Nonparametric Tests of Time Variation in Betas*

Ilze Kalnina†
Université de Montréal

December 31, 2011

Abstract

This paper uses existing and new methods in high frequency financial econometrics to propose new tools to inform the study of time variation in equity betas. We first develop nonparametric tests for time variation of betas. We then build on these tests to obtain a method of nonparameric search for breaks in betas. These methods do not impose any parametric structure across time, they fully account for measurement error in betas, and are robust to market microstructure noise inherent in tick time data. In order to obtain robustness to different specifications of the market microstructure noise, we use the subsampling method of Kalnina (2011) in the construction of the tests. Our context requires a multivariate method, so we extend her univariate subsampling method to the multivariate case. We also show its robustness to long memory in volatility, which is a well documented stylized fact. We implement our tests with a small number of stocks on the NYSE over the year 2006. We find that the use of high frequency data allows to detect significant variation of betas over shorter intervals of time than the estimators relying on data at moderate frequencies, such as 5, 15, or 20 minutes. We strongly reject the hypothesis that beta is constant across quarters in year 2006 for all stocks considered. More power is needed to identify the time of the break, and we find that moderate frequency data does not have enough power to identify a single break in betas in 2006 for most of the stocks, while highest frequency data reveals breaks in most weeks in 2006 for each stock considered.

*I would like to thank Kirill Evdokimov, Vasyl Golosnoy, Silvia Gonçalves, Christian Gourieroux, Benoit Perron, Myung Hwan Seo, and Kevin Sheppard for helpful comments, as well as the participants of the 2011 Canadian Economics Association Meetings, 2011 North American and European summer meetings of the Econometric Society, Second International Panel Data Conference, and the 2011 Meeting of the Canadian Econometrics Study Group.

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

1 Introduction

This paper proposes nonparametric tests for time variation in betas and search for breaks in betas with high frequency data. These methods do not impose any parametric structure across time, they fully account for measurement error in betas, and are robust to market microstructure noise inherent in tick time data. In order to gain robustness to general specifications of the market microstructure noise, we use inference procedure along the lines of Kalnina (2011). We extend her univariate procedure to the multivariate case and demonstrate its robustness to long memory in the spot volatility of asset prices. An empirical study of a small number of stocks from NYSE confirms this to be a very powerful method in detecting the timing of the breaks in betas.

While failure of the unconditional, constant beta Capital Asset Pricing Model of Lintner (1965) and Sharpe (1964) is widely accepted, there is no consensus on the nature of time variation in betas. A traditional way is to estimate betas on rolling windows of say 5 years of monthly data, see, e.g. Fama and MacBeth (1973) and Fama and French (1992). Braun, Nelson, and Sunier (1995) have modeled betas with leverage effects according to an EGARCH model, finding very little leverage effects or even time variation. In Bekaert and Wu (2000), beta dynamics is driven by an asymmetric BEKK specification of the variance-covariance matrices. Meanwhile, access to high-frequency data has driven development of nonparametric volatility estimators allowing also more flexible and precise estimation of betas, see e.g., Barndorff-Nielsen and Shepard (2004). The resulting estimators are often called realized measures, such as realized volatility and realized beta, and are justified in stochastic process theory with an asymptotic scheme where observations are sampled more and more frequently. Several papers have investigated dynamics of betas estimated from high frequency data, for example, Patton and Verardo (2011) investigate effect of news on daily realized betas of individual stocks, Bollerslev and Zhang (2003) model dynamics of monthly realized betas as Fama-French three factor model, Andersen, Bollerslev, Diebold, and Wu (2006) assess dynamics of monthly realized betas of portfolios in a fractional co-integration model, Andersen, Bollerslev, Diebold, and Wu (2005) relate monthly realized betas of portfolios to macroeconomic fundamentals within a linear state space model, and Hansen, Lunde, and Voev (2010) develop a Realized Beta GARCH. All these studies use a parametric model across days and abstract from the estimation error in beta estimators. The aim of this paper is to develop tools to aid analysis of time variation of betas with high frequency data that satisfy four criteria. First, we do not impose any particular parametric structure across time. Second, we fully account for the measurement error in beta estimates. Third, we allow for the presence of market microstructure noise enabling us to use tick time data. Fourth, we allow for general specifications of the market microstructure noise by using a multivariate version of the inference method of Kalnina (2011).

