JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2023, VOL. 41, NO. 2, 538–549
https://doi.org/10.1080/07350015.2022.2040520

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

Ilze Kalnina

Department of Economics, Poole College of Management, North Carolina State University, Raleigh, NC

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. By construction, the subsampling method delivers estimates of the variance-covariance matrices that are always positive semidefinite. Our simulation study indicates that the subsampling method is more robust than the plug-in method based on the asymptotic expression for the variance.

We use our subsampling method to study the dynamics of financial betas of six stocks on the NYSE. We document significant variation in betas, and find that tick data captures more variation in betas than the data sampled at moderate frequencies such as every 5 or 20 min. 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.

ARTICLE HISTORY
Received April 2021
Accepted January 2022

KEYWORDS
Market microstructure noise;
Realized volatility;
Subsampling

1. Introduction

Financial econometrics has proposed many multivariate estimators for high frequency data. Moreover, applied researchers often combine multiple estimators.

Inference in these settings requires estimation of the asymptotic variance-covariance (VCV) matrices of the estimators. The expressions for the VCVs can be complicated, and can be difficult or tedious to derive and/or estimate.

We propose a subsampling method to estimate the VCV matrices of high frequency estimators. The method avoids the need to derive the formulas of these VCVs and to construct their analog estimators, instead automatically estimating the VCVs by the proposed resampling scheme. We prove the validity of this multivariate 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 the stochastic volatility such as a Brownian semimartingale or long memory structure, for leverage effects, and for autocorrelated market microstructure noise. Importantly, the proposed multivariate method has the appealing property that the estimated VCVs are guaranteed to be positive semidefinite.

The use of resampling methods to compute standard errors of estimators is a standard practice in applied research that uses iid and classic weakly dependent data. The approach of this article permits extending this practice to a broad class of (possibly multivariate) estimators in high-frequency settings. For instance, multivariate estimators in the presence of serially correlated market microstructure noise may have asymptotic variances with quite complicated expressions that include numerous cross-parameter and cross-lag terms. Deriving them and implementing the corresponding plug-in estimators may be a time consuming and challenging task. Our procedure automatically adapts to the serial correlation of the noise.

Moreover, our automatic procedure delivers a certain degree of robustness compared to the plug-in methods for the estimation of VCVs. Our simulation study presents an example of this property with the Two Scales (TS) estimator of Zhang, Mykland, and Aït-Sahalia (2005) and Zhang (2011) that uses tuning parameters suggested by Bandi and Russell (2011). The choice of tuning parameters following Bandi and Russell (2011) minimizes the finite sample Mean Squared Error (MSE) of the TS estimator. This value of the tuning parameter is often much lower than what the "rule-of-thumb" would suggest. Our finite sample experiments show that in such cases, the subsampling method delivers estimates of the VCVs that are close to the actual finite sample VCVs, while the plug-in method may substantially underestimate them. The result is not driven by the failure of estimating the components of the plug-in estimator; the plug-in estimator is quite close to the derived asymptotic variance, but the latter can differ significantly from the finite sample variance.

CONTACT Ilze Kalnina ikalnin@ncsu.edu Department of Economics, Poole College of Management, North Carolina State University, Nelson Hall 4164, Box 8110, Raleigh, NC 27695-8110.

Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

We use our multivariate subsampling method to study dynamics in financial betas in two applications.

First, we test the hypothesis of betas being time-invariant. Such tests can help choosing an appropriate window length for a rolling window approach to estimating betas, see, for example, Fama and MacBeth (1973) and Fama and French (1992). 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 min), and the TS with tick data. We document significant variation in betas of every stock across weekly windows, and find that tick data captures more variation in betas than the data sampled at moderate frequencies.

Second, we provide and implement Measurement-Error-Corrected (MEC) regression estimators of the dynamics of betas. A popular approach of modeling the dynamics of betas is to use parametric auto-regressive models, see, for example, Braun, Nelson, and Sunier (1995), Bekaert and Wu (2000), Jostova and Philipov (2005), and Adrian and Franzoni (2009). More recently, Andersen et al. (2005b, 2006) and Patton and Verardo (2012) estimate the lower-frequency, such as quarterly, beta auto-regressions, in which the beta for the individual quarter is estimated using high-frequency data. The problem of such auto-regressions is that the estimates of betas are noisy measures of the true betas, and the regressions that ignore this suffer from the errors-in-variables bias problem. We provide MEC estimators that correct for this bias. Their implementation requires estimation of the VCVs of the estimators of the individual betas. We implement the MEC estimators using our subsampling method to estimate these VCVs. Our empirical results suggest that betas are more persistent than the naive regression estimators would suggest.

Several articles are related to the multivariate subsampling method developed in this article. Kalnina (2011) shows the validity of the univariate subsampling method for a general class of estimators. Compared to that article, our article proves the validity of the proposed multivariate subsampling method under weaker assumptions, numerically illustrates the robustness of the method in the case of the Two Scales estimator, provides MEC estimators of the dynamic regressions, and considers the choice of the subsample size. Our subsampling method is partly related to the classical subsampling in the statistics literature for stationary data, see, for example, Politis and Romano (1994) and Lahiri et al. (1999). The high-frequency literature has also proposed several estimator-specific subsampling methods. Kalnina and Linton (2007) develop two schemes for the realized variance estimator. Christensen et al. (2017) show that one of these schemes is also valid for the power variation estimator, and propose modifications suitable for several alternative volatility estimators. Ikeda (2016) and Varneskov (2016) apply the subsampling scheme of Kalnina (2011) and of this article to the Two Scale and flat-top realized kernels. All of the above methods rely on the squared differences of the estimator calculated on nested subsamples of various forms. Recently, Mykland and Zhang (2017) propose an interesting univariate estimator that is based on comparing adjacent subsamples, followed by an additive bias-correction, instead of nested subsamples. One advantage of the subsampling method in our article is that the estimated VCVs are guaranteed to be positive semidefinite.

The remainder of this article is organized as follows. Section 2 introduces the multivariate subsampling method, and presents the main theoretical results. Section 3 illustrates how the subsampling method applies to the multivariate Two Scales estimator and the Two Scales Realized Beta. Section 4 describes methods of analysis of dynamics in betas using the subsampled variances, including the MEC estimators. Section 5 studies the finite sample properties of the proposed methods and investigates the choice of the tuning parameters. Section 6 contains an empirical illustration. Section 7 concludes. All proofs are collected in the Appendices A–C in the supplementary materials. Appendix D in the supplementary materials contains the tables and figures for the summary statistics of data. Appendix E in the supplementary materials contains the tables and figures for the empirical illustration in Section 6. Appendix M in the supplementary materials contains the tables for the simulations in Section 5. Appendix R in the supplementary materials contains an additional simulation study.

2. A Multivariate Subsampling Method

Suppose we are interested in estimating some volatility measure θ, which is a functional of some spot variance-covariance process c. The data generating process may be an Itô Semimartingale with or without some noise contamination or other distortions, and c is the spot variance-covariance process of the Itô Semimartingale component. (For a definition of an Itô Semimartingale and its spot variance-covariance process, see Definition 1.16 and Equations (1.74)–(1.75) of Aït-Sahalia and Jacod 2014). Suppose we have an estimator $\widehat{\theta}_n$, for which we know that

$$\tau_n\left(\widehat{\theta}_n-\theta\right) \Rightarrow N(0, V), \tag{1}$$

where τ_n is the rate of convergence when n observations are used, $\Rightarrow$ denotes stable convergence in law, and V is the VCV of $\widehat{\theta}_n$. Suppose we would like to estimate V.

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

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

where

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

and where $0=t_0<t_1<\cdots<t_n=1$ denote the observation times. In the above, $K=\lfloor n/m\rfloor$ is the number

of subsamples, and the term $(1 - J/m)^{-1}$ is a finite sample adjustment factor. This adjustment is negligible asymptotically, but improves finite sample behavior; it has the usual motivation in variance estimation. Without the adjustment factor, $\widehat{V}^{\text{sub}}$ in Equation (2) is the multivariate version of the subsampling estimator of Kalnina (2011).

Note that $\widehat{V}^{\text{sub}}$, by construction, is positive semidefinite, which is a key property of any VCV matrix estimator.

A few comments on the intuition of the estimator $\widehat{V}^{\text{sub}}$ are in order. Notice that $\frac{n}{m}\widehat{\theta}_k^{\text{long}}$ and $\frac{n}{J}\widehat{\theta}_k^{\text{short}}$ estimate the same object, the "spot" version of θ, but one uses more observations than the other. Therefore, $\frac{n}{m}\widehat{\theta}_k^{\text{long}}$ can be used to demean $\frac{n}{J}\widehat{\theta}_k^{\text{short}}$. The outer product of the differences is multiplied by the rate of convergence $\frac{n}{J}\tau_n^2$ of $\frac{n}{J}\widehat{\theta}_k^{\text{short}}$, 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, $\widehat{V}^{\text{sub}}$ estimates V under the regularity assumptions (described below).

The estimator $\widehat{V}^{\text{sub}}$ in Equation (2) uses nonoverlapping 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 lth long subsample is t_{ls}. The expression in Equation (2) is obtained by setting $s = m$.

We use the following assumptions to prove the consistency of $\widehat{V}^{\text{sub}}$ in Equation (2). 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 integrated 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 $E\left[\left\|c_{t_2} - c_{t_1}\right\|^2\right] \leq B_c |t_2 - t_1|^\alpha$ 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 $\tau_n^2\, Jm^\alpha \big/ n^{1+\alpha} \to 0$.

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 $s = m$,

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

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, e.g., sec. 4.2 in Mykland and Zhang 2006; Zhang 2012), and principal components (see Aït-Sahalia and Xiu 2019). The asymptotic variances V of the corresponding estimators of such θ typically also satisfy Assumption A1. Counter-examples for Assumption A1 include parameters of interest that include price jumps, for example, the quadratic variation. Assumption A1 does not rule out price jumps, but it does rule out price jumps appearing in the parameter of interest θ or the asymptotic variance of its estimator V. A standard method to deal with price jumps is by truncation, in which case the asymptotic variance expressions do not contain price jumps and Assumption A1 is satisfied, see, for example, chap. 6.2.1. of Aït-Sahalia and Jacod (2014).

Assumption A2 allows long memory in the spot variance process, as illustrated by Lemma 2. Assumption A2 also allows for the more standard Brownian Semimartingale dynamics in the c process. In the latter case, Assumption A2 is satisfied with $\alpha = 1$. Note that Kalnina (2011) assumes a Brownian Semimartingale dynamics for volatility.

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 rate requirement in Assumption A3 arises due to the "discretization bias" in the volatility, that is, 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 consider the case of 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, a similar property is assumed to hold for the case of the shorter subsamples ($s = J$). Note that the smoothness condition on c_t was already imposed in Assumption A2. In contrast to the relevant assumption (A5) in Kalnina (2011), Assumption A4 allows for the leverage effects in returns, which is empirically important. Assumption A4 essentially requires that the estimator on the subsample behaves similarly to the estimator on the full sample, in the sense that its variance is the corresponding portion of the integral V. This is an important property, and subsampling relies on it, but typically, high-frequency estimators have such property.

With the above assumptions, the subsampling estimator is consistent:

Theorem 1. Suppose Assumptions A1–A4 hold. Let $\widehat{V}^{\text{sub}}$ be defined by (2). Then, as $n \to \infty$,

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

Remark. Notice that $\widehat{V}^{\text{sub}}$ is an automatic method in the sense that it does not use any information about the expression for V. As a result, the applied researchers can easily implement it. The availability of an automatic estimator of V is particularly useful in practice for multivariate parameters θ and possibly serially correlated data (e.g., because of the market microstructure noise), since in those cases the expression for V is generally complicated and, moreover, there is a large number of plug-in parameters such as the cross-covariances to be estimated if one wants to estimate the plug-in estimator of V.

Lemma 2 illustrates the fact that Assumption A2 can be satisfied for a long memory spot volatility process.

Lemma 2. 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 $\alpha = \overline{\alpha}$.

2.1. The Choice of Parameters for Subsampling

For practical application, one needs to choose specific values for the subsample sizes (m and J). The choice of such tuning parameters is well known to be a difficult problem, with many alternative approaches proposed even for the standard HAC estimation (see, e.g., Lazarus et al. 2018 for a recent review). For example, Andrews (1991) derives an MSE-optimal HAC estimator, which depends on the unknown spectral density. To obtain a data-driven choice of the tuning parameter, Andrews (1991) suggests estimating the latent spectral density using a simple parametric reference model, AR(1).

Our goal is to provide a method of inference that does not require the researcher to calculate the asymptotic variance V, let alone any higher-order expansions. Therefore, our suggested method is not based on an analytical MSE expansion.

We suggest the following procedure to choose the subsample sizes. It is a variation of the calibration method considered in Politis, Romano, and Wolf (1999). First, the researcher chooses and estimates a parametric reference model of the stock price dynamics for the real data, on which the subsampling method is to be used. Note that one can use any additional available and relevant data, for example, option prices, in addition to the price data. In this article, we illustrate this method using Heston (1993) model as the reference model. The reference model should be chosen to fit the data sufficiently well; if the researcher suspects the Heston model is a poor fit, she may use, for example, a two-factor volatility model instead. Then, the researcher simulates pseudo-data from the estimated model, and calculates the subsampled variances for a range of values of m and J. This simulation is repeated multiple times. Finally, the researcher chooses the values of m and J that provide the best performance of the subsampling estimator, based on the estimation of the asymptotic variance, or the coverage of the confidence intervals.

This suggestion is easy to implement, and it has an intuitive justification. By choosing the reference model, the researchers can tailor the procedure to the applications they have in mind. In Section 5, we study this procedure in a set of finite sample experiments, see also Appendix R in the supplementary materials. We leave theoretical analysis of the procedure for future research.

Importantly, we find that the subsampling procedure is insensitive to the choice of subsample sizes for a wide range of values.

3. Inference for the Two Scales Realized Beta ($\widehat{\beta}^{TS}$) Using Subsampling

This section illustrates the subsampling methodology by applying it to a specific estimator in the setting with the market microstructure noise. In particular, suppose we are interested in constructing confidence intervals for an estimator of beta based on the Two Scales (co)volatility estimators. We choose the Two Scales estimator because of its popularity and the fact that the choice of the (single) Two Scales tuning parameter, as well as the accuracy of asymptotic approximations, has been studied extensively for this estimator (see, e.g., Bandi and Russell 2011). We note that many alternative multivariate volatility estimators that are robust to the market microstructure noise have been proposed. Examples of rate-optimal covolatility estimators (with a rate of convergence $n^{1/4}$) include the preaveraging estimator (Christensen, Kinnebrock, and Podolskij 2010), theQuasi-Maximum Likelihood estimator (Aït-Sahalia, Fan, and Xiu 2010), the Multi-Scale Realized Variance (Bibinger 2012), and the local spectral estimator (Bibinger and Reiss 2014). Themultivariate realized kernel of Barndorff-Nielsen et al. (2011) uses a larger bandwidth to guarantee positive semidefinite estimates and robustness to general noise; in this case the realized kernel is $n^{1/5}$-convergent. The Two Scales estimator has a rate of convergence $n^{1/6}$.

