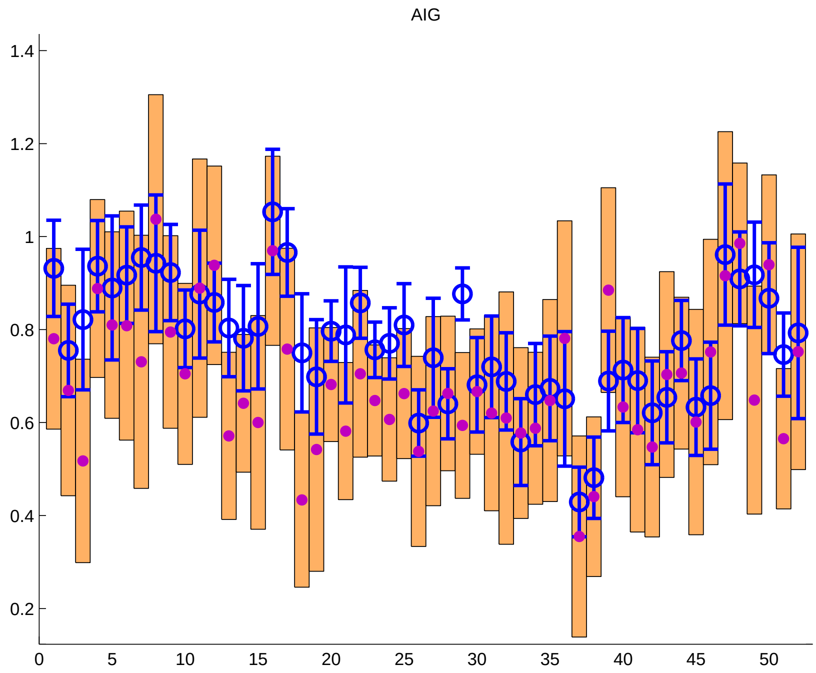


Figure E.4: Average weekly betas of individual stocks against the frequency (in ticks) used in their calculation. Data year 2006.



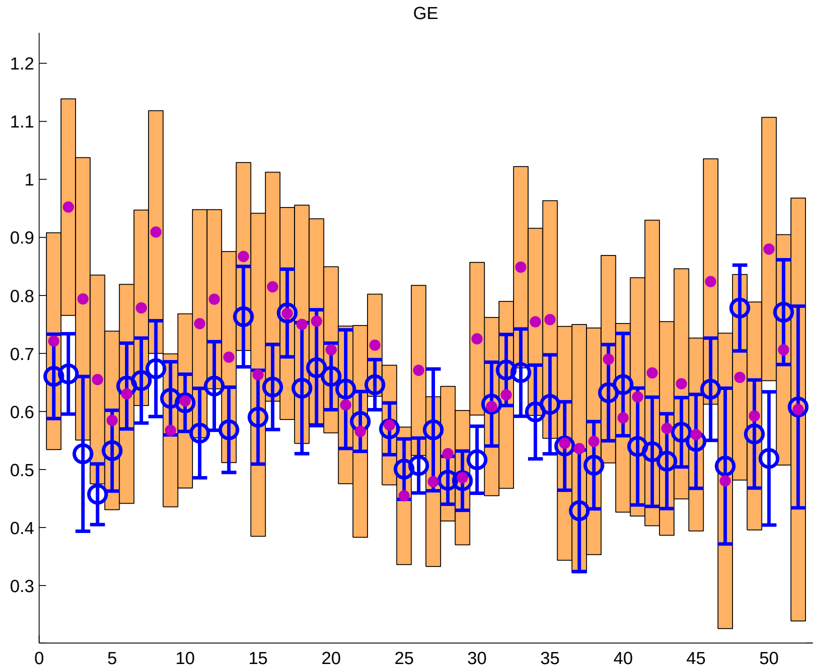


Figure E.5: Estimated betas for AIG and GE with 95% confidence intervals. Filled dots with rectangular CIs correspond to $\widehat{\beta}^{RV}_{10min}$*, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{TS}$*. Weeks on the x-axis.*

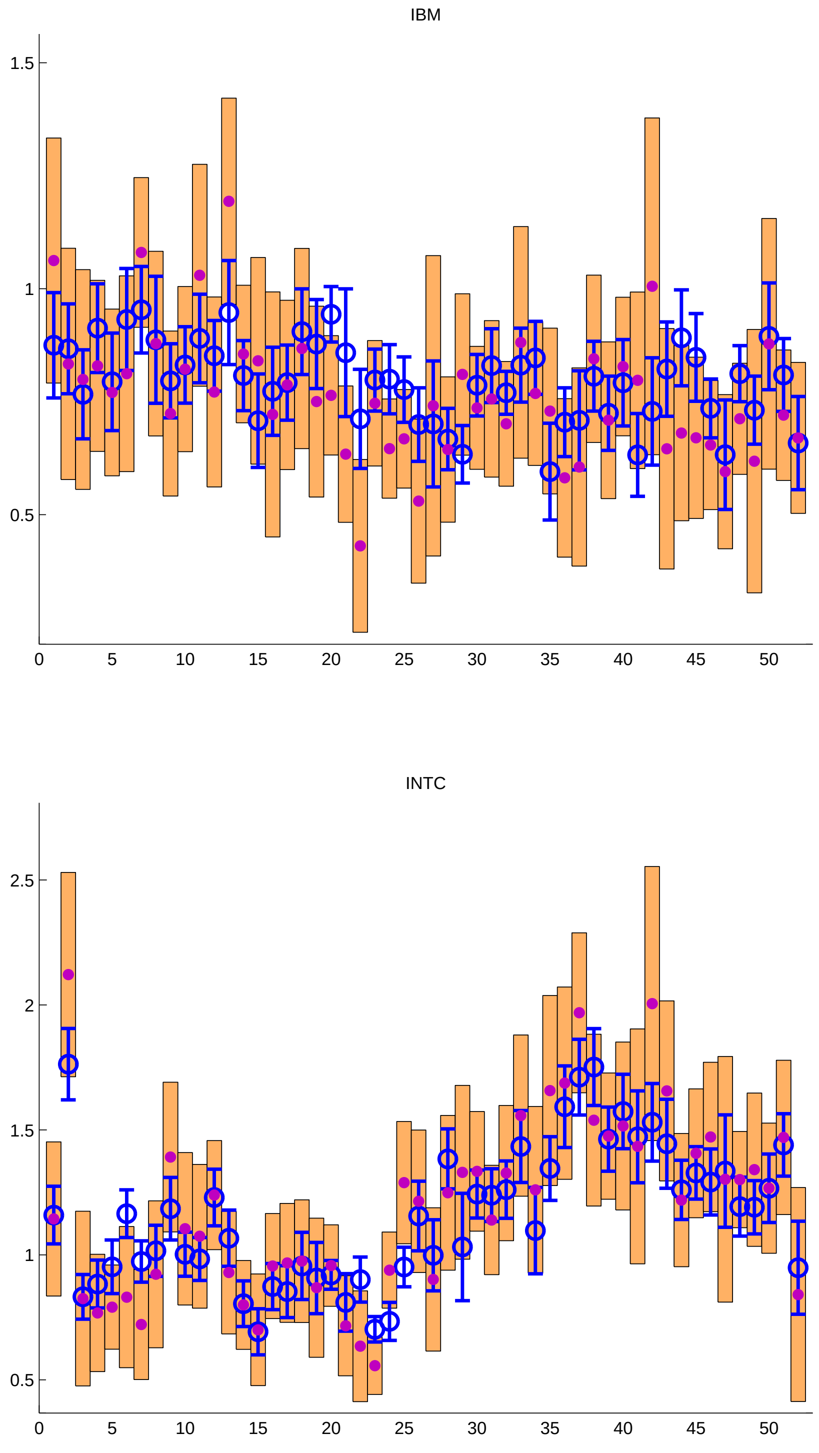
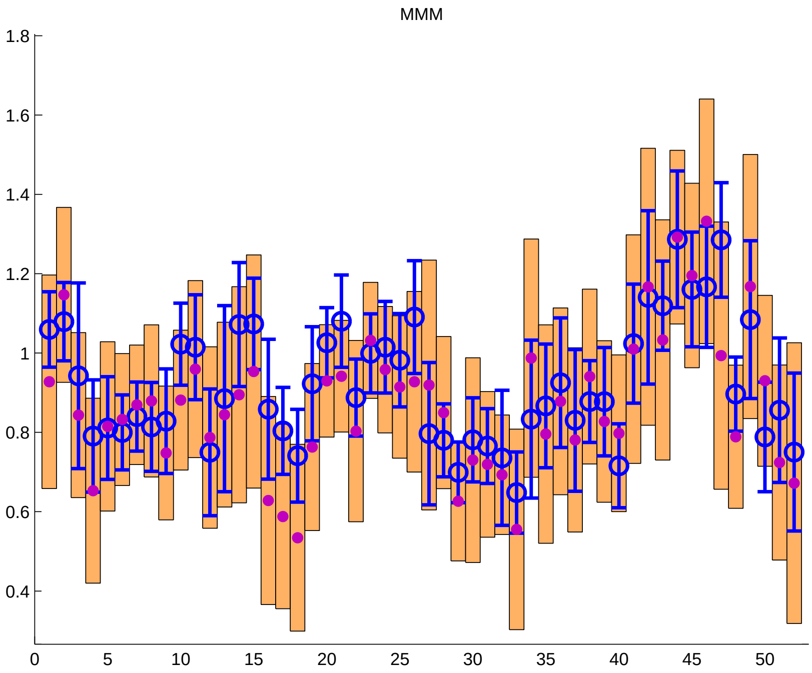