Full record transaction prices are asynchronously observed and known to be contaminated with market microstructure noise, such as bid-ask bounce. Several papers allow estimation of betas in this setting: Barndorff-Nielsen, Hansen, Lunde, and Shephard (2010), Christensen, Kinnebrock, and Podolskij (2010), and Zhang (2011). We construct tests for betas based on the estimator of Zhang (2011). An inference procedure for betas is required to construct any tests. Since none is proposed in Zhang (2011), we use the inference procedure of Kalnina (2011). It has the advantages of being simple to implement and being robust to different specifications of the market microstructure noise such as autocorrelation and heteroscedasticity. These advantages derive from the fact that it does not rely on the exact expression of the asymptotic variance of the beta estimators, which can have complicated expressions in the presence of market microstructure noise.

In order to apply the method of Kalnina (2011) for our purpose, we extend it in two directions.

First, we show the robustness of the subsampling method of Kalnina (2011) to long memory in spot variance process.[1] Long memory in the asset price variances is a well documented stylized fact. Early contributions include Ding, Granger, and Engle (1993) and Bollerslev and Mikkelsen (1996). See also Comte and Renault (1998), Giraitis, Kokoszka, Leipus, and Teyssière (2003), and Sizova (2011). Long memory in spot variance process can explain long memory in realized betas, which has been assumed in, e.g., Andersen et al. (2004), Andersen et al. (2005), and Bandi et al. (2008).

Second, we extend the univariate subsampling method of Kalnina (2011) to the multivariate case. The univariate method cannot be applied directly to the beta estimators even in the case of one asset, because they do not satisfy a basic requirement of her method, which is additivity of the variance of the beta estimator over time. We provide intuition for this in Section 4.2. On the other hand, we can easily derive the asymptotic variance of a beta estimator by the Delta method if we use a multivariate inference method. An appealing feature of this multivariate subsampling method is that it always produces estimates of the variance-covariance matrix that are positive semi-definite.

We then proceed to construct a test of betas being constant over time. This test is applicable for one stock or for several stocks simultaneously. Finally, we determine break times in betas nonparametrically. We do so by accounting for multiple testing among all the possible break dates by controlling familywise error rate. We implement these tests with six stocks on the NYSE in year 2006, a period of relative tranquility in volatilities and hence betas, with Standard and Poors Depositary Receipts (SPIDERS) as a proxy for the market factor. As a benchmark, we first calculate the tests with traditional 5, 10, or 20 minute based estimators, and find they can easily detect significant time variation over the whole year 2006 and also over some quarters for some stocks. High frequency estimators can detect significant time variation in betas for every quarter and every stock we consider. In a joint test across 6 assets, moderate frequency estimators

[1]Long memory in the spot variance process implies long memory in the process of integrated variances with the same persistence parameter (Rossi and Magistris, 2011).

can also detect time variation in betas for every quarter. However, finding the time of the break requires more power, and this is where the real difference between the methods emerges. We find that moderate frequency methods do not have enough power to detect a single break for most of the stocks. On the other hand, methods exploiting highest frequency data find breaks among most of the weeks in 2006 for each of the stocks.

The remainder of this paper is organized as follows. Section 2 introduces the model and defines beta. Section 3 describes the available literature that provides beta estimators in both moderate and high frequency settings. Section 4 shows robustness of the subsampling method of Kalnina (2011) to long memory in volatility and extends it to the multivariate framework. Section 5 describes the tests for constant betas and the nonparametric search for breaks. Section 6 investigates the finite sample properties of the proposed methods via simulations. Section 7 implements these with high frequency data. Section 8 concludes. All proofs are collected in the appendix.

2 The model

Denote X to be the log-price process of the market portfolio, and Y to be a log-price process of an individual stock. Suppose they both follow a continuous bivariate Brownian semimartingale process,

$$\begin{aligned} dX_t &= \mu_t^x dt + \sigma_t^x dW_t^x \\ dY_t &= \mu_t^y dt + \sigma_t^y dW_t^y \end{aligned} \tag{1}$$

over k intervals of length one, i.e., $t \in [i-1, i)$, $i = 1, \ldots, k$, where k is some fixed number. These intervals will correspond to weeks in the empirical application. In the above, μ^x and μ^y are predictable locally bounded drift processes, σ^x and σ^y are adapted càdlàg volatility processes, and W^x and W^y are standard Brownian Motion processes with $\text{Corr}\left(W_t^x, W_t^y\right) = \eta_t$. The case of N stocks corresponds to $(N+1)$-variate process and is used later in Section 5.2, but so far we use one stock for simplicity of exposition.