For simplicity, we introduce the framework for one factor beta; we also consider one time period, such as one week, which we normalize to be $[0, 1]$. Denote by X the vector containing the latent (log-)prices of the asset X^S and of the factor X^F, so that $X_t = \left(X_t^S, X_t^F\right)'$. We follow Zhang, Mykland, and Aït-Sahalia (2005) and Zhang (2011), and assume that X follows a Brownian Semimartingale. For $t \in [0, 1]$,

$$X_t = X_0 + \int_0^t b_u \, du + \int_0^t \sigma_u dW_u, \qquad (3)$$

where W is a d-dimensional Brownian motion, σ_u is a $d \times d$ stochastic volatility process. We are considering the bivariate case $d = 2$. The spot covariance of X is the $d \times d$ matrix $c_u = \sigma_u \sigma_u^\top$ with $c_{u,ij}$ its (i,j) element. We choose the following popular measure of beta of the asset,

$$\beta := \frac{\int_0^1 c_{u,12}du}{\int_0^1 c_{u,22}du} = \frac{\langle X^S, X^F \rangle}{\langle X^F, X^F \rangle}, \tag{4}$$

where $\langle X^S, X^F \rangle$ is the quadratic covariation between X^S and X^F over the time period $[0,1]$. This measure of beta has been used in, for example, Andersen et al. (2005b, 2006), Barndorff-Nielsen and Shephard (2004), and Bandi and Russell (2005). 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. Alternative beta measures include the spot beta (a time-localised version of (4)), see, for example, Mykland and Zhang (2006), Li, Todorov, and Tauchen (2017), Kalnina and Tewou (2019), and Yang (2020), and integrated beta, see, for example, Aït-Sahalia, Kalnina, and Xiu (2020).

Clearly, our parameter of interest β in Equation (4) violates Assumption A1. Hence, there is no estimator of β, for which a univariate subsampling method can be applied directly. However, β is a function, say γ, of another bivariate parameter θ, which does satisfy Assumption A1,

$$\beta = \gamma(\theta) = \frac{\theta_2}{\theta_1}, \text{ where } \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}. \tag{5}$$

In the above, $\theta_1 = \int_0^1 c_{u,22}du$ and $\theta_2 = \int_0^1 c_{u,12}du$. An estimator of the beta is $\widehat{\beta} = \gamma\left(\widehat{\theta}\right)$. If $\widehat{\theta}^{TS}$ is a TS estimator, we call $\widehat{\beta}^{TS} = \gamma(\widehat{\theta}^{TS})$ the Two Scales Realized Beta.

If the subsampling assumptions are satisfied by $\widehat{\theta}$, we can obtain the asymptotic variance-covariance matrix of $\widehat{\theta}$ by subsampling. Corollary 3 shows this is indeed the case for the TS estimator. An application of the Delta method delivers an estimator of the asymptotic variance of $\widehat{\beta} = \gamma\left(\widehat{\theta}\right)$.

We now introduce the Two Scales estimators. The data is generated by a contaminated process $Y = X + \varepsilon$. The additive measurement error ε represents the market microstructure effects such as the bid-ask bounce. Suppose we have $n+1$ equidistant observations on Y over $[0,1]$ at times $0 = t_0 < t_1 < \cdots < t_n = 1$. Then, an example of an estimator of θ_2 in Equation (5) is the Two Scales estimator of Zhang (2011),

$$\widehat{\langle X^S, X^F \rangle}^{\text{TS}} = \left[Y^S, Y^F\right]^{(G_1)} - \frac{\overline{n}_{G_1}}{\overline{n}_{G_2}} \left[Y^S, Y^F\right]^{(G_2)} \tag{6}$$

where

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

To estimate θ_1, we use analogously defined quantities $\widehat{\langle X^F, X^F \rangle}^{\text{TS}}$ and $\left[Y^F, Y^F\right]^{(G_l)}$. In the above, $\overline{n}_{G_j} = \frac{n-G_j+1}{G_j}$ for $j = 1,2$, $G_1 = \lfloor \varphi^{\text{TS}} n^{2/3} \rfloor$ for some tuning parameter φ^{TS}, and $G_2/G_1 \to 0$.

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 (when viewed as a process in index i) stationary and exponentially α-mixing. Also, $E\epsilon^{4+\delta} < \infty$ for some $\delta > 0$.

The asymptotic distribution of the TS estimator of the θ vector in Equation (5) is

$$n^{1/6}\left(\begin{pmatrix} \widehat{\langle X^F, X^F \rangle}^{\text{TS}} \\ \widehat{\langle X^F, X^S \rangle}^{\text{TS}} \end{pmatrix} - \begin{pmatrix} \langle X^F, X^F \rangle \\ \langle X^F, X^S \rangle \end{pmatrix}\right) \Rightarrow MN\left(0, V^{\text{TS}}\right), \tag{7}$$

where

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

was first derived in Aït-Sahalia, Mykland, and Zhang (2011). In Equation (8), the first summand is clearly a smooth function of c_u. The remaining summands 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 expression for V_{22}^{TS} is rather complicated and can be found in Zhang (2011) (Theorem 6 and Equation (52) in the Appendix), while the expression for V^{TS} does not seem to be available in the literature, although it can be derived following the arguments of Zhang (2011). Aït-Sahalia, Mykland, and Zhang (2011) do not provide estimators for the asymptotic variances of the proposed estimators.

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

Corollary 3. Suppose log-price X satisfies Equation (3) with $d = 2$, where b_s and σ_s are adapted and càdlàg. Let $\widehat{\theta}_n$ be the TS estimator defined by (6), 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 $\text{cov}\left(\epsilon_1, \epsilon_{G_2}\right) = o\left(n^{-1/2}\right)$, $G_2 \to \infty$, and $G_2/G_1 \to 0$. Let V be defined by V^{TS} in Equation (7), and $\widehat{V}^{\text{sub}}$ be defined by (2). Suppose Assumptions A2, A3 and N hold. Then,

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

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 $Jm^{\alpha}/n^{2/3+\alpha} \to 0$.

We conclude this section with a remark about using data with nonequidistant and asynchronous observations. Theorem 1 and Assumptions A1–A4 do not rely on any specific sampling scheme or synchronization method for asynchronous data. Corollary 3 assumes that observations are equidistant. If the data is irregular, we conjecture that as long as the TS estimator is implemented in tick time, and under additional regularity assumptions on the durations, the conclusion of Corollary 3 would remain valid, see Aït-Sahalia and Jacod (2014), Barndorff-Nielsen et al. (2008), and Hayashi, Jacod, and Yoshida (2011). If the data is in addition asynchronous, and the Refresh Time sampling is used, Corollary 3 should remain valid under the regularity assumptions on the refresh times, see Barndorff-Nielsen et al. (2011).

4. Econometric Methods of Analysis of the Dynamic Properties of the Betas

We now introduce two sets of tools for studying the dynamics of financial betas, which build on our subsampling method. We illustrate these tools in our empirical illustration in Section 6. First, we briefly outline a test for time-invariance of betas. Second, we introduce the Measurement-Error-Corrected (MEC) estimator, which uses the estimators of the asymptotic variances of betas to correct for the biases of the OLS estimator in a dynamic model for betas.

4.1. Tests of Time-Invariance of Betas

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

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

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 $i = 1, 2, \dots, k$ (see, e.g., sec. 5 of Zhang, Mykland, and Aït-Sahalia 2005), with the asymptotic distribution

$$\tau_{n_1} \Phi \Sigma^{-1/2} \left(\widehat{\beta} - \beta\right) \Rightarrow N(0, I_k), \tag{9}$$

$$\Sigma = \operatorname{diag}(V_1, V_2, \dots, V_k), \text{ and } \Phi = \operatorname{diag}(\phi_1, \phi_2, \dots, \phi_k),$$

where Σ is a diagonal matrix containing the asymptotic variances of $\widehat{\beta}_i$, 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 $\tau_{n_i} = \tau_{n_1}(\phi_i + o(1))$ for some positive finite constants ϕ_i, $i = 2, \dots, k$.

The hypothesis of time-invariant β is

$$H_0 : \beta_1 = \dots = \beta_k, \text{ versus } H_1 : \beta_i \neq \beta_j \text{ for some } i \text{ and } j. \tag{10}$$

Rewriting the null hypothesis as $H_0 : \beta_i = \beta_1$ for $i = 2, \dots, k$ as usual we have $\tau_{n_1}^2 \widehat{\beta}' \underline{\Delta}' \left(\underline{\Delta} \Phi^{-1} \Sigma \Phi^{-1} \underline{\Delta}'\right)^{-1} \underline{\Delta} \widehat{\beta} \Rightarrow \chi^2_{k-1}$, where the matrix $\underline{\Delta}$ is $(k-1) \times k$ and is defined as $\underline{\Delta} = \left(-\mathrm{i}_{k-1}, \mathrm{I}_{k-1}\right)$, 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^2_{k-1} follows from (9). 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

$$\tau_{n_1}^2 \widehat{\beta}' \underline{\Delta}' \left(\underline{\Delta} \Phi^{-1} \widehat{\Sigma} \Phi^{-1} \underline{\Delta}'\right)^{-1} \underline{\Delta} \widehat{\beta} \Rightarrow \chi^2_{k-1}. \tag{11}$$

We can similarly test the null hypothesis of constant betas jointly across stocks, that is,

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

where $\beta_i^{(j)}$ denotes the beta of the jth stock in the ith 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.

4.2. Estimation of Models of the Dynamics: The Measurement Error Correction

When the hypothesis of time-invariant betas is rejected, 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, the unconditional Capital Asset Pricing Model has time-invariant betas and is known not to be empirically realistic; to alleviate this problem, various conditional CAPM models have been proposed, in which betas vary with variables in investors' information set. If the model includes lagged betas, the errors-in-variables problem can arise, which leads to inconsistent parameter estimates. This problem can however, be addressed if we have access to the variance of the estimation error.

To illustrate the measurement error correction, consider the dynamic specification for betas in the conditional CAPM of Adrian and Franzoni (2009). In this specification, the variation in betas has a component that depends on exogenous regressors, an autoregressive component, and an idiosyncratic component. A similar beta specification is estimated by Andersen et al. (2005b). It can be written as

$$\beta_i = \rho_1 \beta_{i-1} + \dots + \rho_p \beta_{i-p} + \gamma' X_i + U_i, \quad i = p+1, \dots, k, \tag{12}$$

where β_i is the value of financial beta in ith time period, and X_i collect the conditioning variables, as well as the intercept (we follow the notation of Section 4.1). Both Andersen et al. (2005) and Adrian and Franzoni (2009) estimate their models using the Kalman Filter.

While the conditioning variables X_i are usually only available at lower frequencies, we can exploit the high-frequency asset price data to estimate (12) by replacing β_i with the high-frequency estimate $\widehat{\beta}_i$ for each i. (For example, Andersen et al. (2005) use daily returns to calculate quarterly Realized Betas.) However, this substitution leads to the problem of measurement errors in covariates, because it puts estimated $\widehat{\beta}_{i-1}$ on the right-hand side of Equation (12). 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 $\rho_1, \dots, \rho_p$, 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, for example, Andersen, Bollerslev, and Meddahi (2005a) who adjust the forecasting loss functions.

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

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

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

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

Let us consider the effect of replacing β_i with $\widehat{\beta}_i$. Remember that $\varepsilon_i = \widehat{\beta}_i - \beta_i \sim_a N\left(0, V_i/\tau_{n_i}^2\right)$, and that $E\left[\varepsilon_i \varepsilon_j\right] = 0$ for $i \neq j$. Usually ε_j are also uncorrelated with X_i and hence, we have

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

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

where

$$\Xi_k = \frac{1}{k-p-1}\sum_{i=p+1}^{k}\operatorname{diag}\left(V_{i-1}/\tau_{n_{i-1}}^2, \ldots, V_{i-p}/\tau_{n_{i-p}}^2, 0, \ldots, 0\right). \quad (13)$$

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 (13). Hence, in the empirical analysis we can use the Measurement-Error-Corrected (MEC) estimator

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

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 for finance applications: for example, uncertainty about betas varies with firm-specific information flows, see Patton and Verardo (2012).

Considering the linear model (12) allows us to obtain explicit expression for the estimator $\widehat{\theta}^{\text{MEC}}$. 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.

5. Simulation Studies

We now illustrate the subsampling method in the example of the Two Scales Realized Beta estimator calculated over one day. This 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 estimator in finite samples. We also discuss the choice of the parameters of the $\widehat{\beta}^{\text{TS}}$ estimator G_1 and G_2. As a robustness check, Appendix R in the supplementary materials repeats the analysis with a two-factor stochastic volatility model. Appendices D and M in the supplementary materials contain the tables.