Figure E.6: Estimated betas for IBM and INTC with 95% confidence intervals. Filled dots with rectangular CIs correspond to $\widehat{\beta}^{RV}_{10min}$*, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{TS}$*. Weeks on the x-axis.*



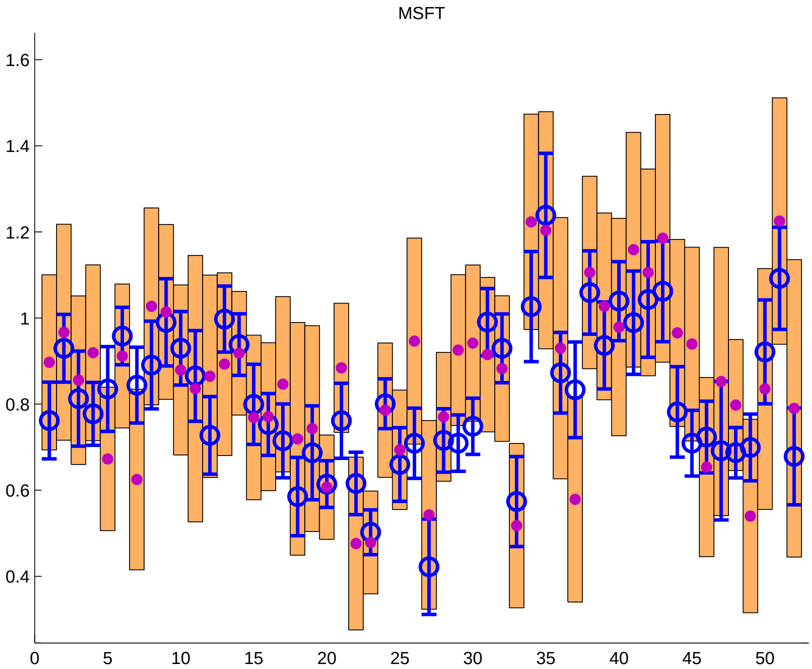


Figure E.7: Estimated betas for MMM and MSFT with 95% confidence intervals. Filled dots with rectangular CIs correspond to $\widehat{\beta}^{RV}_{10min}$*, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{TS}$*. Weeks on the x-axis.*

	2006		Q1		Q2		Q3		Q4	
				Test based on $\widehat{\beta}^{RV}_{5min}$						
AIG	177.32	(0)	20.37	(0.06)	35.09	(0)	46.79	(0)	41.43	(0)
GE	125.94	(0)	12.38	(0.416)	29.49	(0.003)	35.44	(0)	19.73	(0.072)
IBM	109.66	(0)	16.81	(0.157)	27.84	(0.006)	26.07	(0.011)	14.79	(0.253)
INTC	457.06	(0)	75.57	(0)	62.42	(0)	53.76	(0)	35.38	(0)
MMM	183.03	(0)	25.73	(0.012)	38.28	(0)	22.82	(0.029)	51.27	(0)
MSFT	282.25	(0)	19.57	(0.076)	97.53	(0)	87.92	(0)	46.59	(0)
				Test based on $\widehat{\beta}^{RV}_{10min}$						
AIG	101.53	(0)	19.77	(0.071)	19.37	(0.08)	18.26	(0.108)	26.23	(0.01)
GE	98.04	(0)	21.89	(0.039)	29.30	(0.004)	26.84	(0.008)	11.31	(0.503)
IBM	94.37	(0)	22.80	(0.029)	23.04	(0.027)	9.21	(0.685)	10.32	(0.588)
INTC	354.08	(0)	60.16	(0)	56.12	(0)	36.89	(0)	17.61	(0.128)
MMM	109.51	(0)	13.01	(0.368)	26.34	(0.01)	14.62	(0.263)	33.46	(0.001)
MSFT	171.86	(0)	15.58	(0.211)	43.02	(0)	50.03	(0)	33.49	(0.001)
				Test based on $\widehat{\beta}^{RV}_{20min}$						
AIG	79.93	(0.006)	26.59	(0.009)	12.66	(0.394)	9.68	(0.644)	18.99	(0.089)
GE	93.09	(0)	18.93	(0.09)	21.75	(0.04)	19.71	(0.073)	18.80	(0.093)
IBM	98.63	(0)	14.20	(0.288)	29.32	(0.004)	12.89	(0.377)	9.40	(0.668)
INTC	261.08	(0)	64.63	(0)	44.82	(0)	13.34	(0.345)	22.26	(0.035)
MMM	82.87	(0.003)	14.27	(0.284)	14.50	(0.27)	13.41	(0.34)	23.12	(0.027)
MSFT	104.92	(0)	21.23	(0.047)	19.97	(0.068)	24.02	(0.02)	18.53	(0.101)
				Test based on $\widehat{\beta}^{TS}$						
AIG	333.65	(0)	17.59	(0.129)	59.45	(0)	114.00	(0)	41.42	(0)
GE	238.62	(0)	44.07	(0)	69.14	(0)	54.92	(0)	51.61	(0)
IBM	178.97	(0)	14.08	(0.295)	37.63	(0)	43.27	(0)	35.83	(0)
INTC	1111.55	(0)	153.92	(0)	83.24	(0)	106.38	(0)	49.51	(0)
MMM	268.93	(0)	47.89	(0)	36.55	(0)	17.07	(0.147)	82.21	(0)
MSFT	623.50	(0)	45.71	(0)	128.26	(0)	178.41	(0)	121.23	(0)

Table E.8: Values of the Chi-square test; corresponding p-values in parenthesis. The null hypothesis is that true betas are constant over the some time interval. The top row indicates the corresponding time interval.