In continuous time, a natural measure of the variability of the process X over some interval $[i-1, i)$ is its quadratic variation,

$$\langle X, X \rangle_i = \int_{i-1}^{i} \left(\sigma_t^x\right)^2 dt,$$

and similarly covariation

$$\langle X, Y \rangle_i = \int_{i-1}^{i} \sigma_t^x \sigma_t^y \eta_t dt$$

measures the covariability of X and Y. Using these, we can define the integrated beta (or just beta) over interval $[i-1, i)$ as

$$\beta_i := \frac{\langle X, Y \rangle_i}{\langle X, X \rangle_i}. \tag{2}$$

This paper develops nonparametric tests of time variation of integrated beta β_i across time, as well as a procedure to search for breaks in betas. We now briefly discuss several perspectives, from which this integrated beta has meaning in financial economics.

First, consider a one factor (cross-sectional) model,

$$r_s = \alpha_s + \beta_s r_0 + \varepsilon_s, s = 1, ..., N,$$

where N is the number of assets, r_s is the return on the s^{th} asset and r_0 is the return on the systematic risk factor (such as the market), with innovations ε_s uncorrelated with r_0. Then, over a fixed period of time $[i-1, i)$, it would be encompassed by the following model in continuous time,

$$dp_{st} = \alpha_{st} dt + \beta_{is} \sigma_{0t} dW_{0t} + \sigma_{st} dW_{st}, s = 1, ..., N.$$

In the above, the systematic risk factor is represented by $\sigma_{0t} dW_{0t}$, and $\sigma_{st} dW_{st}$ represents the idiosynchratic risk. It can be easily verified that the coefficient β_{is} equals to

$$\beta_{is} = \frac{\langle p_s, p_0 \rangle_i}{\langle p_s, p_s \rangle_i}.$$

Denoting p_0 by X and p_1 by Y to minimise the number of subscripts, the above equality exactly equals the integrated beta definition in equation (2). This interpretation is equivalent to one in Bollerslev and Zhang (2003) and Todorov and Bollerslev (2009). The methods remain the same for multi-factor representations.

Second, knowledge of integrated betas can inform intertemporal asset pricing model formulations. Conditional CAPM requires to specify the information sets of investors, which is impossible in practice. One solution to this problem has been to assume that betas are constant over some relatively short periods of time, such as quarters or years (see, e.g., Lewellen and Nagel, 2006). After such decision, typically no formal tests are made as to whether this assumption is sufficient. Using a general nonparametric diffusion model for stock returns, we can avoid several misspecification sources, and find breaks in betas nonparametrically. However, one should also keep in mind the limitations of the high frequency data. The NYSE Trade and Quote Database, which records intraday transactions data, is only available since 1993. The time span of the data is too small to obtain good tests for whether changes in integrated betas are indeed due to changes in long-run asset market equilibrium. Hence, the current paper is not a direct contribution to this literature. Rather, it presents an additional tool that can aid finding breaks nonparametrically in high frequency betas, which should be useful in formulating parametric asset pricing models and risk management more generally.

Alternatively, integrated betas can aid risk management and porfolio decisions as measures of exposure to systematic risks without reference to specific asset pricing models. As a simple example, suppose an investor aims to hedge exposure to being short of one unit of Y_t by holding ρ_t units of X_t. Denoting by Z_t the gain or loss by following this procedure, we have the following relationship

$$dY_t = \rho_t dX_t + dZ_t, 0 \leq t \leq T.$$

Then, ρ_t, or instantaneous beta, is obtained by minimizing the residual sum of squares. Under continuous observations, the residual sum of squares is the quadratic variation of Z, $\langle Z, Z \rangle_T$.[2] Therefore, instantaneous beta becomes

$$\rho_t = \frac{d \langle X, Y \rangle_t}{d \langle X, X \rangle_t} = \frac{\langle X, Y \rangle_t'}{\langle X, X \rangle_t'},$$