5.1. Monte Carlo Designs

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:

$$dX_t^{(j)} = \left(\alpha_1^{(j)} - c_t^{(j)}/2\right)dt + \sigma_t^{(j)}dW_t^{(j)}$$

$$dc_t^{(j)} = \alpha_2^{(j)}\left(\alpha_3^{(j)} - c_t^{(j)}\right)dt + \alpha_4^{(j)}\left(c_t^{(j)}\right)^{1/2}dB_t^{(j)},\ j = 1, \ldots, 7,$$

where $c_t^{(j)} = \left(\sigma_t^{(j)}\right)^2$, and $W_t^{(j)}$ and $B_t^{(j)}$ are Brownian Motions with $\operatorname{corr}\left(W_t^{(j)}, B_t^{(j)}\right) = \rho^{(j)}$. 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 $X^{(1)}$, …, $X^{(7)}$ 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 6.1 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:

$$\left(\alpha_4^{(j)}\right)^2 = \frac{\left\langle c^{(j)}, c^{(j)}\right\rangle}{\left\langle X^{(j)}, X^{(j)}\right\rangle}. \quad (14)$$

The numerator is the quadratic variation of the spot variance of the jth asset. To estimate it, we use moderate frequency price returns and the estimator of Vetter (2012) with jump truncation, see Jacod and Rosenbaum (2015). (We use 5-min frequency, which is a simple and popular choice. Bandi and Russell (2008) provide the optimal frequency for the realized variance.) We estimate the denominator in the above with truncated realized variance, see Mancini (2001). 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 minimize 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 $\operatorname{corr}\left(W_t^{(j)}, W_t^{(7)}\right) = \varrho^{(j)}, j = 1, \ldots, 6$, to the value of the realized beta with 50-tick observations of the jth stock. The realized beta estimator is defined as $\widehat{\beta}^{\text{RV}} = \widehat{\theta}_2^{\text{RV}}/\widehat{\theta}_1^{\text{RV}}$ where $\widehat{\theta}_1^{\text{RV}} = [X^F, X^F] = \sum_{j=1}^{n}(X_{t_j}^F - X_{t_{j-1}}^F)(X_{t_j}^F - X_{t_{j-1}}^F)$ is the realized variance of the market returns, $\widehat{\theta}_2^{\text{RV}} = [X^S, X^F] = \sum_{j=1}^{n}(X_{t_j}^S - X_{t_{j-1}}^S)(X_{t_j}^F - X_{t_{j-1}}^F)$ is the realized covariance of the stock and the market returns, and where θ is defined in Equation (5). In this model, the beta over [0,1] is

$$\beta^{(j)} = \varrho^{(j)}\int_0^1 \sigma_u^{(j)}\sigma_u^{(7)}du \Bigg/ \int_0^1\left(\sigma_u^{(7)}\right)^2 du,\ j = 1, \ldots, 6. \quad (15)$$

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

Microstructure noise is simulated as a normally distributed white noise with variance $\xi^{(j)} IV^{(j)}$, where $\xi^{(j)}$ is a noise-to-signal ratio, and $IV^{(j)}$ is the weekly integrated volatility of the jth stock, see, for example, Hansen and Lunde (2006). We set $\xi^{(j)}$ to be either 0 or 0.0001; all estimated values of the noise-to-signal ratio are between these two values, see Table D.1, in the supplementary materials. We simulate the noise to be iid to minimize 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, $n_j, j = 1, \dots, 7$, to the data (see the first column in Table D.1 in the supplementary materials). 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, (see Barndorff-Nielsen et al. 2011) resulting in a random number of observations to be used for estimation.

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

For the TS parameter G_1, we use a range of values that include the main recommendations in the literature, and the current section provides the implementation details for choosing this range. We calculate the Two Scale estimator following the Equation (6) with $G_2 = 1$ (for the iid noise) and the finite sample adjustment suggested in Zhang, Mykland, and Aït-Sahalia (2005).

As a rough starting point for choosing a range of G_1 for TS-beta, we use the two data-driven principles of selecting G_1 parameter for the TS variance estimator that are available in the literature: by minimization of the asymptotic variance (Zhang, Mykland, and Aït-Sahalia 2005) and by minimization of the finite sample MSE, see Bandi and Russell (2011), henceforth BR. We denote the theoretical values given by these rules by $G_1^\star$ and G_1^{BR}, and the estimated values by $\widehat{G}_1^\star$ and $\widehat{G}_1^{\mathrm{BR}}$. Simplifying the notation of Section 3 to the case of one variable, the precise expressions are as follows

- $G_1^\star = \varphi^\star n^{2/3}$ with $\varphi^\star = \left(\frac{16\omega^4}{\frac{4}{3}E(IQ)}\right)^{1/3}$ where $\omega^2 = E\epsilon^2$ and $IQ = \int_0^1 c_u^2 du$.
- $\widehat{G}_1^\star$ replaces the unobservable ω^2 and IQ in $G_1^\star$ with their estimators.
- $G_1^{\mathrm{BR}} = \mathrm{argmin}_{G_1} \mathrm{MSE}(G_1)$ where $\mathrm{MSE}(G_1) = \mathrm{MSE}\left(G_1, n, IQ, IV, \omega^2\right)$ is the finite sample MSE of the TS variance estimator, presented in the Corollary 3 of Bandi and Russell (2011). $IV = \int_0^1 c_u du$.
- $\widehat{G}_1^{\mathrm{BR}}$ replaces the unobservable ω^2, IV, and IQ in G_1^{BR} with their estimators.

We use the following estimators to obtain $\widehat{G}_1^\star$ and $\widehat{G}_1^{\mathrm{BR}}$. To estimate the variance of the noise ω^2, we use the estimator (17) calculated with the highest frequency data. To estimate IV and IQ, for simplicity, we use Realized Variance and Realized Quarticity (Barndorff-Nielsen and Shephard 2004) calculated with 5-min data. Table M.1 in the supplementary materials shows the values of the four different rules for choosing G_1 in the six simulated scenarios.

Another, more heuristic, recommendation for the choice of G_1 is to choose the slow scale to correspond to 5-min frequency, see, for example, sec. 4.2 of Aït-Sahalia and Mancini (2008). This rule-of-thumb implies that G_1 depends only on the number of observations, and it is different across stocks and weeks. Table D.2 in the supplementary materials gives an idea of the average values of the resulting G_1. For example, INTC, our most liquid stock, has 24,601 observations after pairwise synchronization with the market, which implies an average value of G_1 of 63.1. Other stocks have smaller number of observations, especially after joint synchronization with all stocks and the market.

Table M.1 in the supplementary materials shows the finite sample MSE across simulations of the $\widehat{\beta}^{TS}$ estimator for a range of G_1. In all cases considered, the BR rule results in a smaller finite sample MSE than the rule based on the asymptotic variance, which in turn gives more precise estimates than the 5-min rule-of-thumb. (These favorable properties are not a priori guaranteed as G_1^{BR} and $G_1^\star$ are theoretically optimal for the TS variance, not TS-beta.)

We conclude that a relatively large range of values of G_1 is needed to include all practically relevant implementation versions of $\widehat{\beta}^{\mathrm{TS}}$: on the one hand, very small values of G_1 deliver precise estimates of $\widehat{\beta}^{\mathrm{TS}}$; on the other hand, the 5-min rule-of-thumb results in relatively large values of G_1. Hence, the simulation exercises below use values ranging from $G_1 = 3$ to $G_1 = 50$.

5.3. Choice of the Subsampling Parameters (Tables M.2–M.5)

To obtain a data-driven choice for m and J, we fit a parametric reference model to our data, see Section 2.1 for details.

Tables M.2–M.3 in the supplementary materials show the average over simulations of the subsampled variances $\widehat{V}^{\mathrm{sub}}$ of the $\widehat{\beta}^{\mathrm{TS}}$ estimator. (The quantities V^{FS}, V^{Theo}, and $\hat{V}^{\mathrm{pl}}$ are discussed below.) We find 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 use the same m and J for every week.

We also consider the coverage of the nominal 95% confidence interval based on the $\widehat{\beta}^{TS}$ estimator. The results are given in Tables M.4 and M.5 in the supplementary materials. We again find that the results are relatively flat over wide ranges of m and J. To limit the number tables, Tables M.4–M.5 consider scenario (3) (where the design is calibrated to IBM data), but the same conclusions hold for other scenarios.

5.4. Performance of the Subsampling Estimator for $\widehat{\beta}^{TS}$ (Tables M.6–M.9)

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, for example, in the case of $\widehat{\beta}^{\mathrm{TS}}$, it is the average over simulations of $n^{1/3}\left(\widehat{\beta}^{\mathrm{TS}}-\beta\right)^2$. 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^{Theo}, as well as its estimated counterpart $\hat{V}^{\mathrm{pl}}$ using the same estimation approach as for $\widehat{G}_1^{\star}$ and $\widehat{G}_1^{\mathrm{BR}}$.

Tables M.6 and M.7 in the supplementary materials contain the results for the $\widehat{\beta}^{\mathrm{TS}}$ estimator. 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 in the supplementary materials, these smaller values of G_1 where subsampling method performs particularly well compared to $\widehat{V}^{\mathrm{pl}}$ (and V^{Theo}), are often close to the values of G_1 that minimize the finite sample MSE of the $\widehat{\beta}^{\mathrm{TS}}$.

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

5.5. Performance of the Subsampling Estimator for TSRV (Tables M.10–M.13)

Finally, we repeat the above analysis for the TS estimator of the variance of the stock (TSRV). As opposed to $\widehat{\beta}^{\mathrm{TS}}$, this estimator does not involve any nonlinear transformations. The results are collected in Tables M.10–M.13 in the supplementary materials. Again, the subsampling method is more robust than the plug-in method $\widehat{V}^{\mathrm{pl}}$.

6. 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 min), 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 $\widehat{\theta}^{\mathrm{RV}}$ (recall notation in Equation (5)). 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.

6.1. Description of the 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), 3M Co. (MMM), and Microsoft Co. (MSFT). To proxy for the market portfolio, we use Standard and Poor's Depository Receipts (SPIDERS for short, ticker symbol SPY), which is a 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 et al. (2009) and remove jumps with the thresholding methodology, see Mancini (2001).

We start by analyzing the high frequency data. Table D.1 in the supplementary materials 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, for example, Griffin and Oomen (2008). In the full dataset 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. Columns 2 and 3 in Table D.1 in the supplementary materials show two estimators of the noise variance. The first estimator of the noise variance is the realized variance divided by twice the sample size,

$$\widehat{\omega}^2 = [X, X]/2n, \tag{16}$$

see Bandi and Russell (2006). The second estimator of the noise variance corrects a small sample bias,

$$\widetilde{\omega}^2 = \left([X, X] - \widehat{\langle X, X\rangle}^{\mathrm{TSRV}}\right)/2n, \tag{17}$$

similarly to, for example, Hansen and Lunde (2006). The estimated values of ω^2 are along the lines of those reported previously, see, for example, Bandi and Russell (2006).

Table D.2 in the supplementary materials 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 $\widetilde{\omega}^2$ are the same with and without synchronization.

Figure D.3 in the supplementary materials contains the signature plots of the RV-based beta (or realized beta) for each stock, that is, plots of RV-based beta against the frequency used in its calculation. (RV-based beta (or realized beta) is defined

Table 1. Estimated dynamics of the betas. OLS and MEC estimators are described in Sections 4.2 and 6.3.

Stock	ρ_1^{OLS}	ρ_2^{OLS}	γ_1^{OLS}	$\gamma_{\text{MKT}}^{\text{OLS}}$	$\gamma_{\text{TERM}}^{\text{OLS}}$	$\gamma_{\text{HML}}^{\text{OLS}}$	ρ_1^{MEC}	ρ_2^{MEC}	γ_1^{MEC}	$\gamma_{\text{MKT}}^{\text{MEC}}$	$\gamma_{\text{TERM}}^{\text{MEC}}$	$\gamma_{\text{HML}}^{\text{MEC}}$
AIG	0.60		0.31				0.73		0.21			
AIG	0.56	0.09	0.27				0.81	−0.08	0.20			
AIG	0.61		0.32	−0.27	0.06	−0.03	0.76		0.19	−0.43	0.04	−0.04
AIG	0.59	0.07	0.27	0.16	0.07	−0.03	0.89	−0.11	0.20	−0.02	0.07	−0.04
GE	0.24		0.45				0.32		0.41			
GE	0.21	0.05	0.43				0.29	0.05	0.39			
GE	0.24		0.45	1.60	0.08	0.01	0.34		0.40	1.69	0.07	0.02
GE	0.23	0.05	0.43	1.55	0.07	0.02	0.32	0.04	0.38	1.65	0.06	0.02
IBM	0.39		0.48				0.57		0.34			
IBM	0.37	0.03	0.47				0.61	−0.09	0.38			
IBM	0.33		0.55	0.17	0.13	−0.01	0.49		0.42	0.21	0.10	−0.01
IBM	0.33	−0.03	0.57	0.08	0.13	−0.01	0.57	−0.17	0.49	0.11	0.12	−0.01
INTC	0.64		0.42				0.68		0.37			
INTC	0.44	0.31	0.28				0.46	0.31	0.25			
INTC	0.40		0.60	3.73	−0.55	−0.01	0.45		0.56	3.70	−0.51	−0.00
INTC	0.31	0.17	0.51	1.69	−0.58	0.00	0.34	0.17	0.47	1.68	−0.55	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	0.56		0.41	−0.89	−0.03	−0.04	0.73		0.25	−0.68	−0.00	−0.04
MMM	0.46	0.15	0.36	−1.62	−0.03	−0.04	0.70	0.01	0.27	−0.93	−0.02	−0.04
MSFT	0.43		0.46				0.48		0.43			
MSFT	0.42	0.04	0.44				0.47	0.02	0.42			
MSFT	0.35		0.53	0.18	−0.17	−0.07	0.39		0.50	0.19	−0.16	−0.07
MSFT	0.36	−0.02	0.53	−0.16	−0.19	−0.07	0.41	−0.05	0.51	−0.11	−0.18	−0.07

NOTE: Here ρ_j are the AR(p) coefficients, γ_1 is the intercept, and γ_{MKT}, γ_{TERM} and γ_{HML} are the coefficients on MKT_{i-1}, TERM_{i-1}, and HML_{i-1}.

as $[X, Y]/[X, X]$, and the interval of estimation is the whole year 2006.) There is a strong bias toward zero at the highest frequencies. This bias is a combination of, first, the realized volatilities exploding due to the microstructure noise dominating at the highest frequencies, second, the realized covariances being biased toward zero at these frequencies, that is, the Epps effect. This confirms that the realized betas should not 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).

6.2. Testing the Constant Beta Hypothesis

Figure E.3, in the supplementary materials shows plots of estimated betas using $\widehat{\beta}^{\text{RV}}_{10\text{min}}$ and $\widehat{\beta}^{\text{TS}}$ 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 $\widehat{\beta}^{\text{RV}}_{10\text{min}}$ was also graphed by Andersen et al. (2006) in their Figures 13–15, except they used 10 min and daily data to calculate estimated betas over intervals of one quarter. In Figure E.3 in the supplementary materials, we see that beta is estimated more precisely using full record transaction prices. The two parameters in $\widehat{\beta}^{\text{TS}}$ were chosen as follows. We set $G_2 = 3$ as no stocks (after synchronization) display autocorrelated returns beyond the second lag. G_1 was chosen as $5G_2$. Recall that we choose $J = 500$ and $m = 3000$, see Section 5.3.

Table E.2 in the supplementary materials 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 (10). Four different tests are implemented based on four estimators: $\widehat{\beta}^{\text{RV}}_{5\text{min}}$, $\widehat{\beta}^{\text{RV}}_{10\text{min}}$, $\widehat{\beta}^{\text{RV}}_{20\text{min}}$ and $\widehat{\beta}^{\text{TS}}$. 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 $\widehat{\beta}^{\text{TS}}$ 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.1, in the supplementary materials 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.