	$\widehat{\beta}^{RV}_{5min}$		$\widehat{\beta}^{RV}_{10min}$		$\widehat{\beta}^{RV}_{20min}$		$\widehat{\beta}^{TS}$	
2006	937.6	(0)	700.8	(0)	545.2	(0)	3286.8	(0)
Q1	107.2	(0.005)	98.1	(0.022)	85.2	(0.138)	445.7	(0)
Q2	147.1	(0)	114.9	(0.001)	106.9	(0.005)	289.2	(0)
Q3	296.8	(0)	222.5	(0)	140.6	(0)	671.3	(0)
Q4	172.4	(0)	145.9	(0)	154.7	(0)	542.9	(0)

Table E.9: Values of the joint Chi-square test (see section 3.1); corresponding p-values in parenthesis. The null hypothesis is that true betas for all 6 stocks are constant over a particular time interval, which is indicated in the first column. First three methods (labelled $\widehat{\beta}^{RV}_{5min}$, $\widehat{\beta}^{RV}_{10min}$, and $\widehat{\beta}^{RV}_{20min}$) are based on the realized covariance and the estimator of Barndorff-Nielsen and Shephard (2004) of its asymptotic variance; for the last column, Two Scale method is used for point estimates of betas, and subsampling method is used to estimate their asymptotic variance-covariance matrices.

		AIG	*GE*	*IBM*	*INTC*	*MMM*	*MSFT*
$\widehat{\beta}^{RV}_{5min}$	α = *0.01*	1 (1,0,0)	0 (0,0,0)	0 (0,0,0)	2 (2,0,0)	0 (0,0,0)	4 (4,0,0)
	α = *0.05*	2 (2,0,0)	0 (0,0,0)	0 (0,0,0)	3 (3,0,0)	1 (1,0,0)	6 (6,0,0)
	α = *0.10*	2 (2,0,0)	0 (0,0,0)	0 (0,0,0)	5 (5,0,0)	1 (1,0,0)	6 (6,0,0)
$\widehat{\beta}^{RV}_{10min}$	α = *0.01*	0 (0,0,0)	0 (0,0,0)	0 (0,0,0)	3 (3,0,0)	0 (0,0,0)	1 (1,0,0)
	α = *0.05*	0 (0,0,0)	0 (0,0,0)	0 (0,0,0)	3 (3,0,0)	0 (0,0,0)	1 (1,0,0)
	α = *0.10*	1 (1,0,0)	0 (0,0,0)	0 (0,0,0)	3 (3,0,0)	0 (0,0,0)	4 (4,0,0)
$\widehat{\beta}^{RV}_{20min}$	α = *0.01*	0 (0,0,0)	0 (0,0,0)	0 (0,0,0)	2 (2,0,0)	0 (0,0,0)	0 (0,0,0)
	α = *0.05*	0 (0,0,0)	0 (0,0,0)	0 (0,0,0)	2 (2,0,0)	0 (0,0,0)	1 (1,0,0)
	α = *0.10*	0 (0,0,0)	0 (0,0,0)	0 (0,0,0)	3 (3,0,0)	0 (0,0,0)	1 (1,0,0)
$\widehat{\beta}^{TS}$	α = *0.01*	2 (2,0,0)	3 (2,1,0)	1 (1,0,0)	8 (7,1,0)	1 (1,0,0)	10 (10,0,0)
	α = *0.05*	4 (4,0,0)	4 (4,0,0)	2 (2,0,0)	9 (8,1,0)	2 (2,0,0)	12 (12,0,0)
	α = *0.10*	5 (5,0,0)	5 (5,0,0)	2 (2,0,0)	14 (13,1,0)	2 (2,0,0)	14 (13,1,0)

Table E.10: Number of the detected breaks in weekly betas during 2006, i.e., number of the rejected null hypotheses out of 51. See Section 3.2 for the description of the test. Numbers in parenthesis are the breakdown of the total detected breaks across first, second, and third step of testing; and α is the significance level. First three methods (labelled $\widehat{\beta}^{RV}_{5min}$, $\widehat{\beta}^{RV}_{10min}$, and $\widehat{\beta}^{RV}_{20min}$) are based on the realized covariance and the estimator of Barndorff-Nielsen and Shephard (2004) of its asymptotic variance; in the last three rows, the Two Scale method is used for point estimates of betas, and subsampling method is used to estimate their asymptotic variance-covariance matrices.

Stock	ρ_1^{OLS}	ρ_2^{OLS}	γ_1^{OLS}	γ_{VIX}^{OLS}	ρ_1^{MEC}	ρ_2^{MEC}	γ_1^{MEC}	γ_{VIX}^{MEC}
AIG	0.60		0.31		0.73		0.21	
AIG	0.56	0.09	0.27		0.81	-0.08	0.20	
AIG	0.59		0.34	-0.00	0.73		0.23	-0.00
AIG	0.56	0.08	0.30	-0.00	0.82	-0.08	0.24	-0.00
GE	0.24		0.45		0.32		0.41	
GE	0.21	0.05	0.43		0.29	0.05	0.39	
GE	0.20		0.60	-0.01	0.27		0.56	-0.01
GE	0.18	0.05	0.58	-0.01	0.25	0.04	0.54	-0.01
IBM	0.39		0.48		0.57		0.34	
IBM	0.37	0.03	0.47		0.61	-0.09	0.38	
IBM	0.39		0.51	-0.00	0.56		0.35	-0.00
IBM	0.37	0.02	0.50	-0.00	0.61	-0.10	0.40	-0.00
INTC	0.64		0.42		0.68		0.37	
INTC	0.44	0.31	0.28		0.46	0.31	0.25	
INTC	0.62		0.54	-0.01	0.66		0.46	-0.01
INTC	0.44	0.31	0.26	0.00	0.47	0.31	0.20	0.00
MMM	0.58		0.39		0.74		0.23	
MMM	0.51	0.10	0.35		0.78	-0.07	0.26	
MMM	0.55		0.54	-0.01	0.72		0.35	-0.01
MMM	0.50	0.08	0.50	-0.01	0.77	-0.09	0.39	-0.01
MSFT	0.43		0.46		0.48		0.43	
MSFT	0.42	0.04	0.44		0.47	0.02	0.42	
MSFT	0.37		0.71	-0.01	0.41		0.66	-0.01
MSFT	0.38	-0.02	0.70	-0.01	0.44	-0.05	0.67	-0.01

Table E.11: Estimated dynamics of the betas. OLS and MEC estimators are described in Section 5.3. Here ρ_j are the AR(p) coefficients, γ_1 is the intercept, and γ_{VIX} is the coefficient on VIX_{i-1}.