where $d \langle X, Y \rangle_t$ is the derivative of $\langle X, Y \rangle_t$ with respect to time (in our setup ρ_t simplifies to $(\sigma_t^x)^{-1} \sigma_t^y \eta_t$), see Mykland and Zhang (2006). The relationship between this instantaneous beta ρ_t and the integrated beta β_i defined by equation (2) is straighforward. If ρ_t is constant across any interval $[i-1, i+1)$, then integrated betas over $[i-1, i)$ and $[i, i+1)$ would have to be equal, $\beta_i = \beta_{i+1}$. Therefore, rejection of the hypothesis $\beta_i = \beta_{i+1}$ implies rejection of the hypothesis of ρ_t being constant over the interval $[i-1, i+1)$. The opposite is not true, the intuition being that some higher frequency movements of ρ_t over time would be missed by the integrated beta.[3]

Before proceeding, we note that some care should be taken of what a null hypothesis of constant betas means. The integrated betas are random, so the unconditional probability that two different betas happen to be the same is zero. Therefore, a better interpretation is to conduct all analysis conditional on the realized path of the spot volatility. Conditional on the volatility path, betas are constant. Thus, when we talk about a null hypothesis of two betas being the same, we mean that the volatility path has realized such that betas are equal.[4] Having said that, all analysis in this paper is written without conditioning on the volatility path. It does not change the form of our test statistics or their asymptotic distribution.[5]

[2] Recall that quadratic variation of Z can be defined as

$$\langle Z, Z \rangle_T = \lim_{\pi \downarrow 0} \sum_{i=1}^{n} \left(Z_{t_i} - Z_{t_{i-1}} \right)^2$$

where we consider some partition $0 = t_0 < t_1 < ... < t_n = T$ of the interval $[0, T]$, and $\pi = \max_i |t_i - t_{i-1}|$.

[3] Spot beta can be estimated by localization in time of the integrated beta estimators, see Bandi and Reno (2010), Kristensen (2010), and Zu and Boswijk (2009).

[4] I am grateful to Christian Gourieroux for suggesting this interpretation.

[5] This is because the extra randomness that arises in the limit is independent of the spot volatility process, see proof of Proposition 2 for an example.

3 Existing theory for estimation of beta

We first review the existing methods we aim to use. While the focus of the paper is on data at the highest frequency, the setting of moderate frequencies serves as an important benchmark. We start with this benchmark case due to its relative simplicity.

3.1 Noise-free and synchronous data

We refer to moderate frequency data a situation when sampling frequencies are low enough for the effects of market microstructure to be negligible, e.g., 5, 15, or 20 minutes.[6] In these cases, the data is interpolated to a regular grid of calendar time, but the effect of interpolation is usually negligible. Thus, the underlying assumption is that we have discrete equi-distant observations on the continuous time log-price process in equation (1). Denote by n the number of observations in each time period. Normalizing the length of period to be 1, the distance between observations becomes $1/n$. In this relatively simple scenario we can estimate quadratic variation of X by realized variance

$$[X,X]_i = \sum_{j=1}^{n} \left(X_{(i-1)n+j} - X_{(i-1)n+j-1}\right)^2$$

and quadratic covariation between X and Y by realized covariance,

$$[X,Y]_i = \sum_{j=1}^{n} \left(X_{(i-1)n+j} - X_{(i-1)n+j-1}\right)\left(Y_{(i-1)n+j} - Y_{(i-1)n+j-1}\right).$$

These can be used to calculate the so-called realized beta as explored by Andersen et al. (2004), Andersen et al. (2005), and Barndorff-Nielsen and Shephard (2004),

$$\widehat{\beta}_i^{RV} := \frac{[X,Y]_i}{[X,X]_i}.$$

It is consistent for the true beta as the sampling becomes more and more frequent (i.e., as $1/n \to 0$), under the assumption of no noise and synchronous observations. Asymptotic distribution of realized covariation matrix was first derived in Barndorff-Nielsen and Shephard (2004). We have

$$\sqrt{n}\begin{pmatrix} [X,X]_i - \langle X,X\rangle_i \\ [X,Y]_i - \langle X,Y\rangle_i \end{pmatrix} \Rightarrow MN\left(0,\Psi_i\right)$$

where

$$\Psi_i = \int_{i-1}^{i} \begin{pmatrix} 2\sigma_x^4(u) & 2\sigma_x^3(u)\,\sigma_y(u)\,\eta(u) \\ 2\sigma_x^3(u)\,\sigma_y(u)\,\eta(u) & \sigma_x^2(u)\,\sigma_y^2(u)\left(1+\eta^2(u)\right) \end{pmatrix} du. \tag{3}$$