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 min similarly contains more information than RV-based beta at 20 min. It is well known that RV-based beta cannot be used at frequencies much higher than 5 min without accounting for the market microstructure noise. Our results suggest that the noise-robust estimators such as the $\widehat{\beta}^{\text{TS}}$ are able to extract additional information on time variation in betas while being robust to the noise contaminations.

6.3. Estimation of the Dynamics of Betas and the Impact of MEC

In this section, we illustrate empirically the OLS and MEC estimators introduced in Section 4.2. We follow Adrian and Franzoni (2009) and choose the following three regressors X

in Equation (12): return on value-weighted market, the term spread, and the value spread. The market is chosen as the S&P500 index; the term spread is constructed as the difference between the ten-year and the three-month constant maturity Treasury yield; the value spread is chosen as the return to the HML factor from Fama and French (1992). For each of the six stocks we estimate model (12) with $p = 1$ and $p = 2$, where the betas are estimated by $\widehat{\beta}^{\mathrm{TS}}$. We first consider the case with only the intercept ($X_i = 1$). Then, we consider $X_i = (1, \mathrm{MKT}_{i-1}, \mathrm{TERM}_{i-1}, \mathrm{HML}_{i-1})'$, where MKT_{i-1}, TERM_{i-1}, and HML_{i-1} are the (low frequency) S&P500 return, the term spread and the value spread at the end of week $i-1$. For comparison, we include both the results for the naive OLS estimator $\widehat{\theta}^{\mathrm{OLS}}$ and for the measurement error corrected estimator $\widehat{\theta}^{\mathrm{MEC}}$. Table 1 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) and Adrian and Franzoni (2009). 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 finding is in line with Hansen and Lunde (2014) who use an instrumental variable approach. Our conditioning variables do not appear to be an important determinants of the betas in our dataset. For the OLS estimator, MKT, TERM, and HML are 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 fromzero 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 article (see Andersen, Bollerslev, and Meddahi 2005a for a discussion).

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.

7. Conclusion

We propose a subsampling method to estimate the VCVs of high frequency estimators that does not rely on knowing the expressions of these matrices. We prove the validity of the multivariate inference method for a general class of estimators including many estimators of the integrated covariance matrices; our assumptions are weaker than those available in the literature in several empirically relevant directions. This multivariate method has the appealing property of always giving estimates of the VCVs that are positive semidefinite. Our simulation study suggests that when the asymptotic variance differs from the finite sample variance, the subsampling-based inference is more reliable than the plug-in approach.

We consider two applications of the multivariate subsampling method to the study of time variation in financial betas. First, we implement tests of beta time-invariance of six stocks on the NYSE. Second, we provide and implement Measurement-Error-Corrected estimators of the dynamics of betas in Adrian and Franzoni (2009). After the bias-correction, the betas appear to be substantially more persistent than the naive estimators suggest.

Supplementary Materials

All proofs are collected in the Appendices A–C in the supplementary materials. Appendix D in the supplementary materials contains the tables and figures for the summary statistics of data. Appendix E in the supplementary materials contains the tables and figures for the empirical illustration in Section 6. Appendix M in the supplementary materials contains the tables for the simulations in Section 5. Appendix R in the supplementary materials contains an additional simulation study.

Acknowledgments

I would like to thank an associate editor and four 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 various seminars and conferences. Earlier drafts of this article were previously circulated under the title "Nonparametric Tests of Time Variation in Betas."

Funding

I gratefully acknowledge financial support from Institut de finance mathématique de Montréal, SSHRC, and FQRSC.

References

Adrian, T., and Franzoni, F. (2009), "Learning about beta: Time-Varying Factor Loadings, Expected Returns, and the Conditional CAPM," *Journal of Empirical Finance*, 16, 537–556. [539,543,547,548]

Aït-Sahalia, Y., Fan, J., and Xiu, D. (2010), "High-Frequency Covariance Estimates with Noisy and Asynchronous Data," *Journal of the American Statistical Association*, 105, 1504–1517. [541]

Aït-Sahalia, Y., and Jacod, J. (2014), *High-Frequency Financial Econometrics*, Princeton, NJ: Princeton University Press. [539,540,542]

Aït-Sahalia, Y., Kalnina, I., and Xiu, D. (2020), "High-Frequency Factors Models and Regressions," *Journal of Econometrics*, 216, 86–105. [542]

Aït-Sahalia, Y., and Mancini, L. (2008), "Our of Sample Forecasts of Quadratic Variation," *Journal of Econometrics*, 147, 17–33. [545]

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011), "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise," *Journal of Econometrics*, 160, 190–203. [542,546]

Aït-Sahalia, Y., and Xiu, D. (2019), "Principal Component Analysis of High Frequency Data," *Journal of the American Statistical Association*, 114, 287–303. [540]

Andersen, T., T. Bollerslev, and N. Meddahi (2005a), "Correcting the Errors: Volatility Forecast Evaluation Using High Frequency Data and Realized Volatilities," *Econometrica*, 73, 279–296. [543,548]

——— (2005b), "A Framework for Exploring the Macroeconomic Determinants of Systematic Risk," *American Economic Review*, 95, 398–404. [539,542,543]

——— (2006), "Realized Beta: Persistence and Predictability," in *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger*, eds. T. Fomby and D. Terrell (pp. 1–39). Bingley: Emerald Publishing. [539,542,547,548]

Andrews, D. (1991), "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858. [541]

Bandi, F. M., and Russell, J. R. (2005), "Realized Covariation, Realized Beta, and Microstructure Noise," Technical Reports, University of Chicago. [542]

——— (2006), "Separating Microstructure Noise from Volatility," *Journal of Financial Economics*, 79, 655–692. [546]

——— (2008), "Microstructure Noise, Realized Volatility and Optimal Sampling," *Review of Economic Studies*, 75, 339–369. [544]

——— (2011), "Market Microstructure Noise, Integrated Variance Estimators, and the Accuracy of Asymptotic Approximations," *Journal of Econometrics*, 160, 145–159. [538,541,545]

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008), "Designing Realized Kernels to Measure Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536. [542]

——— (2009), "Realised Kernels in Practice: Trades and Quotes," *Econometrics Journal*, 12, C1–C32. [546]

——— (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. [541,542,545]

Barndorff-Nielsen, O. E., and Shephard, N. (2004), "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72, 885–925. [542,545,546,547]

Bekaert, G., and Wu, G. (2000), "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, 1–42. [539]

Bibinger, M. (2012), "An Estimator for the Quadratic Covariation of Asynchronously Observed Ito Processes with Noise: Asymptotic Distribution Theory," *Stochastic Processes and their Applications*, 122, 2411–2453. [541]

Bibinger, M., and Reiss, M. (2014), "Spectral Estimation of Covolatility from Noisy Observations Using Local Weights," *Scandinavian Journal of Statistics*, 41, 23–50. [541]

Bollerslev, T., and Zhang, B. Y. B. (2003), "Measuring and Modelling Systematic Risk in Factor Pricing Models Using High-Frequency Data," *Journal of Empirical Finance*, 10, 533–558. [542]

Braun, P. A., Nelson, D. B., and Sunier, A. M. (1995), "Good News, Bad News, Volatility, and Betas," *Journal of Finance*, 50, 1575–1603. [539]

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010), "Pre-Averaging Estimators of the Ex-Post Covariance Matrix in Noisy Diffusion Models with Non-Synchronous Data," *Journal of Econometrics*, 159, 116–133. [541]

Christensen, K., Podolskij, M., Thamrongrat, N., and Veliyev, B. (2017), "Inference from High-Frequency Data: A Subsampling Approach," *Journal of Econometrics*, 197, 245–272. [539]

Fama, E. F., and French, K. R. (1992), "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47, 427–465. [539,548]

Fama, E. F., and MacBeth, J. D. (1973), "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636. [539]

Griffin, J. E., and Oomen, R. C. (2008), "Sampling Returns for Realized Variance Calculations: Tick Time or Transaction Time?" *Econometric Reviews*, 27, 230–253. [546]

Hansen, P. R., and Lunde, A. (2006), "Realized Variance and Market Microstructure Noise," *Journal of Business and Economic Statistics*, 24, 127–161. [545,546]

——— (2014), "Estimating the Persistence and the Autocorrelation Function of a Time Series that is Measured with Error," *Econometric Theory*, 30, 60–93. [548]

Hayashi, T., Jacod, J., and Yoshida, N. (2011), "Irregular Sampling and Central Limmit Theorems for Power Variations: The Continuous Case," *Annales de l'Institut Henri Poincaré*, 47, 1197–1218. [542]

Heston, S. (1993), "A Closed-form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options," *Review of Financial Studies*, 6, 327–343. [541,544]

Ikeda, S. S. (2016), "A Bias-Corrected Estimator of the Covariation Matrix of Multiple Security Prices when both Microstructure Effects and Sampling Durations are Persistent and Endogenous," *Journal of Econometrics*, 193, 203–214. [539]

Jacod, J., and Rosenbaum, M. (2015), "Estimation of Volatility Functionals: The Case of a $\sqrt{n}$ Window," in *Large Deviations and Asymptotic Methods in Finance*, eds. P. K. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, and J. Teichmann (pp. 559–590), Cham: Springer. [544]

Jostova, G., and Philipov, A. (2005), "Bayesian Analysis of Stochastic Betas," *Journal of Financial and Quantitative Analysis*, 40, 747–778. [539]

Kalnina, I. (2011), "Subsampling High Frequency Data," *Journal of Econometrics*, 161, 262–283. [539,540]

Kalnina, I., and Linton, O. (2007), "Inference about Realized Volatiolity Using Infill Subsampling," Technical Report, London School of Economics. [539]

Kalnina, I., and Tewou, K. (2019), "Cross-Sectional Dependence in Idiosyncratic Volatility," Technical Report, North Carolina State University. [542]

Lahiri, S. N., Kaiser, M. S., Cressie, N., and Hsu, N.-J. (1999), "Prediction of Spatial Cumulative Distribution Functions Using Subsampling," *Journal of the American Statistical Association*, 94, 86–97. [539]

Lazarus, E., Lewis, D. J., Stock, J. H., and Watson, M. W. (2018), "HAR Inference: Recommendations for Practice," *Journal of Business and Economic Statistics*, 36, 541–559. [541]

Li, J., Todorov, V., and Tauchen, G. (2017), "Adaptive Estimation of Continuous-Time Regression Models using High-Frequency Data," *Journal of Econometrics*, 200, 36–47. [542]

Mancini, C. (2001), "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47. [544,546]

Mykland, P. A., and Zhang, L. (2006), "ANOVA for Diffusions and Itô Processes," *Annals of Statistics*, 34, 1931–1963. [540,542]

——— (2017), "Assessment of Uncertainty in High Frequency Data: The Observed Asymptotic Variance," *Econometrica*, 85, 197–31. [539]

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [548]

Patton, A., and Verardo, M. (2012), "Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability," *Review of Financial Studies*, 25, 2789–2839. [539,544]

Politis, D. N., and Romano, J. P. (1994), "Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *Annals of Statistics*, 22, 2031–2050. [539]

Politis, D. N., Romano, J. P., and Wolf, M. (1999), *Subsampling*, New York: Springer-Verlag. [541]

Todorov, V., and Bollerslev, T. (2010), "Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks," *Journal of Econometrics*, 157, 220–235. [542]

Varneskov, R. T. (2016), "Flat-Top Realized Kernel Estimation of Quadratic Covariation with Nonsynchronous and NOisy Asset Prices," *Journal of Business and Economic Statistics*, 34, 1–22. [539]

Vetter, M. (2012), "Estimation of Integrated Volatility of Volatility with Applications to Goodness-of-fit Testing," Technical Report, Ruhr-Universität Bochum. [544]

Yang, X. (2020), "Time-Invariant Restrictions of Volatility Functionals: Efficient Estimation and Specification Tests," *Journal of Econometrics*, 215, 486–516. [542]

Zhang, L. (2011), "Estimating Covariation: Epps Effect and Microstructure Noise," *Journal of Econometrics*, 160, 33–47. [538,541,542,547]

——— (2012), "Implied and Realized Volatility: Empirical Model Selection," *Annals of Finance*, 8, 259–275. [540]

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [538,541,543,544,545]

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

by Ilze Kalnina

A Proofs

A.1 Proof of Theorem 1

Let

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

Sections A.1.1 and A.1.2 establish the following facts:

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

and

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

Introduce the following notation,

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

Assumption A4 with $s = J$ and (A.1) imply $\widehat{V}^{\text{infeasible}} - V = o_p(1)$. It remains to prove that

$$\widehat{V}^{sub} - \widehat{V}^{\text{infeasible}} = o_p(1). \tag{A.3}$$

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 (A.3) as follows,

$$\begin{aligned}\widehat{V}^{sub} - \widehat{V}^{\text{infeasible}} &= \frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\Big(-R_l^{short}R_l^{long\prime} - R_l^{short}d_l^{\prime} - R_l^{long}R_l^{short\prime} + R_l^{long}R_l^{long\prime} \\ &\quad + R_l^{long}d_l^{\prime} - d_l R_l^{short\prime} + d_l R_l^{long\prime} + d_l^{\prime}d_l^{\prime}\Big). \qquad \text{(A.4)}\end{aligned}$$

If we can show that

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 R_l^{short}R_l^{short\prime} = O_p(1), \qquad \text{(A.5)}$$

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 R_l^{long}R_l^{long\prime} = o_p(1), \qquad \text{(A.6)}$$

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 d_l d_l^{\prime} = o_p(1), \qquad \text{(A.7)}$$

then all terms in (A.4) are negligible by the Cauchy-Schwarz inequality, e.g.,

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

Equation (A.5) follows from $\widehat{V}^{\text{infeasible}} = V + o_p(1)$, since the left-hand-side of equation (A.5) is $\widehat{V}^{\text{infeasible}}$.