M Results of Monte Carlo Simulations

	$\xi = 0.000$						$\xi = 0.001$					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
$G_1 = 3$	0.10	0.09	0.11	0.05	0.11	0.08	0.14	0.12	0.14	0.17	0.15	0.13
$G_1 = 5$	0.11	0.10	0.12	0.06	0.13	0.09	0.13	0.12	0.13	0.11	0.14	0.11
$G_1 = 10$	0.14	0.13	0.15	0.07	0.15	0.11	0.15	0.13	0.15	0.09	0.16	0.12
$G_1 = 15$	0.17	0.15	0.17	0.09	0.18	0.13	0.17	0.16	0.18	0.10	0.19	0.13
$G_1 = 20$	0.19	0.17	0.20	0.10	0.21	0.15	0.19	0.18	0.20	0.11	0.21	0.15
$G_1 = 50$	0.29	0.26	0.30	0.15	0.32	0.22	0.29	0.27	0.31	0.16	0.33	0.23
$G_1 = 70$	0.34	0.30	0.36	0.18	0.38	0.26	0.35	0.32	0.36	0.19	0.38	0.27
G_1^{BR}	2.00	1.90	2.00	1.67	1.74	1.75	4.64	4.19	3.99	11.99	4.37	6.07
$\widehat{G}_1^{BR}$	2.01	2.01	2.01	2.01	2.01	2.01	2.17	2.16	2.11	3.22	2.16	2.37
$G_1^{\star}$	0.00	0.00	0.00	0.00	0.00	0.00	2.62	2.32	2.00	9.24	2.40	3.89
$\widehat{G}_1^{\star}$	0.37	0.36	0.36	0.36	0.36	0.36	0.73	0.72	0.65	1.93	0.71	0.99

Table M.1: The last four rows denote the averages across simulations of G_1^{BR}, $\widehat{G}_1^{BR}$, $G_1^{\star}$, and $\widehat{G}_1^{\star}$, respectively. The rows above them show the RMSE times a factor of 10 of the TS-beta for different values of the TS parameter G_1 and different scenarios (1)-(6). ξ is the noise-to-signal ratio.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3,\ V^{FS} = 0.294\ (V^{pl} = 0.155,\ V^{Asy} = 0.145)$								
100	0.286	0.284	0.283	0.283	0.283	0.284	0.283	0.284	0.284
200	0.295	0.294	0.294	0.293	0.293	0.293	0.293	0.295	0.295
300	0.299	0.297	0.297	0.297	0.296	0.296	0.296	0.298	0.299
400	0.302	0.299	0.299	0.299	0.298	0.297	0.299	0.300	0.302
500		0.300	0.300	0.299	0.299	0.298	0.300	0.302	0.304
600		0.301	0.301	0.301	0.299	0.298	0.300	0.303	0.305
800		0.303	0.302	0.302	0.300	0.299	0.301	0.305	0.308
1000			0.302	0.302	0.300	0.299	0.302	0.306	0.309
	$G_1 = 5,\ V^{FS} = 0.378\ (V^{pl} = 0.247,\ V^{Asy} = 0.242)$								
100	0.359	0.357	0.356	0.356	0.355	0.356	0.355	0.356	0.357
200	0.376	0.374	0.374	0.373	0.372	0.372	0.373	0.374	0.376
300	0.381	0.380	0.380	0.378	0.378	0.377	0.378	0.380	0.381
400	0.383	0.383	0.383	0.381	0.381	0.380	0.382	0.384	0.386
500		0.384	0.384	0.383	0.382	0.381	0.384	0.385	0.388
600		0.386	0.385	0.384	0.383	0.382	0.384	0.387	0.390
800		0.388	0.387	0.385	0.384	0.383	0.386	0.389	0.394
1000			0.387	0.386	0.384	0.383	0.386	0.391	0.396
	$G_1 = 10,\ V^{FS} = 0.585\ (V^{pl} = 0.489,\ V^{Asy} = 0.483)$								
100	0.542	0.537	0.534	0.533	0.534	0.534	0.533	0.533	0.535
200	0.586	0.583	0.581	0.580	0.580	0.580	0.581	0.582	0.583
300	0.600	0.598	0.596	0.595	0.595	0.595	0.597	0.597	0.599
400	0.596	0.605	0.604	0.601	0.602	0.602	0.605	0.605	0.608
500		0.607	0.607	0.605	0.606	0.606	0.610	0.609	0.613
600		0.610	0.609	0.608	0.609	0.609	0.612	0.613	0.616
800		0.612	0.614	0.612	0.613	0.613	0.617	0.617	0.624
1000			0.614	0.614	0.616	0.615	0.618	0.622	0.627
	$G_1 = 20,\ V^{FS} = 1.025\ (V^{pl} = 0.978,\ V^{Asy} = 0.967)$								
100	0.836	0.823	0.815	0.813	0.812	0.812	0.810	0.808	0.810
200	0.978	0.971	0.964	0.962	0.962	0.962	0.961	0.960	0.964
300	1.016	1.020	1.013	1.012	1.012	1.012	1.012	1.010	1.010
400	0.991	1.041	1.036	1.034	1.036	1.036	1.038	1.032	1.036
500		1.051	1.048	1.048	1.050	1.050	1.051	1.047	1.051
600		1.055	1.057	1.057	1.060	1.060	1.060	1.057	1.059
800		1.049	1.070	1.070	1.075	1.075	1.076	1.070	1.076
1000			1.073	1.076	1.082	1.083	1.081	1.080	1.084
	$G_1 = 50,\ V^{FS} = 2.430\ (V^{pl} = 2.456,\ V^{Asy} = 2.417)$								
100	1.079	1.007	0.975	0.958	0.950	0.942	0.940	0.929	0.932
200	1.840	1.814	1.783	1.768	1.762	1.757	1.754	1.745	1.747
300	2.019	2.082	2.056	2.042	2.039	2.033	2.031	2.023	2.023
400	1.852	2.205	2.187	2.176	2.174	2.169	2.168	2.159	2.164
500		2.266	2.261	2.253	2.253	2.251	2.251	2.245	2.241
600		2.285	2.307	2.303	2.305	2.305	2.306	2.300	2.298
800		2.200	2.360	2.366	2.372	2.377	2.380	2.374	2.367
1000			2.373	2.403	2.414	2.421	2.426	2.422	2.415