[6]Additional precision can be gained by sampling at data-driven moderate frequencies, see Bandi and Russell (2008).

and where $MN(0, \Psi_i)$ means mixed normal distribution with random conditional variance Ψ_i independent of the underlying normal distribution.[7] The above convergence is stable in law, which is slightly stronger than the usual convergence in law, see e.g., Aldous and Eagleson (1978), which means that confidence intervals intervals can be constructed as usual even though Ψ_i is random. By the Delta method, provided $\langle X, X \rangle_i > 0$ for $i = 1, .., k$,

$$\sqrt{n}\left(\widehat{\beta}_i^{RV} - \beta_i\right) \Rightarrow MN\left(0, V_i^{RV}\right)$$

where

$$V_i^{RV} = \langle X, X \rangle_i^{-2} \begin{pmatrix} -\beta_i & 1 \end{pmatrix} \Psi_i \begin{pmatrix} -\beta_i \\ 1 \end{pmatrix}.$$

From Barndorff-Nielsen and Shephard (2004), we know that the asymptotic variance of $\widehat{\beta}_i^{RV}$ can be estimated by

$$[X, X]_i^{-2} \left(\sum_{j=1}^{n} \kappa_{i,j}^2 - \sum_{j=1}^{n-1} \kappa_{i,j}\kappa_{i,j+1} \right) \tag{4}$$

where

$$\begin{aligned} \kappa_{i,j} &= \left(X_{(i-1)n+j} - X_{(i-1)n+j-1}\right)\left(Y_{(i-1)n+j} - Y_{(i-1)n+j-1}\right) \\ &\quad -\widehat{\beta}_i^{RV}\left(X_{(i-1)n+j} - X_{(i-1)n+j-1}\right)^2. \end{aligned}$$

Joint distribution of estimated betas across different time periods can be obtained from marginals, since asymptotic distributions of $\sqrt{n}\left(V_i^{RV}\right)^{-1/2}\left(\widehat{\beta}_i^{RV} - \beta_i\right)$ are independent for any $i \neq j$ (see e.g. Mykland and Zhang, 2006).

The estimator of Barndorff-Nielsen (2004) is the first and most popular estimator of Ψ_i and hence of the asymptotic variance of $\widehat{\beta}_i^{RV}$ in practice. Potentially better approximations are provided by bootstrapping of Donovon, Gonçalves, and Meddahi (2008) and blocking of Mykland and Zhang (2009).

3.2 Noisy and asynchronously observed data

We now turn to the case of highest frequency data, which is full record transaction prices. These are irregular (observed at non-equidistant times), asynchronous across stocks, and are known to be contaminated by market microstructure noise. The same structure applies to every interval (such as a day or a week), so to facilitate the notation in this section, we omit the reference to multiple intervals, use a generic time interval $[0, T]$, and use notation similar to Zhang (2011). The market microstructure noise is typically modeled as

[7] In other words, the limiting distribution is that of the random variable $\Psi_i^{-1/2} Z$, which is a product of two (multivariate) independent random variables, $\Psi_i^{-1/2}$ and a standard normal random variable Z. Conditional on the volatility path $\{\sigma_t\}_{t \geq 0}$, Ψ_i is nonstochastic and the distribution of $\sqrt{\Psi_i} Z$ is normal, $N(0, \Psi_i)$.

an additive measurement error, so one assumes that latent prices X^* and Y^* still follow diffusions as in equation (1), but we only have noisy observations on them, $X = X^* + \epsilon^x$ and $Y = Y^* + \epsilon^y$. Assumptions about the noise are collected in the Assumption N below.