Equation (A.6) follows from

$$\frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2 R_l^{long}R_l^{long\prime} = \frac{J}{m}\left[\frac{1}{K}\frac{n}{m}V + o_p(1)\right] = o_p(1). \qquad \text{(A.8)}$$

The first equality in the above follows by Assumption A4 with $s = m$, which can be written as

$$\frac{1}{K}\sum_{l=1}^{K}\frac{m}{n}\tau_n^2 R_l^{long}R_l^{long\prime} - \frac{1}{K}\frac{n}{m}V = o_p(1),$$

since $V = \sum_{l=1}^{K} V_l^{long}$. The second equality in (A.8) follows by $K = \lfloor n/m \rfloor$ and $J/m = o(1)$. Equation (A.7) follows from equation (A.2), since the left-hand-side of (A.7) equals the left-hand-side of (A.2) times $n/(Km)$. □

A.1.1 Proof of equation (A.1)

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

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

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

Lemma A.1. *Suppose φ is a continuously differentiable functional. Let $A_n \le a_n < b_n \le B_n$ 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*

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

To prove equation (A.1), 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 A.1 we obtain

$$\begin{aligned} \left| V_{k_1k_2} - \sum_{l=1}^{K} \frac{m}{J} V^{short}_{l,k_1k_2} \right| &\le \frac{m}{n} \sum_{l=1}^{K} \left| \frac{1}{m/n} H^{\varphi}_{(l-1)m/n,lm/n} - \frac{1}{J/n} H^{\varphi}_{(l-1)m/n,[(l-1)m+J]/n} \right| \\ &\le \overline{\varphi_\nabla} \frac{m}{n} \sum_{l=1}^{K} \left\{ \frac{I_{(l-1)m/n,lm/n}}{m/n} + \frac{I_{(l-1)m/n,[(l-1)m+J]m/n}}{J/n} \right\}. \qquad (A.9) \end{aligned}$$

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

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

Hence

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

as long as $J/n\to 0$ and $m/n\to 0$, since $K\frac{m}{n}\to 1$. Using Markov inequality and equation (A.9) we obtain $V_{k_1k_2}-\sum_{l=1}^{K}\frac{m}{J}V_{l,k_1k_2}^{short}=o_P\left(1\right)$ and hence equation (A.1) holds. □
Proof of Lemma A.1: We omit subscript n and replace A_n,a_n,b_n,B_n with A,a,b,B, respectively.

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

A.1.2 Proof of equation (A.2)

Fix any k and let $\varphi\left(\cdot\right)$ denote $f_k\left(\cdot\right)$. Then, using Lemma A.1 in Section A.1.1,

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

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,

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

$$\leq B_c \frac{J^2}{n^2} K \left(\left(\frac{m}{n}\right)^{\alpha} + \left(\frac{J}{n}\right)^{\alpha}\right) \leq 2B_c \frac{J^2}{nm}\left(\frac{m}{n}\right)^{\alpha}.$$

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

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

so equation (A.2) follows from Assumption A3. $\square$

B Proof of Lemma 2

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 3

We prove Corollary 3 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

$$g_{11}(u) = \varphi^{TS}\frac{4}{3}c_{u,11}^2 + 8(\varphi^{TS})^{-2}Var\left(\epsilon^F\right)^2 + 16(\varphi^{TS})^{-2}\lim_{n\to\infty}\sum_{i=1}^{n}Cov\left(\epsilon_0^F, \epsilon_{i/n}^F\right)^2, \qquad \text{(A.10)}$$

and other elements of $g(u)$ are defined similarly (see Footnote 6 for details). Assumption A2 is directly assumed in the statement of Corollary (sufficient conditions are discussed after the statement of Assumption A2 in Section 2). Assumption A3 is the restriction on the subsample sizes, which is also directly assumed by the statement of the Corollary. For the case $s = J$, Assumption A4 is verified by the analogue of Lemma 7 of Kalnina (2011) for the multivariate case, where instead of adapting the results of Zhang et al. (2005) to a shrinking subsample, one uses the corresponding results of Zhang (2011). The proof of the case $s = m$ of Assumption A4 is the same. □

D Data: Summary Statistics

	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 D.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 using estimators in (16) and (17). Fourth column contains estimated noise-to-signal ratio, $\widehat{\xi} = \widehat{\omega}^2/\widehat{IV}$ where IV is estimated by the TSRV. 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 D.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 D.1 annotation for the meaning of the other column entries.

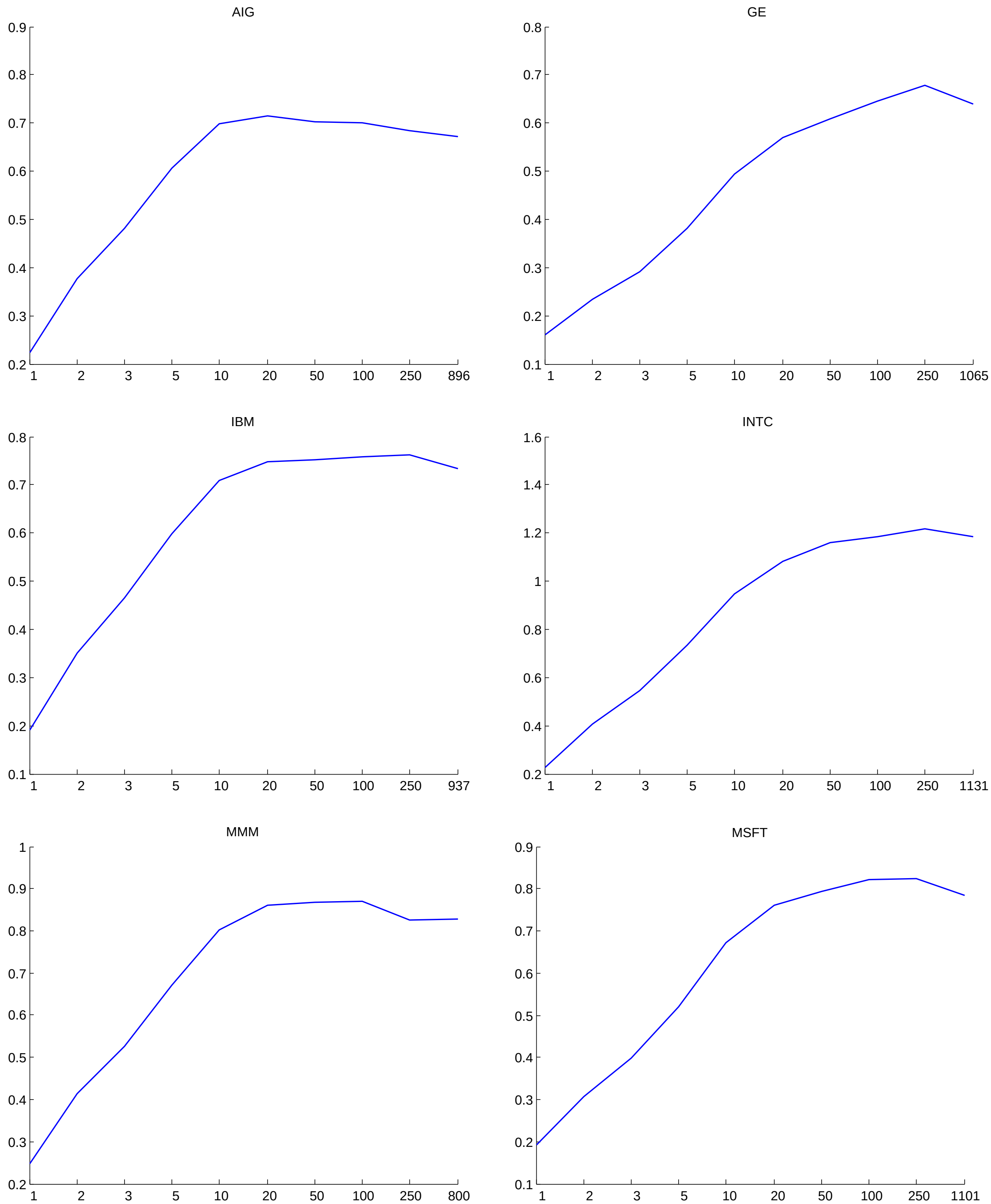

Figure D.3: Average weekly realised betas of individual stocks against the frequency (in ticks) used in their calculation.

E Empirical Analysis: Figures and Tables

	$\widehat{\beta}^{RV}_{5min}$		$\widehat{\beta}^{RV}_{10min}$		$\widehat{\beta}^{RV}_{20min}$		$\widehat{\beta}^{TS}$	
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.1: Values of the joint Chi-square test statistic (see section 4.1); corresponding p-values in parenthesis. The null hypothesis is that the true betas for all 6 stocks are time-invariant over the time interval specified 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.

	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.2: Values of the Chi-square test statistic; corresponding p-values in parenthesis. The null hypothesis is that true betas are constant over the time interval specified in the top row (2016, Q1, Q2, Q3, or Q4).

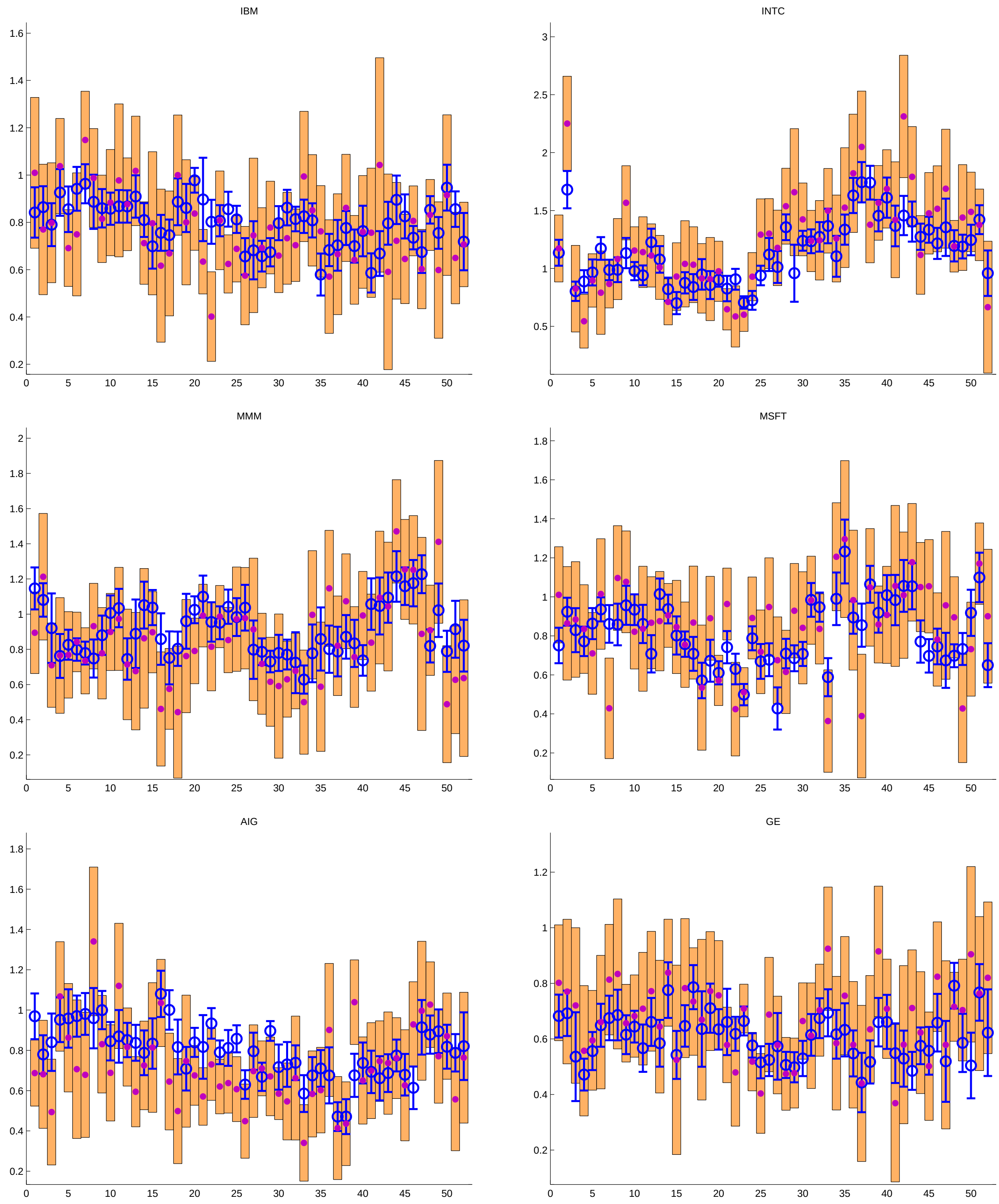

Figure E.3: Estimated betas for AIG, GE, IBM, INTC, MMM, and MSFT with 95% confidence intervals. Filled dots with rectangular CIs correspond to $\widehat{\beta}^{RV}_{10min}$*, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{TS}$*. Weeks on the x-axis.*

M Monte Carlo Simulations: Tables

	$\xi = 0.0000$						$\xi = 0.0001$					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
	$RMSE \times 10$ of $\widehat{\beta}^{TS}$											
$G_1 = 3$	0.10	0.10	0.11	0.06	0.11	0.08	0.17	0.19	0.19	0.13	0.18	0.17
$G_1 = 5$	0.12	0.11	0.13	0.06	0.13	0.09	0.15	0.14	0.16	0.09	0.16	0.12
$G_1 = 10$	0.15	0.14	0.16	0.08	0.17	0.12	0.16	0.15	0.17	0.09	0.18	0.13
$G_1 = 20$	0.21	0.19	0.22	0.11	0.23	0.16	0.21	0.19	0.23	0.11	0.24	0.16
$G_1 = 50$	0.31	0.30	0.34	0.17	0.35	0.25	0.31	0.30	0.34	0.18	0.35	0.25
$G_1 = 70$	0.37	0.35	0.40	0.20	0.41	0.29	0.37	0.35	0.40	0.21	0.41	0.29
	Mean values of G_1 selected by each rule											
G_1^{BR}	1.71	1.71	1.71	1.71	1.71	1.71	5.22	6.21	5.44	6.56	4.69	6.38
$\widehat{G}_1^{BR}$	1.77	1.77	1.77	1.76	1.77	1.77	2.05	2.15	2.06	2.19	1.98	2.18
$G_1^{\star}$	0.00	0.00	0.00	0.00	0.00	0.00	3.17	4.04	3.36	4.36	2.72	4.20
$\widehat{G}_1^{\star}$	0.27	0.27	0.27	0.26	0.26	0.26	0.71	0.84	0.73	0.90	0.62	0.88