Table M.2: Sensitivity and choice of subsampling parameters J and m for the estimation of the variance of $\hat{\beta}^{TS}$ (times 100). The numbers in the table represent average variance estimated by the subsampling estimator over $N = 1000$ replications. Scenario (3). These numbers are to be compared with the actual variability of $\hat{\beta}^{TS}$, V^{FS}, which is the average over simulations of $n^{1/3}\left(\widehat{\beta}^{TS} - \beta\right)^2$. Noise-to-signal ratio is $\xi = 0.000$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3,\ V^{FS} = 0.496\ (V^{pl} = 0.223,\ V^{Asy} = 0.169)$								
100	0.467	0.466	0.464	0.464	0.463	0.462	0.461	0.462	0.463
200	0.482	0.481	0.479	0.478	0.478	0.478	0.478	0.476	0.476
300	0.490	0.487	0.485	0.484	0.482	0.482	0.482	0.480	0.481
400	0.498	0.490	0.487	0.486	0.485	0.485	0.484	0.483	0.483
500		0.491	0.488	0.486	0.486	0.486	0.485	0.483	0.483
600		0.492	0.489	0.486	0.486	0.486	0.486	0.484	0.484
800		0.496	0.492	0.488	0.488	0.488	0.488	0.485	0.485
1000			0.492	0.489	0.489	0.489	0.488	0.485	0.486
	$G_1 = 5,\ V^{FS} = 0.447\ (V^{pl} = 0.287,\ V^{Asy} = 0.251)$								
100	0.435	0.434	0.432	0.431	0.430	0.431	0.429	0.430	0.431
200	0.455	0.453	0.452	0.451	0.450	0.450	0.449	0.448	0.448
300	0.462	0.460	0.458	0.457	0.456	0.455	0.455	0.453	0.454
400	0.467	0.463	0.460	0.460	0.458	0.458	0.457	0.456	0.456
500		0.464	0.462	0.460	0.459	0.459	0.459	0.456	0.457
600		0.465	0.464	0.461	0.460	0.460	0.459	0.456	0.458
800		0.469	0.467	0.463	0.461	0.462	0.461	0.459	0.460
1000			0.468	0.463	0.462	0.462	0.460	0.458	0.460
	$G_1 = 10,\ V^{FS} = 0.625\ (V^{pl} = 0.532,\ V^{Asy} = 0.486)$								
100	0.574	0.569	0.566	0.564	0.563	0.564	0.561	0.563	0.563
200	0.621	0.617	0.615	0.613	0.610	0.610	0.610	0.609	0.611
300	0.634	0.634	0.629	0.628	0.625	0.625	0.625	0.623	0.624
400	0.632	0.640	0.636	0.635	0.631	0.630	0.631	0.628	0.631
500		0.644	0.641	0.638	0.634	0.632	0.634	0.629	0.632
600		0.645	0.645	0.640	0.636	0.634	0.635	0.630	0.633
800		0.646	0.650	0.643	0.639	0.638	0.638	0.633	0.636
1000			0.652	0.644	0.640	0.639	0.637	0.633	0.637
	$G_1 = 20,\ V^{FS} = 1.063\ (V^{pl} = 1.055,\ V^{Asy} = 0.968)$								
100	0.852	0.837	0.830	0.825	0.824	0.825	0.821	0.822	0.822
200	0.996	0.986	0.981	0.977	0.973	0.972	0.971	0.970	0.973
300	1.031	1.034	1.029	1.025	1.020	1.019	1.021	1.018	1.022
400	1.009	1.056	1.051	1.048	1.043	1.040	1.043	1.039	1.045
500		1.067	1.065	1.061	1.056	1.053	1.055	1.050	1.054
600		1.072	1.075	1.071	1.066	1.062	1.064	1.057	1.063
800		1.065	1.088	1.082	1.078	1.076	1.075	1.069	1.074
1000			1.092	1.084	1.083	1.083	1.077	1.074	1.081
	$G_1 = 50,\ V^{FS} = 2.569\ (V^{pl} = 2.657,\ V^{Asy} = 2.418)$								
100	1.087	1.010	0.978	0.963	0.954	0.947	0.942	0.935	0.939
200	1.853	1.816	1.788	1.777	1.768	1.761	1.757	1.751	1.749
300	2.025	2.076	2.054	2.048	2.040	2.033	2.026	2.017	2.020
400	1.860	2.198	2.184	2.183	2.179	2.170	2.160	2.149	2.154
500		2.256	2.260	2.264	2.261	2.254	2.242	2.228	2.232
600		2.279	2.310	2.319	2.320	2.315	2.301	2.284	2.289
800		2.204	2.370	2.390	2.397	2.396	2.374	2.356	2.361
1000			2.383	2.422	2.438	2.441	2.411	2.398	2.401

Table M.3: Sensitivity and choice of subsampling parameters J and m for the estimation of the variance of $\hat{\beta}^{TS}$ (times 100). The numbers in the table represent average variance estimated by the subsampling estimator over $N = 1000$ replications. Scenario (3). These numbers are to be compared with the actual variability of $\hat{\beta}^{TS}$, V^{FS}, which is the average over simulations of $n^{1/3}\big(\widehat{\beta}^{TS} - \beta\big)^2$. Noise-to-signal ratio is $\xi = 0.001$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$ (coverage is 0.839 for V^{pl} and 0.826 for V^{Asy})								
100	0.951	0.953	0.950	0.953	0.950	0.948	0.952	0.943	0.947
200	0.952	0.952	0.950	0.950	0.951	0.949	0.952	0.952	0.952
300	0.956	0.952	0.951	0.950	0.950	0.951	0.953	0.949	0.951
400	0.956	0.951	0.947	0.945	0.949	0.950	0.950	0.949	0.948
500		0.950	0.945	0.945	0.945	0.946	0.949	0.953	0.954
600		0.952	0.947	0.946	0.948	0.947	0.944	0.951	0.951
800		0.954	0.947	0.948	0.945	0.945	0.947	0.950	0.950
1000			0.955	0.949	0.946	0.944	0.946	0.947	0.950
	$G_1 = 5$ (coverage is 0.890 for V^{pl} and 0.885 for V^{Asy})								
100	0.947	0.946	0.948	0.944	0.948	0.942	0.945	0.950	0.947
200	0.948	0.945	0.948	0.945	0.946	0.944	0.947	0.946	0.948
300	0.954	0.945	0.944	0.947	0.947	0.944	0.946	0.947	0.949
400	0.951	0.946	0.943	0.948	0.948	0.945	0.948	0.948	0.952
500		0.948	0.944	0.946	0.948	0.946	0.946	0.954	0.956
600		0.950	0.948	0.947	0.946	0.945	0.951	0.950	0.955
800		0.955	0.949	0.945	0.942	0.943	0.950	0.953	0.956
1000			0.950	0.947	0.943	0.948	0.949	0.951	0.950
	$G_1 = 10$ (coverage is 0.933 for V^{pl} and 0.936 for V^{Asy})								
100	0.950	0.948	0.945	0.947	0.950	0.947	0.944	0.944	0.946
200	0.956	0.954	0.953	0.951	0.954	0.953	0.953	0.956	0.951
300	0.958	0.954	0.953	0.953	0.953	0.954	0.951	0.951	0.944
400	0.958	0.955	0.952	0.952	0.955	0.953	0.954	0.952	0.949
500		0.957	0.954	0.953	0.954	0.952	0.951	0.953	0.946
600		0.959	0.952	0.952	0.952	0.952	0.950	0.953	0.946
800		0.959	0.956	0.950	0.950	0.945	0.945	0.953	0.952
1000			0.953	0.952	0.948	0.947	0.945	0.950	0.948
	$G_1 = 20$ (coverage is 0.942 for V^{pl} and 0.942 for V^{Asy})								
100	0.919	0.913	0.911	0.914	0.914	0.918	0.915	0.913	0.916
200	0.939	0.939	0.934	0.942	0.942	0.944	0.934	0.937	0.940
300	0.947	0.945	0.945	0.948	0.949	0.947	0.944	0.942	0.947
400	0.947	0.949	0.945	0.945	0.947	0.947	0.945	0.942	0.945
500		0.947	0.943	0.946	0.949	0.945	0.947	0.940	0.949
600		0.948	0.943	0.947	0.942	0.944	0.943	0.937	0.948
800		0.948	0.947	0.944	0.942	0.947	0.942	0.942	0.944
1000			0.953	0.947	0.945	0.940	0.939	0.935	0.937
	$G_1 = 50$ (coverage is 0.945 for V^{pl} and 0.949 for V^{Asy})								
100	0.806	0.791	0.789	0.778	0.780	0.777	0.782	0.765	0.773
200	0.896	0.892	0.888	0.890	0.891	0.887	0.884	0.881	0.881
300	0.909	0.914	0.915	0.913	0.912	0.914	0.909	0.907	0.909
400	0.897	0.925	0.919	0.923	0.926	0.924	0.916	0.919	0.913
500		0.929	0.924	0.930	0.930	0.926	0.923	0.923	0.923
600		0.928	0.927	0.931	0.929	0.927	0.921	0.926	0.924
800		0.931	0.936	0.933	0.930	0.925	0.921	0.922	0.925
1000			0.942	0.937	0.931	0.925	0.922	0.921	0.922