Table M.1: The last four rows denote the averages across simulations of G_1^{BR}, $\widehat{G}_1^{BR}$, $G_1^{\star}$, and $\widehat{G}_1^{\star}$ defined in Section 5.2. The top rows show the RMSE times a factor of 10 of the $\widehat{\beta}^{TS}$ 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
	$G_1 = 3,\ \mathbf{V^{FS}} = \mathbf{0.755}\ \ (V^{pl} = 0.363,\ V^{Theo} = 0.357)$								
100	0.703	0.699	0.697	0.697	0.697	0.697	0.693	0.695	0.693
200	0.726	0.721	0.719	0.719	0.720	0.719	0.716	0.719	0.717
300	0.734	0.728	0.726	0.726	0.727	0.727	0.723	0.726	0.723
400	0.738	0.732	0.730	0.730	0.732	0.733	0.728	0.729	0.729
500		0.734	0.732	0.734	0.736	0.736	0.732	0.733	0.735
600		0.737	0.734	0.736	0.739	0.740	0.733	0.736	0.734
800		0.744	0.737	0.739	0.742	0.744	0.738	0.740	0.744
1000			0.739	0.742	0.746	0.745	0.741	0.743	0.748
	$G_1 = 5,\ \mathbf{V^{FS}} = \mathbf{0.983}\ \ (V^{pl} = 0.604,\ V^{Theo} = 0.595)$								
100	0.885	0.881	0.879	0.878	0.877	0.877	0.873	0.873	0.872
200	0.926	0.923	0.919	0.920	0.920	0.920	0.914	0.918	0.916
300	0.941	0.939	0.935	0.934	0.935	0.934	0.930	0.932	0.929
400	0.939	0.944	0.941	0.940	0.942	0.942	0.937	0.939	0.938
500		0.947	0.945	0.943	0.946	0.947	0.942	0.945	0.946
600		0.949	0.946	0.946	0.949	0.951	0.945	0.949	0.947
800		0.953	0.952	0.950	0.953	0.956	0.951	0.955	0.959
1000			0.954	0.954	0.957	0.958	0.956	0.960	0.966
	$G_1 = 10,\ \mathbf{V^{FS}} = \mathbf{1.594}\ \ (V^{pl} = 1.208,\ V^{Theo} = 1.189)$								
100	1.337	1.329	1.323	1.322	1.317	1.318	1.309	1.315	1.310
200	1.448	1.447	1.441	1.439	1.436	1.435	1.432	1.436	1.433
300	1.483	1.489	1.484	1.481	1.477	1.477	1.474	1.479	1.475
400	1.469	1.509	1.503	1.497	1.495	1.496	1.495	1.499	1.498
500		1.516	1.515	1.507	1.505	1.508	1.508	1.514	1.513
600		1.519	1.521	1.513	1.511	1.516	1.515	1.523	1.521
800		1.511	1.528	1.521	1.519	1.525	1.528	1.536	1.538
1000			1.529	1.524	1.526	1.531	1.539	1.546	1.549
	$G_1 = 20,\ \mathbf{V^{FS}} = \mathbf{2.778}\ \ (V^{pl} = 2.415,\ V^{Theo} = 2.379)$								
100	2.060	2.029	2.015	2.007	1.998	1.996	1.988	1.991	1.975
200	2.415	2.405	2.391	2.381	2.373	2.369	2.367	2.364	2.359
300	2.509	2.533	2.520	2.508	2.497	2.496	2.493	2.497	2.483
400	2.445	2.591	2.577	2.563	2.553	2.552	2.548	2.554	2.539
500		2.614	2.611	2.597	2.586	2.586	2.582	2.588	2.580
600		2.627	2.635	2.619	2.608	2.610	2.609	2.616	2.601
800		2.592	2.657	2.639	2.635	2.637	2.642	2.646	2.639
1000			2.660	2.650	2.653	2.654	2.663	2.668	2.662
	$G_1 = 50,\ \mathbf{V^{FS}} = \mathbf{6.236}\ \ (V^{pl} = 6.071,\ V^{Theo} = 5.947)$								
100	2.660	2.474	2.403	2.362	2.332	2.313	2.291	2.282	2.248
200	4.529	4.452	4.399	4.349	4.323	4.307	4.281	4.257	4.244
300	4.965	5.110	5.072	5.023	4.993	4.984	4.951	4.934	4.911
400	4.560	5.415	5.396	5.349	5.315	5.309	5.274	5.265	5.239
500		5.565	5.579	5.536	5.505	5.499	5.469	5.457	5.432
600		5.632	5.699	5.664	5.635	5.634	5.601	5.590	5.564
800		5.437	5.814	5.803	5.790	5.791	5.762	5.748	5.727
1000			5.820	5.869	5.886	5.880	5.859	5.837	5.821

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 $n^{1/3}\big(\widehat{\beta}^{TS} - \beta\big)^2$. Noise-to-signal ratio is $\xi = 0.0000$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$, $\mathbf{V^{FS} = 2.549}$ ($V^{pl} = 0.879$, $V^{Theo} = 0.798$)								
100	2.442	2.429	2.420	2.421	2.416	2.416	2.418	2.420	2.415
200	2.527	2.511	2.495	2.494	2.497	2.493	2.490	2.499	2.500
300	2.572	2.541	2.522	2.518	2.523	2.517	2.513	2.524	2.520
400	2.620	2.551	2.530	2.528	2.533	2.530	2.528	2.530	2.534
500		2.553	2.534	2.533	2.533	2.533	2.527	2.535	2.540
600		2.558	2.537	2.536	2.536	2.539	2.529	2.542	2.535
800		2.601	2.557	2.549	2.548	2.552	2.543	2.551	2.553
1000			2.569	2.557	2.556	2.553	2.550	2.554	2.559
	$G_1 = 5$, $\mathbf{V^{FS} = 1.694}$ ($V^{pl} = 0.883$, $V^{Theo} = 0.754$)								
100	1.504	1.495	1.487	1.489	1.483	1.486	1.480	1.483	1.474
200	1.571	1.556	1.544	1.549	1.548	1.544	1.539	1.541	1.539
300	1.599	1.581	1.567	1.568	1.569	1.563	1.558	1.560	1.556
400	1.619	1.589	1.574	1.575	1.578	1.575	1.570	1.572	1.568
500		1.593	1.577	1.577	1.581	1.579	1.575	1.578	1.575
600		1.596	1.581	1.583	1.585	1.585	1.577	1.582	1.577
800		1.617	1.596	1.592	1.594	1.595	1.588	1.592	1.591
1000			1.606	1.601	1.599	1.597	1.594	1.600	1.598
	$G_1 = 10$, $\mathbf{V^{FS} = 1.867}$ ($V^{pl} = 1.485$, $V^{Theo} = 1.229$)								
100	1.541	1.531	1.523	1.522	1.517	1.517	1.508	1.514	1.508
200	1.665	1.659	1.655	1.653	1.649	1.648	1.644	1.647	1.643
300	1.707	1.708	1.703	1.701	1.696	1.693	1.691	1.694	1.692
400	1.704	1.729	1.724	1.721	1.716	1.715	1.714	1.718	1.716
500		1.739	1.737	1.732	1.728	1.728	1.729	1.734	1.731
600		1.745	1.747	1.742	1.737	1.739	1.737	1.743	1.739
800		1.740	1.757	1.750	1.745	1.748	1.750	1.756	1.753
1000			1.759	1.753	1.751	1.754	1.759	1.763	1.759
	$G_1 = 20$, $\mathbf{V^{FS} = 2.917}$ ($V^{pl} = 2.900$, $V^{Theo} = 2.389$)								
100	2.139	2.106	2.091	2.083	2.075	2.071	2.063	2.066	2.050
200	2.506	2.493	2.477	2.467	2.459	2.452	2.452	2.448	2.443
300	2.603	2.624	2.609	2.596	2.585	2.580	2.579	2.585	2.570
400	2.545	2.684	2.666	2.651	2.640	2.638	2.633	2.640	2.625
500		2.708	2.701	2.685	2.672	2.671	2.667	2.674	2.665
600		2.722	2.725	2.707	2.694	2.695	2.694	2.701	2.686
800		2.688	2.748	2.728	2.721	2.722	2.728	2.730	2.723
1000			2.757	2.740	2.739	2.739	2.749	2.753	2.746
	$G_1 = 50$, $\mathbf{V^{FS} = 6.317}$ ($V^{pl} = 7.263$, $V^{Theo} = 5.949$)								
100	2.680	2.490	2.419	2.378	2.348	2.329	2.307	2.296	2.264
200	4.556	4.475	4.420	4.371	4.345	4.330	4.302	4.279	4.267
300	4.994	5.132	5.095	5.046	5.016	5.007	4.973	4.958	4.935
400	4.593	5.439	5.419	5.372	5.339	5.334	5.296	5.289	5.263
500		5.590	5.602	5.560	5.530	5.523	5.491	5.482	5.456
600		5.659	5.724	5.690	5.662	5.660	5.624	5.616	5.589
800		5.468	5.840	5.830	5.817	5.817	5.787	5.774	5.754
1000			5.848	5.897	5.915	5.907	5.887	5.864	5.849

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 $\xi = 0.0001$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$ (coverage is 0.825 for V^{pl} and 0.822 for V^{Theo})								
100	0.943	0.939	0.939	0.939	0.944	0.938	0.935	0.937	0.938
200	0.950	0.944	0.944	0.947	0.949	0.947	0.947	0.943	0.944
300	0.949	0.948	0.946	0.946	0.948	0.947	0.944	0.940	0.942
400	0.951	0.945	0.947	0.949	0.946	0.944	0.941	0.938	0.936
500		0.944	0.944	0.943	0.942	0.942	0.941	0.936	0.930
600		0.945	0.945	0.941	0.940	0.939	0.936	0.938	0.930
800		0.949	0.947	0.941	0.935	0.933	0.933	0.933	0.927
1000			0.949	0.940	0.935	0.935	0.930	0.933	0.924
	$G_1 = 5$ (coverage is 0.870 for V^{pl} and 0.874 for V^{Theo})								
100	0.938	0.938	0.937	0.937	0.936	0.940	0.935	0.938	0.934
200	0.945	0.943	0.943	0.946	0.946	0.944	0.941	0.938	0.941
300	0.945	0.943	0.939	0.942	0.941	0.942	0.941	0.941	0.938
400	0.943	0.942	0.939	0.941	0.938	0.939	0.937	0.938	0.939
500		0.945	0.944	0.939	0.937	0.935	0.937	0.939	0.937
600		0.945	0.945	0.935	0.937	0.934	0.934	0.935	0.933
800		0.945	0.947	0.935	0.933	0.933	0.931	0.931	0.930
1000			0.945	0.938	0.932	0.933	0.931	0.926	0.926
	$G_1 = 10$ (coverage is 0.907 for V^{pl} and 0.907 for V^{Theo})								
100	0.929	0.929	0.926	0.926	0.924	0.922	0.925	0.927	0.925
200	0.943	0.937	0.935	0.933	0.937	0.934	0.937	0.936	0.935
300	0.943	0.941	0.935	0.938	0.938	0.935	0.939	0.938	0.938
400	0.940	0.943	0.938	0.939	0.937	0.937	0.934	0.939	0.939
500		0.941	0.938	0.936	0.934	0.934	0.934	0.935	0.934
600		0.939	0.935	0.934	0.932	0.937	0.930	0.933	0.936
800		0.943	0.937	0.936	0.929	0.935	0.931	0.932	0.929
1000			0.945	0.934	0.934	0.931	0.923	0.929	0.926
	$G_1 = 20$ (coverage is 0.940 for V^{pl} and 0.938 for V^{Theo})								
100	0.916	0.915	0.912	0.911	0.909	0.905	0.908	0.906	0.899
200	0.942	0.937	0.938	0.938	0.936	0.936	0.938	0.938	0.934
300	0.944	0.941	0.945	0.941	0.942	0.941	0.942	0.938	0.937
400	0.940	0.939	0.942	0.941	0.944	0.940	0.939	0.943	0.939
500		0.944	0.942	0.944	0.943	0.940	0.941	0.941	0.940
600		0.946	0.943	0.946	0.941	0.942	0.940	0.942	0.936
800		0.949	0.946	0.939	0.937	0.938	0.939	0.945	0.938
1000			0.946	0.941	0.937	0.935	0.938	0.942	0.937
	$G_1 = 50$ (coverage is 0.947 for V^{pl} and 0.945 for V^{Theo})								
100	0.798	0.780	0.771	0.760	0.765	0.763	0.756	0.759	0.750
200	0.900	0.897	0.895	0.893	0.891	0.893	0.892	0.891	0.887
300	0.914	0.915	0.913	0.910	0.909	0.913	0.912	0.914	0.912
400	0.909	0.924	0.922	0.919	0.919	0.923	0.917	0.920	0.917
500		0.927	0.924	0.921	0.925	0.928	0.926	0.920	0.921
600		0.929	0.928	0.926	0.932	0.930	0.929	0.924	0.921
800		0.929	0.933	0.932	0.931	0.930	0.926	0.927	0.921
1000			0.937	0.935	0.931	0.930	0.921	0.922	0.919

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 $\xi = 0.0000$.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$ (coverage is 0.754 for V^{pl} and 0.740 for V^{Theo})								
100	0.949	0.947	0.950	0.946	0.947	0.948	0.946	0.948	0.945
200	0.952	0.951	0.946	0.946	0.943	0.946	0.944	0.944	0.943
300	0.951	0.953	0.948	0.944	0.945	0.944	0.945	0.944	0.943
400	0.956	0.951	0.949	0.943	0.943	0.943	0.943	0.944	0.942
500		0.950	0.948	0.942	0.939	0.936	0.938	0.935	0.932
600		0.948	0.947	0.940	0.941	0.936	0.933	0.936	0.934
800		0.952	0.942	0.941	0.937	0.937	0.931	0.931	0.930
1000			0.949	0.942	0.937	0.933	0.926	0.929	0.931
	$G_1 = 5$ (coverage is 0.846 for V^{pl} and 0.818 for V^{Theo})								
100	0.933	0.933	0.929	0.932	0.932	0.931	0.934	0.926	0.927
200	0.938	0.937	0.936	0.935	0.940	0.933	0.934	0.934	0.935
300	0.936	0.939	0.935	0.938	0.943	0.937	0.936	0.933	0.932
400	0.937	0.941	0.934	0.935	0.940	0.934	0.931	0.934	0.931
500		0.941	0.933	0.936	0.938	0.932	0.926	0.928	0.929
600		0.938	0.934	0.935	0.934	0.933	0.927	0.928	0.927
800		0.941	0.936	0.936	0.930	0.931	0.924	0.929	0.925
1000			0.944	0.941	0.932	0.927	0.922	0.922	0.925
	$G_1 = 10$ (coverage is 0.917 for V^{pl} and 0.882 for V^{Theo})								
100	0.919	0.917	0.916	0.920	0.915	0.916	0.918	0.914	0.917
200	0.931	0.924	0.927	0.927	0.925	0.926	0.922	0.926	0.927
300	0.938	0.929	0.932	0.929	0.926	0.925	0.928	0.929	0.923
400	0.933	0.931	0.929	0.930	0.927	0.928	0.924	0.926	0.922
500		0.931	0.929	0.929	0.928	0.925	0.925	0.923	0.923
600		0.932	0.932	0.926	0.927	0.926	0.926	0.925	0.921
800		0.934	0.937	0.928	0.926	0.926	0.925	0.927	0.920
1000			0.934	0.932	0.927	0.923	0.925	0.924	0.919
	$G_1 = 20$ (coverage is 0.959 for V^{pl} and 0.937 for V^{Theo})								
100	0.913	0.903	0.904	0.903	0.899	0.899	0.902	0.897	0.903
200	0.937	0.931	0.931	0.939	0.934	0.935	0.934	0.933	0.932
300	0.945	0.933	0.940	0.941	0.940	0.941	0.938	0.937	0.936
400	0.942	0.938	0.942	0.939	0.938	0.943	0.937	0.939	0.936
500		0.940	0.946	0.943	0.937	0.936	0.936	0.939	0.936
600		0.945	0.945	0.940	0.935	0.936	0.938	0.937	0.933
800		0.950	0.945	0.936	0.933	0.935	0.938	0.937	0.939
1000			0.947	0.940	0.937	0.936	0.941	0.942	0.937
	$G_1 = 50$ (coverage is 0.968 for V^{pl} and 0.943 for V^{Theo})								
100	0.798	0.784	0.776	0.768	0.771	0.764	0.757	0.761	0.747
200	0.893	0.891	0.890	0.888	0.887	0.888	0.886	0.888	0.889
300	0.912	0.911	0.910	0.908	0.904	0.911	0.913	0.912	0.911
400	0.907	0.921	0.923	0.919	0.918	0.921	0.921	0.921	0.919
500		0.926	0.925	0.923	0.920	0.924	0.922	0.925	0.920
600		0.931	0.926	0.928	0.929	0.929	0.928	0.927	0.918
800		0.925	0.929	0.932	0.931	0.926	0.926	0.920	0.921
1000			0.936	0.932	0.927	0.923	0.920	0.916	0.914

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 $\xi = 0.0001$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	(1)				(2)			
$G_1 = 3$	1.005	0.493	0.485	1.031	1.379	0.700	0.692	1.434
$G_1 = 5$	1.294	0.822	0.808	1.340	1.782	1.166	1.154	1.886
$G_1 = 10$	2.054	1.643	1.615	2.143	2.864	2.332	2.307	3.050
$G_1 = 20$	3.513	3.286	3.231	3.668	5.003	4.661	4.614	5.521
$G_1 = 50$	7.450	8.273	8.076	8.335	10.717	11.675	11.535	12.180
	(3)				(4)			
$G_1 = 3$	0.736	0.363	0.357	0.755	1.213	0.584	0.572	1.207
$G_1 = 5$	0.947	0.604	0.595	0.983	1.563	0.973	0.953	1.569
$G_1 = 10$	1.508	1.208	1.189	1.594	2.463	1.945	1.907	2.620
$G_1 = 20$	2.586	2.415	2.379	2.778	4.165	3.893	3.814	4.435
$G_1 = 50$	5.499	6.071	5.947	6.236	8.782	9.842	9.534	10.230
	(5)				(6)			
$G_1 = 3$	1.213	0.584	0.572	1.207	1.085	0.539	0.529	1.079
$G_1 = 5$	1.563	0.973	0.953	1.569	1.406	0.898	0.881	1.425
$G_1 = 10$	2.463	1.945	1.907	2.620	2.237	1.796	1.762	2.373
$G_1 = 20$	4.165	3.893	3.814	4.435	3.823	3.590	3.524	4.071
$G_1 = 50$	8.782	9.842	9.534	10.230	8.131	9.044	8.810	9.154

Table M.6: The average over simulations of the following measures of dispersion of $\hat{\beta}^{TS}$ (times 10^2): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asymptotic variance V^{Theo}, 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 5. $J = 500$, $m = 3000$, and $\xi = 0.0000$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	(1)				(2)			
$G_1 = 3$	3.118	1.100	0.988	3.410	8.029	2.738	2.581	8.429
$G_1 = 5$	2.061	1.167	0.989	2.279	3.969	2.084	1.834	4.399
$G_1 = 10$	2.326	2.016	1.661	2.472	3.579	2.977	2.477	4.004
$G_1 = 20$	3.619	3.952	3.242	3.822	5.267	5.654	4.657	5.906
$G_1 = 50$	7.485	9.930	8.078	8.412	10.791	14.075	11.542	12.380
	(3)				(4)			
$G_1 = 3$	2.533	0.879	0.798	2.549	2.997	1.080	0.953	3.186
$G_1 = 5$	1.579	0.883	0.754	1.694	2.247	1.298	1.090	2.376
$G_1 = 10$	1.728	1.485	1.229	1.867	2.711	2.356	1.941	2.933
$G_1 = 20$	2.671	2.900	2.389	2.917	4.270	4.657	3.822	4.596
$G_1 = 50$	5.523	7.263	5.949	6.317	8.814	11.753	9.536	10.290
	(5)				(6)			
$G_1 = 3$	2.997	1.080	0.953	3.186	3.209	1.143	1.022	3.346
$G_1 = 5$	2.247	1.298	1.090	2.376	2.178	1.252	1.058	2.387
$G_1 = 10$	2.711	2.356	1.941	2.933	2.507	2.192	1.806	2.746
$G_1 = 20$	4.270	4.657	3.822	4.596	3.931	4.306	3.535	4.227
$G_1 = 50$	8.814	11.753	9.536	10.290	8.167	10.822	8.812	9.225

Table M.7: The average over simulations of the following measures of dispersion of $\hat{\beta}^{TS}$ (times 10^2): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asymptotic variance V^{Theo}, 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 5. $J = 500$, $m = 3000$, and $\xi = 0.0001$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
		(1)			(2)	
$G_1 = 3$	0.938	0.821	0.824	0.938	0.834	0.820
$G_1 = 5$	0.939	0.865	0.874	0.941	0.877	0.874
$G_1 = 10$	0.937	0.908	0.909	0.939	0.910	0.918
$G_1 = 20$	0.939	0.942	0.940	0.930	0.932	0.930
$G_1 = 50$	0.922	0.944	0.943	0.928	0.948	0.954
		(3)			(4)	
$G_1 = 3$	0.942	0.825	0.822	0.941	0.821	0.819
$G_1 = 5$	0.935	0.870	0.874	0.933	0.872	0.870
$G_1 = 10$	0.934	0.907	0.907	0.934	0.913	0.912
$G_1 = 20$	0.940	0.940	0.938	0.942	0.941	0.937
$G_1 = 50$	0.928	0.947	0.945	0.914	0.941	0.938
		(5)			(6)	
$G_1 = 3$	0.941	0.821	0.819	0.942	0.832	0.827
$G_1 = 5$	0.933	0.872	0.870	0.938	0.883	0.876
$G_1 = 10$	0.934	0.913	0.912	0.940	0.919	0.914
$G_1 = 20$	0.942	0.941	0.937	0.940	0.940	0.940
$G_1 = 50$	0.914	0.941	0.938	0.927	0.943	0.943

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) asymptotic variance V^{Theo}. Scenarios (1)-(6) are described in Section 5. $J = 500$, $m = 3000$, and $\xi = 0.0000$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
	(1)			(2)		
$G_1 = 3$	0.928	0.746	0.720	0.952	0.751	0.734
$G_1 = 5$	0.924	0.842	0.815	0.937	0.817	0.792
$G_1 = 10$	0.940	0.925	0.895	0.940	0.907	0.885
$G_1 = 20$	0.943	0.955	0.931	0.934	0.948	0.924
$G_1 = 50$	0.920	0.961	0.939	0.928	0.965	0.950
	(3)			(4)		
$G_1 = 3$	0.936	0.754	0.740	0.925	0.755	0.724
$G_1 = 5$	0.932	0.846	0.818	0.924	0.844	0.812
$G_1 = 10$	0.925	0.917	0.882	0.935	0.915	0.892
$G_1 = 20$	0.936	0.959	0.937	0.943	0.956	0.941
$G_1 = 50$	0.924	0.968	0.943	0.914	0.962	0.943
	(5)			(6)		
$G_1 = 3$	0.925	0.755	0.724	0.939	0.752	0.729
$G_1 = 5$	0.924	0.844	0.812	0.923	0.844	0.820
$G_1 = 10$	0.935	0.915	0.892	0.934	0.923	0.894
$G_1 = 20$	0.943	0.956	0.941	0.934	0.952	0.935
$G_1 = 50$	0.914	0.962	0.943	0.924	0.959	0.942

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) asymptotic variance V^{Theo}. Scenarios (1)-(6) are described in Section 5. $J = 500$, $m = 3000$, and $\xi = 0.0001$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.076	0.035	0.034	0.068	0.229	0.104	0.103	0.210
$G_1 = 5$	0.097	0.059	0.057	0.091	0.294	0.174	0.172	0.269
$G_1 = 10$	0.152	0.117	0.114	0.148	0.460	0.348	0.345	0.440
$G_1 = 20$	0.259	0.234	0.228	0.263	0.789	0.696	0.689	0.811
$G_1 = 50$	0.573	0.585	0.569	0.631	1.737	1.739	1.723	1.809
	(3)				(4)			
$G_1 = 3$	0.044	0.020	0.020	0.038	0.117	0.046	0.045	0.093
$G_1 = 5$	0.056	0.034	0.033	0.052	0.145	0.077	0.075	0.123
$G_1 = 10$	0.087	0.067	0.066	0.085	0.216	0.155	0.149	0.199
$G_1 = 20$	0.149	0.134	0.131	0.151	0.354	0.309	0.298	0.346
$G_1 = 50$	0.329	0.336	0.328	0.358	0.758	0.773	0.745	0.820
	(5)				(6)			
$G_1 = 3$	0.117	0.046	0.045	0.093	0.105	0.043	0.043	0.081
$G_1 = 5$	0.145	0.077	0.075	0.123	0.131	0.072	0.071	0.110
$G_1 = 10$	0.216	0.155	0.149	0.199	0.199	0.145	0.142	0.188
$G_1 = 20$	0.354	0.309	0.298	0.346	0.333	0.289	0.285	0.326
$G_1 = 50$	0.758	0.773	0.745	0.820	0.722	0.724	0.712	0.773

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) asymptotic variance V^{Theo}, 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 5. $J = 500$, $m = 3000$, and $\xi = 0.0000$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	(1)				(2)			
$G_1 = 3$	0.176	0.061	0.054	0.166	0.850	0.279	0.262	0.813
$G_1 = 5$	0.136	0.075	0.064	0.134	0.515	0.258	0.229	0.516
$G_1 = 10$	0.167	0.137	0.116	0.164	0.540	0.416	0.359	0.524
$G_1 = 20$	0.265	0.271	0.228	0.274	0.821	0.808	0.693	0.869
$G_1 = 50$	0.575	0.676	0.569	0.633	1.746	2.011	1.724	1.835
	(3)				(4)			
$G_1 = 3$	0.109	0.037	0.034	0.100	0.211	0.070	0.061	0.197
$G_1 = 5$	0.080	0.044	0.038	0.073	0.182	0.095	0.081	0.161
$G_1 = 10$	0.097	0.079	0.067	0.095	0.231	0.179	0.151	0.221
$G_1 = 20$	0.153	0.155	0.131	0.158	0.360	0.354	0.299	0.357
$G_1 = 50$	0.330	0.388	0.328	0.361	0.759	0.885	0.745	0.825
	(5)				(6)			
$G_1 = 3$	0.211	0.070	0.061	0.197	0.217	0.072	0.065	0.197
$G_1 = 5$	0.182	0.095	0.081	0.161	0.174	0.092	0.079	0.158
$G_1 = 10$	0.231	0.179	0.151	0.221	0.216	0.169	0.144	0.206
$G_1 = 20$	0.360	0.354	0.299	0.357	0.340	0.335	0.285	0.339
$G_1 = 50$	0.759	0.885	0.745	0.825	0.724	0.835	0.712	0.777

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) asymptotic variance V^{Theo}, 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 5. $J = 500$, $m = 3000$, and $\xi = 0.0001$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
		(1)			(2)	
$G_1 = 3$	0.956	0.829	0.830	0.955	0.834	0.838
$G_1 = 5$	0.950	0.880	0.886	0.955	0.877	0.879
$G_1 = 10$	0.950	0.914	0.924	0.944	0.909	0.915
$G_1 = 20$	0.941	0.933	0.939	0.941	0.933	0.934
$G_1 = 50$	0.921	0.932	0.942	0.928	0.936	0.939
		(3)			(4)	
$G_1 = 3$	0.956	0.847	0.842	0.963	0.827	0.832
$G_1 = 5$	0.955	0.875	0.878	0.958	0.875	0.871
$G_1 = 10$	0.938	0.914	0.909	0.943	0.902	0.911
$G_1 = 20$	0.942	0.940	0.937	0.928	0.923	0.926
$G_1 = 50$	0.923	0.930	0.944	0.916	0.925	0.935
		(5)			(6)	
$G_1 = 3$	0.963	0.827	0.832	0.964	0.840	0.839
$G_1 = 5$	0.958	0.875	0.871	0.955	0.887	0.882
$G_1 = 10$	0.943	0.902	0.911	0.950	0.911	0.916
$G_1 = 20$	0.928	0.923	0.926	0.940	0.930	0.933
$G_1 = 50$	0.916	0.925	0.935	0.917	0.928	0.938

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) asymptotic variance V^{Theo}. Scenarios (1)-(6) are described in Section 5. $J = 500$, $m = 3000$, and $\xi = 0.0000$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
		(1)			(2)	
$G_1 = 3$	0.957	0.763	0.733	0.951	0.736	0.728
$G_1 = 5$	0.944	0.847	0.807	0.952	0.840	0.824
$G_1 = 10$	0.945	0.927	0.902	0.952	0.914	0.899
$G_1 = 20$	0.946	0.945	0.938	0.940	0.933	0.923
$G_1 = 50$	0.922	0.954	0.945	0.923	0.952	0.937
		(3)			(4)	
$G_1 = 3$	0.954	0.760	0.752	0.956	0.767	0.730
$G_1 = 5$	0.950	0.871	0.846	0.948	0.859	0.838
$G_1 = 10$	0.945	0.931	0.907	0.932	0.916	0.898
$G_1 = 20$	0.938	0.951	0.930	0.929	0.942	0.924
$G_1 = 50$	0.921	0.953	0.943	0.913	0.948	0.931
		(5)			(6)	
$G_1 = 3$	0.956	0.767	0.730	0.957	0.754	0.746
$G_1 = 5$	0.948	0.859	0.838	0.946	0.845	0.822
$G_1 = 10$	0.932	0.916	0.898	0.946	0.924	0.906
$G_1 = 20$	0.929	0.942	0.924	0.942	0.948	0.934
$G_1 = 50$	0.913	0.948	0.931	0.919	0.951	0.941

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) asymptotic variance V^{Theo}. Scenarios (1)-(6) are described in Section 5. $J = 500$, $m = 3000$, and $\xi = 0.0001$.