Table M.4: Sensitivity and choice of subsampling parameters J and m for the coverage of the confidence interval for $\hat{\beta}^{TS}$. The numbers in the table represent empirical coverage rate of the subsampling estimator over $N = 1000$ replications. Scenario (3). The nominal coverage rate is 0.95. Noise-to-signal ratio is $\xi = 0.000$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$ (coverage is 0.801 for V^{pl} and 0.736 for V^{Asy})								
100	0.939	0.941	0.942	0.943	0.937	0.939	0.938	0.940	0.938
200	0.941	0.946	0.946	0.945	0.945	0.944	0.941	0.941	0.941
300	0.943	0.948	0.950	0.947	0.945	0.942	0.944	0.942	0.940
400	0.953	0.945	0.949	0.947	0.945	0.943	0.945	0.943	0.947
500		0.948	0.946	0.948	0.944	0.941	0.942	0.943	0.944
600		0.952	0.949	0.945	0.944	0.939	0.941	0.941	0.944
800		0.955	0.951	0.942	0.941	0.942	0.939	0.938	0.942
1000			0.951	0.946	0.946	0.941	0.937	0.938	0.936
	$G_1 = 5$ (coverage is 0.879 for V^{pl} and 0.863 for V^{Asy})								
100	0.941	0.939	0.938	0.937	0.938	0.939	0.937	0.936	0.936
200	0.940	0.943	0.943	0.940	0.943	0.944	0.943	0.941	0.941
300	0.947	0.942	0.943	0.945	0.949	0.945	0.946	0.943	0.942
400	0.951	0.943	0.944	0.946	0.948	0.947	0.945	0.948	0.947
500		0.943	0.947	0.946	0.945	0.948	0.947	0.948	0.948
600		0.944	0.947	0.947	0.946	0.949	0.945	0.947	0.946
800		0.949	0.949	0.946	0.942	0.942	0.946	0.945	0.945
1000			0.952	0.947	0.943	0.940	0.942	0.940	0.942
	$G_1 = 10$ (coverage is 0.926 for V^{pl} and 0.914 for V^{Asy})								
100	0.940	0.938	0.935	0.934	0.936	0.935	0.934	0.934	0.933
200	0.947	0.944	0.942	0.942	0.944	0.945	0.944	0.943	0.943
300	0.947	0.946	0.944	0.942	0.944	0.946	0.946	0.944	0.946
400	0.951	0.946	0.944	0.942	0.944	0.945	0.942	0.944	0.943
500		0.947	0.944	0.942	0.942	0.944	0.938	0.941	0.942
600		0.946	0.944	0.941	0.939	0.941	0.938	0.940	0.939
800		0.950	0.942	0.941	0.939	0.936	0.937	0.937	0.938
1000			0.944	0.944	0.939	0.935	0.932	0.937	0.937
	$G_1 = 20$ (coverage is 0.945 for V^{pl} and 0.939 for V^{Asy})								
100	0.924	0.914	0.911	0.910	0.911	0.913	0.913	0.912	0.910
200	0.940	0.933	0.933	0.929	0.931	0.932	0.931	0.928	0.930
300	0.940	0.937	0.939	0.935	0.934	0.936	0.932	0.932	0.929
400	0.940	0.942	0.941	0.936	0.937	0.937	0.933	0.933	0.937
500		0.941	0.944	0.944	0.939	0.939	0.937	0.935	0.936
600		0.947	0.947	0.944	0.938	0.939	0.938	0.936	0.936
800		0.946	0.942	0.942	0.937	0.934	0.935	0.935	0.937
1000			0.947	0.940	0.936	0.934	0.930	0.931	0.930
	$G_1 = 50$ (coverage is 0.955 for V^{pl} and 0.940 for V^{Asy})								
100	0.801	0.788	0.778	0.776	0.773	0.768	0.770	0.759	0.763
200	0.897	0.890	0.890	0.889	0.890	0.889	0.885	0.884	0.885
300	0.909	0.915	0.905	0.905	0.904	0.902	0.903	0.899	0.901
400	0.898	0.920	0.916	0.913	0.915	0.910	0.911	0.909	0.908
500		0.920	0.920	0.920	0.921	0.920	0.912	0.913	0.908
600		0.923	0.922	0.921	0.922	0.922	0.916	0.911	0.913
800		0.918	0.930	0.922	0.924	0.924	0.920	0.918	0.914
1000			0.931	0.926	0.925	0.921	0.922	0.919	0.918

Table M.5: Sensitivity and choice of subsampling parameters J and m for the coverage of the confidence interval for $\hat{\beta}^{TS}$. The numbers in the table represent empirical coverage rate of the subsampling estimator over $N = 1000$ replications. Scenario (3). The nominal coverage rate is 0.95. Noise-to-signal ratio is $\xi = 0.001$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.256	0.135	0.126	0.259	0.088	0.047	0.044	0.081
$G_1 = 5$	0.330	0.215	0.209	0.334	0.114	0.075	0.073	0.107
$G_1 = 10$	0.524	0.425	0.419	0.535	0.181	0.148	0.147	0.170
$G_1 = 20$	0.909	0.851	0.837	0.935	0.316	0.296	0.293	0.306
$G_1 = 50$	1.938	2.137	2.093	2.159	0.680	0.743	0.734	0.718
	(3)				(4)			
$G_1 = 3$	0.245	0.130	0.122	0.232	0.305	0.155	0.145	0.301
$G_1 = 5$	0.315	0.207	0.204	0.301	0.391	0.248	0.241	0.379
$G_1 = 10$	0.502	0.410	0.407	0.474	0.616	0.490	0.483	0.574
$G_1 = 20$	0.874	0.819	0.814	0.812	1.050	0.981	0.965	1.014
$G_1 = 50$	1.895	2.057	2.035	1.903	2.255	2.477	2.414	2.414
	(5)				(6)			
$G_1 = 3$	0.298	0.155	0.145	0.294	0.183	0.098	0.092	0.172
$G_1 = 5$	0.381	0.247	0.242	0.378	0.236	0.155	0.153	0.224
$G_1 = 10$	0.606	0.489	0.483	0.585	0.376	0.307	0.305	0.355
$G_1 = 20$	1.050	0.978	0.967	1.025	0.654	0.614	0.610	0.635
$G_1 = 50$	2.251	2.456	2.417	2.430	1.410	1.539	1.526	1.496

Table M.6: The average over simulations of the following measures of dispersion of $\hat{\beta}^{TS}$ (times 100): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asysmptotic variance V^{Asy}, and an average over simulations of $n^{1/3}\left(\widehat{\beta}^{TS}-\beta\right)^2$ denoted by V^{FS}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.000$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.477	0.216	0.159	0.470	0.950	0.387	0.298	0.879
$G_1 = 5$	0.421	0.263	0.222	0.406	0.399	0.219	0.165	0.386
$G_1 = 10$	0.561	0.479	0.423	0.555	0.273	0.234	0.170	0.258
$G_1 = 20$	0.930	0.947	0.841	0.939	0.351	0.416	0.299	0.350
$G_1 = 50$	1.986	2.382	2.100	2.172	0.693	1.027	0.735	0.755
	(3)				(4)			
$G_1 = 3$	0.465	0.207	0.156	0.448	0.508	0.236	0.174	0.515
$G_1 = 5$	0.402	0.246	0.216	0.390	0.472	0.301	0.252	0.461
$G_1 = 10$	0.534	0.445	0.410	0.527	0.639	0.556	0.486	0.648
$G_1 = 20$	0.881	0.879	0.815	0.893	1.055	1.103	0.967	1.095
$G_1 = 50$	1.884	2.205	2.036	2.125	2.233	2.793	2.415	2.548
	(5)				(6)			
$G_1 = 3$	0.486	0.223	0.169	0.496	0.531	0.225	0.164	0.508
$G_1 = 5$	0.459	0.287	0.251	0.447	0.367	0.216	0.179	0.365
$G_1 = 10$	0.632	0.532	0.486	0.625	0.423	0.355	0.312	0.415
$G_1 = 20$	1.053	1.055	0.968	1.063	0.671	0.691	0.612	0.670
$G_1 = 50$	2.254	2.657	2.418	2.569	1.415	1.727	1.527	1.563

Table M.7: The average over simulations of the following measures of dispersion of $\hat{\beta}^{TS}$ (times 100): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asysmptotic variance V^{Asy}, and an average over simulations of $n^{1/3}\left(\widehat{\beta}^{TS} - \beta\right)^2$ denoted by V^{FS}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.001$.

	t^{sub}	t^{pl}	t^{Asy}	t^{sub}	t^{pl}	t^{Asy}
		(1)			(2)	
$G_1 = 3$	0.941	0.838	0.820	0.955	0.865	0.850
$G_1 = 5$	0.938	0.876	0.885	0.947	0.898	0.894
$G_1 = 10$	0.943	0.923	0.927	0.955	0.939	0.938
$G_1 = 20$	0.947	0.944	0.942	0.952	0.951	0.949
$G_1 = 50$	0.925	0.946	0.940	0.931	0.952	0.949
		(3)			(4)	
$G_1 = 3$	0.952	0.860	0.843	0.945	0.844	0.822
$G_1 = 5$	0.951	0.891	0.892	0.950	0.891	0.881
$G_1 = 10$	0.954	0.932	0.933	0.946	0.932	0.927
$G_1 = 20$	0.951	0.944	0.951	0.935	0.941	0.938
$G_1 = 50$	0.937	0.944	0.949	0.921	0.945	0.941
		(5)			(6)	
$G_1 = 3$	0.946	0.839	0.826	0.956	0.855	0.852
$G_1 = 5$	0.946	0.890	0.885	0.946	0.899	0.902
$G_1 = 10$	0.952	0.933	0.936	0.956	0.936	0.945
$G_1 = 20$	0.945	0.942	0.942	0.948	0.946	0.947
$G_1 = 50$	0.926	0.945	0.949	0.933	0.945	0.947

Table M.8: Coverage of the confidence interval of $\hat{\beta}^{TS}$ based on the following measures of dispersion of $\hat{\beta}^{TS}$: the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, and the (unobserved) asysmptotic variance V^{Asy}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.000$.

	t^{sub}	t^{pl}	t^{Asy}	t^{sub}	t^{pl}	t^{Asy}
		(1)			(2)	
$G_1 = 3$	0.945	0.809	0.724	0.956	0.800	0.725
$G_1 = 5$	0.943	0.884	0.853	0.950	0.866	0.802
$G_1 = 10$	0.949	0.937	0.922	0.949	0.930	0.884
$G_1 = 20$	0.948	0.953	0.946	0.941	0.963	0.925
$G_1 = 50$	0.928	0.950	0.944	0.931	0.977	0.946
		(3)			(4)	
$G_1 = 3$	0.947	0.816	0.760	0.934	0.810	0.738
$G_1 = 5$	0.953	0.881	0.863	0.945	0.888	0.846
$G_1 = 10$	0.942	0.931	0.924	0.939	0.925	0.907
$G_1 = 20$	0.938	0.939	0.930	0.935	0.946	0.925
$G_1 = 50$	0.928	0.949	0.938	0.914	0.958	0.944
		(5)			(6)	
$G_1 = 3$	0.941	0.801	0.736	0.940	0.816	0.758
$G_1 = 5$	0.948	0.879	0.863	0.949	0.868	0.834
$G_1 = 10$	0.944	0.926	0.914	0.941	0.926	0.912
$G_1 = 20$	0.939	0.945	0.939	0.938	0.950	0.931
$G_1 = 50$	0.920	0.955	0.940	0.924	0.960	0.945

Table M.9: Coverage of the confidence interval of $\hat{\beta}^{TS}$ based on the following measures of dispersion of $\hat{\beta}^{TS}$: the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, and the (unobserved) asysmptotic variance V^{Asy}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.001$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.073	0.035	0.033	0.061	0.222	0.108	0.101	0.212
$G_1 = 5$	0.094	0.057	0.055	0.080	0.284	0.173	0.169	0.272
$G_1 = 10$	0.147	0.112	0.110	0.126	0.447	0.342	0.337	0.429
$G_1 = 20$	0.253	0.224	0.220	0.232	0.773	0.682	0.675	0.783
$G_1 = 50$	0.557	0.560	0.551	0.555	1.701	1.705	1.687	1.797
	(3)				(4)			
$G_1 = 3$	0.032	0.016	0.015	0.032	0.115	0.047	0.044	0.095
$G_1 = 5$	0.041	0.026	0.025	0.042	0.143	0.076	0.073	0.120
$G_1 = 10$	0.066	0.051	0.051	0.067	0.212	0.150	0.146	0.196
$G_1 = 20$	0.114	0.103	0.101	0.123	0.349	0.300	0.292	0.356
$G_1 = 50$	0.253	0.257	0.254	0.280	0.744	0.749	0.731	0.736
	(5)				(6)			
$G_1 = 3$	0.042	0.021	0.019	0.040	0.068	0.032	0.030	0.064
$G_1 = 5$	0.054	0.033	0.032	0.050	0.086	0.051	0.050	0.083
$G_1 = 10$	0.085	0.065	0.064	0.081	0.135	0.101	0.101	0.130
$G_1 = 20$	0.147	0.131	0.128	0.152	0.232	0.201	0.202	0.236
$G_1 = 50$	0.323	0.327	0.320	0.342	0.509	0.502	0.505	0.538