R Monte Carlo Simulations with a Two-Factor Stochastic Volatility Model

As a robustness check, we now document the finite sample properties of the subsampling procedure in a two-factor stochastic volatility model.
Both the efficient log-price of the stock $X^{(1)}$ and the market $X^{(2)}$ follow the two-factor stochastic volatility model of Bollerslev and Todorov (2011) without jumps. For $i = 1, 2$,

$$\begin{aligned}
dX_t^{(i)} &= \sigma_t^{(i)} dW_t^{(i)} \\
\left(\sigma_t^{(i)}\right)^2 &= V_{1,t}^{(i)} + V_{2,t}^{(i)} \\
dV_{1,t}^{(i)} &= 0.0128(0.4068 - V_{1,t}^{(i)})dt + 0.0954\sqrt{V_{1,t}^{(i)}} \left(\rho dW_t^{(i)} + \sqrt{1-\rho^2} dB_{1,t}^{(i)}\right) \\
dV_{2,t}^{(i)} &= 0.6930(0.4068 - V_{2,t}^{(i)})dt + 0.7023\sqrt{V_{2,t}^{(i)}} \left(\rho dW_t^{(i)} + \sqrt{1-\rho^2} dB_{2,t}^{(i)}\right),
\end{aligned}$$

where $\rho = -0.7$. In addition, we set $Corr(dW_t^{(1)}, dW_t^{(2)}) = 0.5$, so that the correlation between the spot returns of $X^{(1)}$ and $X^{(2)}$ (and the spot beta) equals 0.5. We generate the market microstructure noise as before, see Section 5 for details. We generate asynchronous observations as in Section 5 with n_1 and n_2 from MSFT and SPY.
From Tables R.1 and R.2, we find that the values of subsampled variances and coverage probabilities 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. These ranges include the values $J = 500$ and $m = 3000$, which were chosen based on the simulations in Section 5. For brevity, we have only included the designs with the market microstructure, but the zero market microstructure noise results lend themselves to the same conclusion. Further numerical results in Tables R.3-R.6 use $J = 500$ and $m = 3000$.
The results in Tables R.1-R.6 indicate that 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 $\widehat{V}^{pl}$ estimates the theoretical asymptotic variance V^{Theo} well, the asymptotic variance itself is not very close to the finite sample variability V^{FS} for smaller 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.
Hence, this simulation exercises suggests some degree of robustness of the conclusions of Section 5 to model misspecification.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$, $\mathbf{V^{FS}} = \mathbf{1.646}$ $(V^{pl} = 0.568,\ V^{Theo} = 0.509)$								
100	1.570	1.558	1.553	1.555	1.552	1.552	1.558	1.553	1.550
200	1.627	1.614	1.605	1.603	1.598	1.598	1.596	1.602	1.601
300	1.653	1.628	1.620	1.613	1.610	1.607	1.611	1.617	1.613
400	1.681	1.638	1.626	1.620	1.614	1.611	1.618	1.624	1.623
500		1.643	1.628	1.623	1.614	1.613	1.620	1.629	1.628
600		1.650	1.632	1.625	1.613	1.611	1.625	1.630	1.629
800		1.666	1.642	1.627	1.612	1.611	1.627	1.634	1.633
1000			1.652	1.627	1.613	1.614	1.630	1.640	1.638
	$G_1 = 5$, $\mathbf{V^{FS}} = \mathbf{1.128}$ $(V^{pl} = 0.580,\ V^{Theo} = 0.488)$								
100	0.971	0.965	0.960	0.960	0.959	0.959	0.957	0.959	0.955
200	1.014	1.008	1.004	1.002	1.000	0.997	0.994	0.998	0.994
300	1.034	1.025	1.017	1.014	1.010	1.006	1.004	1.006	1.001
400	1.051	1.031	1.023	1.017	1.012	1.007	1.005	1.007	1.004
500		1.034	1.024	1.019	1.013	1.009	1.005	1.009	1.007
600		1.036	1.026	1.020	1.013	1.008	1.004	1.008	1.005
800		1.043	1.032	1.021	1.012	1.007	1.003	1.005	1.003
1000			1.036	1.021	1.011	1.004	1.001	1.001	1.001
	$G_1 = 10$, $\mathbf{V^{FS}} = \mathbf{1.204}$ $(V^{pl} = 0.984,\ V^{Theo} = 0.802)$								
100	1.006	0.997	0.992	0.992	0.990	0.987	0.988	0.988	0.985
200	1.088	1.082	1.078	1.077	1.073	1.071	1.071	1.073	1.070
300	1.111	1.110	1.105	1.103	1.098	1.095	1.096	1.098	1.094
400	1.111	1.123	1.119	1.115	1.110	1.108	1.110	1.111	1.107
500		1.130	1.127	1.123	1.117	1.115	1.117	1.119	1.116
600		1.135	1.132	1.128	1.122	1.118	1.121	1.124	1.117
800		1.129	1.139	1.132	1.125	1.121	1.127	1.126	1.122
1000			1.142	1.136	1.126	1.125	1.129	1.129	1.125
	$G_1 = 20$, $\mathbf{V^{FS}} = \mathbf{1.812}$ $(V^{pl} = 1.929,\ V^{Theo} = 1.560)$								
100	1.397	1.373	1.365	1.359	1.358	1.351	1.356	1.354	1.349
200	1.632	1.623	1.615	1.615	1.610	1.607	1.602	1.606	1.601
300	1.687	1.706	1.698	1.700	1.693	1.691	1.688	1.684	1.681
400	1.657	1.744	1.741	1.743	1.735	1.733	1.731	1.730	1.728
500		1.764	1.763	1.766	1.759	1.753	1.757	1.756	1.753
600		1.773	1.780	1.783	1.777	1.770	1.774	1.768	1.762
800		1.745	1.801	1.802	1.798	1.792	1.796	1.791	1.785
1000			1.808	1.812	1.807	1.808	1.809	1.799	1.795
	$G_1 = 50$, $\mathbf{V^{FS}} = \mathbf{3.838}$ $(V^{pl} = 4.835,\ V^{Theo} = 3.886)$								
100	1.759	1.634	1.588	1.562	1.545	1.536	1.518	1.522	1.512
200	2.996	2.948	2.914	2.892	2.873	2.865	2.849	2.836	2.830
300	3.274	3.382	3.365	3.349	3.334	3.325	3.307	3.295	3.286
400	3.002	3.587	3.586	3.578	3.568	3.555	3.541	3.535	3.521
500		3.692	3.711	3.713	3.707	3.691	3.678	3.665	3.651
600		3.740	3.792	3.806	3.801	3.786	3.772	3.752	3.741
800		3.587	3.875	3.911	3.911	3.903	3.889	3.868	3.848
1000			3.887	3.966	3.974	3.972	3.955	3.931	3.908

Table R.1: 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 $\xi = 0.0001$. The two-factor stochastic volatility model.

$J \setminus m$	500	1000	1500	2000	2500	3000	4000	5000	6000
	$G_1 = 3$ (coverage is 0.747 for V^{pl} and 0.727 for V^{Theo})								
100	0.939	0.939	0.935	0.934	0.934	0.932	0.934	0.935	0.933
200	0.945	0.942	0.936	0.940	0.937	0.939	0.937	0.937	0.934
300	0.941	0.942	0.937	0.940	0.935	0.937	0.942	0.943	0.940
400	0.945	0.944	0.938	0.942	0.939	0.938	0.940	0.938	0.936
500		0.942	0.936	0.944	0.939	0.935	0.938	0.935	0.940
600		0.942	0.940	0.942	0.935	0.937	0.933	0.937	0.938
800		0.942	0.942	0.942	0.933	0.934	0.932	0.935	0.932
1000			0.946	0.941	0.936	0.935	0.929	0.930	0.934
	$G_1 = 5$ (coverage is 0.839 for V^{pl} and 0.811 for V^{Theo})								
100	0.933	0.931	0.928	0.931	0.930	0.931	0.928	0.929	0.932
200	0.936	0.939	0.938	0.937	0.933	0.932	0.931	0.933	0.932
300	0.937	0.936	0.936	0.936	0.931	0.933	0.928	0.931	0.931
400	0.935	0.937	0.936	0.935	0.930	0.929	0.930	0.926	0.928
500		0.937	0.934	0.931	0.926	0.927	0.925	0.925	0.923
600		0.940	0.934	0.929	0.924	0.925	0.924	0.925	0.927
800		0.941	0.936	0.926	0.916	0.918	0.921	0.920	0.919
1000			0.937	0.924	0.919	0.917	0.917	0.912	0.918
	$G_1 = 10$ (coverage is 0.910 for V^{pl} and 0.888 for V^{Theo})								
100	0.922	0.917	0.917	0.918	0.916	0.915	0.917	0.916	0.913
200	0.931	0.928	0.925	0.923	0.923	0.922	0.922	0.919	0.921
300	0.932	0.933	0.924	0.927	0.922	0.924	0.920	0.924	0.923
400	0.930	0.932	0.921	0.926	0.922	0.921	0.921	0.922	0.924
500		0.931	0.924	0.925	0.922	0.925	0.925	0.922	0.923
600		0.930	0.925	0.926	0.924	0.924	0.924	0.924	0.921
800		0.932	0.929	0.927	0.922	0.922	0.922	0.922	0.920
1000			0.930	0.928	0.926	0.921	0.920	0.919	0.919
	$G_1 = 20$ (coverage is 0.949 for V^{pl} and 0.924 for V^{Theo})								
100	0.914	0.906	0.907	0.908	0.904	0.902	0.907	0.903	0.903
200	0.929	0.931	0.929	0.932	0.932	0.932	0.931	0.930	0.928
300	0.935	0.935	0.934	0.935	0.933	0.934	0.930	0.932	0.929
400	0.935	0.936	0.934	0.933	0.930	0.931	0.928	0.931	0.928
500		0.940	0.938	0.933	0.931	0.931	0.927	0.928	0.927
600		0.942	0.940	0.932	0.928	0.928	0.925	0.923	0.927
800		0.943	0.940	0.931	0.927	0.928	0.923	0.921	0.927
1000			0.941	0.930	0.926	0.928	0.923	0.920	0.921
	$G_1 = 50$ (coverage is 0.974 for V^{pl} and 0.955 for V^{Theo})								
100	0.818	0.796	0.786	0.784	0.778	0.784	0.779	0.778	0.779
200	0.914	0.902	0.906	0.903	0.904	0.909	0.904	0.904	0.905
300	0.927	0.932	0.929	0.929	0.927	0.927	0.923	0.930	0.932
400	0.919	0.937	0.938	0.934	0.936	0.933	0.936	0.941	0.942
500		0.940	0.942	0.943	0.939	0.938	0.941	0.943	0.942
600		0.939	0.943	0.944	0.939	0.940	0.940	0.942	0.942
800		0.937	0.947	0.944	0.937	0.941	0.937	0.938	0.940
1000			0.945	0.942	0.941	0.940	0.936	0.932	0.932

Table R.2: 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 $\xi = 0.0001$. The two-factor stochastic volatility model.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	$\xi = 0.0000$				$\xi = 0.0001$			
$G_1 = 3$	0.483	0.238	0.233	0.527	1.613	0.568	0.509	1.646
$G_1 = 5$	0.617	0.396	0.389	0.669	1.009	0.580	0.488	1.128
$G_1 = 10$	0.978	0.791	0.777	0.997	1.115	0.984	0.802	1.204
$G_1 = 20$	1.699	1.586	1.554	1.715	1.753	1.929	1.560	1.812
$G_1 = 50$	3.678	3.991	3.885	3.801	3.691	4.835	3.886	3.838

Table R.3: The average over simulations of the following measures of dispersion of $\hat{\beta}^{TS}$ (times 10^2): the subsampling estimator $\widehat{V}^{sub}$, the plug-in estimated value of the asymptotic variance $\widehat{V}^{pl}$, the (unobserved) asymptotic variance V^{Theo}, and an average over simulations of $n^{1/3}\big(\widehat{\beta}^{TS} - \beta\big)^2$ denoted by V^{FS}. The two-factor stochastic volatility model. $J = 500$ and $m = 3000$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
	$\xi = 0.0000$			$\xi = 0.0001$		
$G_1 = 3$	0.934	0.794	0.799	0.935	0.747	0.727
$G_1 = 5$	0.938	0.854	0.857	0.927	0.839	0.811
$G_1 = 10$	0.930	0.915	0.911	0.925	0.910	0.888
$G_1 = 20$	0.933	0.934	0.935	0.931	0.949	0.924
$G_1 = 50$	0.940	0.959	0.959	0.938	0.974	0.955

Table R.4: 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) asymptotic variance V^{Theo}. The two-factor stochastic volatility model. $J = 500$ and $m = 3000$.

	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}	$\widehat{V}^{sub}$	$\widehat{V}^{pl}$	V^{Theo}	V^{FS}
	$\xi = 0.0000$				$\xi = 0.0001$			
$G_1 = 3$	35.479	16.290	15.777	35.180	81.857	28.181	25.108	81.230
$G_1 = 5$	45.343	27.150	26.295	46.200	63.201	34.793	29.654	63.040
$G_1 = 10$	70.567	54.299	52.591	70.280	77.275	63.717	53.430	76.010
$G_1 = 20$	121.959	108.599	105.181	116.700	124.780	125.967	105.391	118.800
$G_1 = 50$	269.536	271.497	262.953	263.200	270.633	314.427	262.986	262.700

Table R.5: 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) asymptotic variance V^{Theo}, 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}. The two-factor stochastic volatility model. $J = 500$ and $m = 3000$.

	t^{sub}	t^{pl}	t^{Theo}	t^{sub}	t^{pl}	t^{Theo}
	$\xi = 0.0000$			$\xi = 0.0001$		
$G_1 = 3$	0.942	0.811	0.814	0.945	0.753	0.725
$G_1 = 5$	0.941	0.862	0.860	0.938	0.849	0.826
$G_1 = 10$	0.946	0.916	0.916	0.951	0.930	0.904
$G_1 = 20$	0.943	0.940	0.938	0.942	0.956	0.931
$G_1 = 50$	0.935	0.954	0.954	0.938	0.970	0.955

Table R.6: 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) asymptotic variance V^{Theo}. The two-factor stochastic volatility model. $J = 500$ and $m = 3000$.