Table M.10: The average over simulations of the following measures of dispersion of $\widehat{\langle X, X\rangle}^{TS}$ (times 10^7): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asysmptotic variance V^{Asy}, and an average over simulations of $n^{1/3}\big(\widehat{\langle X, X\rangle}^{TS} - \langle X, X\rangle\big)^2$ denoted by V^{FS}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.000$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Asy}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.102	0.046	0.036	0.099	0.723	0.296	0.225	0.655
$G_1 = 5$	0.106	0.063	0.057	0.105	0.470	0.260	0.215	0.483
$G_1 = 10$	0.151	0.121	0.112	0.152	0.518	0.407	0.353	0.542
$G_1 = 20$	0.255	0.240	0.223	0.265	0.807	0.787	0.687	0.854
$G_1 = 50$	0.559	0.599	0.556	0.596	1.734	1.957	1.710	1.869
	(3)				(4)			
$G_1 = 3$	0.046	0.021	0.017	0.045	0.149	0.061	0.048	0.138
$G_1 = 5$	0.048	0.028	0.026	0.051	0.157	0.086	0.076	0.153
$G_1 = 10$	0.069	0.054	0.051	0.073	0.219	0.164	0.149	0.209
$G_1 = 20$	0.116	0.106	0.102	0.126	0.354	0.327	0.297	0.355
$G_1 = 50$	0.255	0.266	0.254	0.283	0.764	0.817	0.742	0.791
	(5)				(6)			
$G_1 = 3$	0.055	0.025	0.021	0.054	0.125	0.052	0.041	0.108
$G_1 = 5$	0.060	0.036	0.033	0.061	0.111	0.062	0.055	0.111
$G_1 = 10$	0.088	0.068	0.064	0.091	0.146	0.112	0.103	0.153
$G_1 = 20$	0.148	0.136	0.128	0.158	0.240	0.221	0.205	0.250
$G_1 = 50$	0.324	0.339	0.321	0.351	0.518	0.551	0.513	0.557

Table M.11: The average over simulations of the following measures of dispersion of $\widehat{\langle X, X\rangle}^{TS}$ (times 10^7): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asysmptotic variance V^{Asy}, and an average over simulations of $n^{1/3}\big(\widehat{\langle X, X\rangle}^{TS} - \langle X, X\rangle\big)^2$ denoted by V^{FS}. Scenarios (1)-(6) are described in Section 4. $J = 500$, $m = 3000$, and $\xi = 0.001$.

	t^{sub}	t^{pl}	t^{Asy}	t^{sub}	t^{pl}	t^{Asy}
		(1)			(2)	
$G_1 = 3$	0.961	0.870	0.856	0.948	0.837	0.825
$G_1 = 5$	0.956	0.898	0.887	0.955	0.887	0.882
$G_1 = 10$	0.951	0.923	0.921	0.952	0.918	0.917
$G_1 = 20$	0.947	0.944	0.942	0.936	0.922	0.922
$G_1 = 50$	0.945	0.949	0.951	0.931	0.936	0.939
		(3)			(4)	
$G_1 = 3$	0.952	0.833	0.821	0.965	0.837	0.820
$G_1 = 5$	0.941	0.874	0.872	0.958	0.877	0.878
$G_1 = 10$	0.932	0.899	0.906	0.942	0.903	0.910
$G_1 = 20$	0.931	0.918	0.933	0.934	0.918	0.915
$G_1 = 50$	0.925	0.933	0.936	0.918	0.935	0.951
		(5)			(6)	
$G_1 = 3$	0.950	0.833	0.830	0.948	0.844	0.827
$G_1 = 5$	0.943	0.877	0.877	0.954	0.876	0.883
$G_1 = 10$	0.939	0.905	0.912	0.949	0.909	0.918
$G_1 = 20$	0.936	0.917	0.927	0.936	0.920	0.929
$G_1 = 50$	0.928	0.932	0.941	0.927	0.932	0.936

Table M.12: Coverage of the confidence interval of $\widehat{\langle X, X\rangle}^{TS}$ *based on the following measures of dispersion of* $\widehat{\langle X, X\rangle}^{TS}$*: the subsampling estimator* $\widehat{V}^{sub}$*, the plug-in estimated value of the asymptotic variance* $\widehat{V}^{pl}$*, and the (unobserved) asysmptotic variance* V^{Asy}*. Scenarios (1)-(6) are described in Section 4.* $J = 500$*,* $m = 3000$*, and* $\xi = 0.000$*.*

	t^{sub}	t^{pl}	t^{Asy}	t^{sub}	t^{pl}	t^{Asy}
		(1)			(2)	
$G_1 = 3$	0.943	0.810	0.756	0.949	0.799	0.755
$G_1 = 5$	0.943	0.865	0.844	0.942	0.843	0.812
$G_1 = 10$	0.936	0.903	0.905	0.950	0.909	0.897
$G_1 = 20$	0.933	0.927	0.929	0.938	0.933	0.917
$G_1 = 50$	0.919	0.936	0.935	0.928	0.944	0.933
		(3)			(4)	
$G_1 = 3$	0.942	0.820	0.774	0.949	0.807	0.754
$G_1 = 5$	0.940	0.845	0.848	0.946	0.845	0.838
$G_1 = 10$	0.928	0.886	0.888	0.946	0.912	0.899
$G_1 = 20$	0.924	0.923	0.917	0.925	0.925	0.928
$G_1 = 50$	0.925	0.936	0.936	0.918	0.936	0.938
		(5)			(6)	
$G_1 = 3$	0.944	0.815	0.761	0.955	0.827	0.779
$G_1 = 5$	0.940	0.854	0.844	0.944	0.848	0.830
$G_1 = 10$	0.940	0.909	0.899	0.953	0.899	0.902
$G_1 = 20$	0.937	0.926	0.924	0.935	0.919	0.917
$G_1 = 50$	0.920	0.936	0.936	0.924	0.940	0.933

Table M.13: Coverage of the confidence interval of $\widehat{\langle X, X\rangle}^{TS}$ *based on the following measures of dispersion of* $\widehat{\langle X, X\rangle}^{TS}$*: the subsampling estimator* $\widehat{V}^{sub}$*, the plug-in estimated value of the asymptotic variance* $\widehat{V}^{pl}$*, and the (unobserved) asysmptotic variance* V^{Asy}*. Scenarios (1)-(6) are described in Section 4.* $J = 500$*,* $m = 3000$*, and* $\xi = 0.001$*.*