We start by describing the Refresh Time synchronisation process, which was introduced by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2010). Suppose we have n_x observations on X and n_y observations on Y. Let $\mathcal{T}_{n_x}$ and Γ_{n_y} be the sets that contain observations on X and Y,

$$\mathcal{T}_{n_x} := \{0 = \tau_{n_x,0} < \tau_{n_x,1} < \ldots < \tau_{n_x,n_x} = T\}, \Gamma_{n_y} := \left\{0 = \gamma_{n_y,0} < \gamma_{n_y,1} < \ldots < \gamma_{n_y,n_y} = T\right\}.$$

Then, Refresh Times $0 = v_0 < v_1 < \ldots < v_n = T$ are those times when both stocks have traded again. In other words, v_i is set to be the maximum of $\min\{\tau \in \mathcal{T}_{n_x} : \tau > v_{i-1}\}$ and $\min\{\gamma \in \Gamma_{n_y} : \gamma > v_{i-1}\}$. To obtain observations at Refresh Times, one uses previous tick interpolation,

$$t_i = \max\{\tau \in \mathcal{T}_{n_x} : \tau \le v_i\} \text{ and } s_i = \max\{\gamma \in \Gamma_{n_y} : \gamma \le v_i\}, \tag{5}$$

so that t_i and s_i become the new sampling points of X and Y.[8] n becomes the new number of observations. One needs to assume that distances between sampling points are not too large, see Condition C2 of Zhang (2011).

Beta can be estimated with high frequency data using Two Scale Realized Volatility or Multi Scale Realized Volatility estimator of Zhang (2011), Realized Kernels of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2010), or pre-averaging estimators of Christensen, Kinnebrock, and Podolskij (2010). We use the Two Scale estimator. It uses the following assumptions about the noise:

Assumption N. *The noise $\left(\epsilon^x_{t_i}, \epsilon^y_{s_i}\right)$ is independent of the efficient price (X, Y), it is stationary, exponentially α-mixing, and both $\epsilon^x_{t_i}$ and $\epsilon^y_{s_i}$ have finite $(4+\delta)^{th}$ moment for some $\delta > 0$.*

Beta can be estimated as follows using the Two Scale estimators[9]

$$\widehat{\beta}^{AMZ} = \frac{\widehat{\langle X, Y\rangle}^{AMZ}}{\widehat{\langle X, X\rangle}^{AMZ}}. \tag{6}$$

In the above equation (6), $\widehat{\langle X, Y\rangle}^{AMZ}$ is defined as

$$\widehat{\langle X, Y\rangle}^{AMZ} = [X, Y]^{(G_1)} - \frac{\overline{n}_{G_1}}{\overline{n}_{G_2}} [X, Y]^{(G_2)} \tag{7}$$

[8] For practical implementation, the algorithm in Palandri (2006) representing this synchronization process is useful.

[9] $[0, T]$ represents time intervals $[i, i+1)$, $i = 1, \ldots, k$. For estimation in Section 7 these correspond to weeks, while synchronisation is done for every day separately. While in theory differences in interval definitions do not matter, in practice there are overnight jumps, which are not easily treated in the framework of semimartingales. The overnight jumps are deleted together with other jumps in our empirical application.

where $[X,Y]^{(G_l)}$ is subsampled (i.e., calculated on sparse data) and averaged realized variance:

$$
\begin{aligned}
[X,Y]^{(G_l)} &= \frac{1}{G_l}\sum_{j=G_l}^{n}\left(X_{t_j}-X_{t_{j-G_l}}\right)\left(Y_{s_j}-Y_{s_{j-G_l}}\right)\ ,\ l=1,2 \\
\overline{n}_{G_l} &= \frac{n-G_l-1}{G_l},\ l=1,2
\end{aligned}
$$

with G_1 and G_2 two constants determining the two 'time scales", i.e., frequency of returns used in $[X,Y]^{(G_l)}$.

From the joint asymptotic distribution

$$
n^{1/6}\left(\left(\begin{array}{c}\widehat{\langle X,X\rangle}^{AMZ}\\ \widehat{\langle X,Y\rangle}^{AMZ}\end{array}\right)-\left(\begin{array}{c}\langle X,X\rangle\\ \langle X,Y\rangle\end{array}\right)\right)\Rightarrow MN\left(0,\Sigma^{AMZ}\right), \tag{8}
$$

we obtain the asymptotic distribution for realized beta by the Delta method (provided $\langle X,X\rangle>0$ a.s.),

$$
n^{1/6}\left(\widehat{\beta}^{AMZ}-\beta\right)\Rightarrow MN\left(0,V^{AMZ}\right) \tag{9}
$$

where

$$
V^{AMZ}=\langle X,X\rangle^{-2}\left(\begin{array}{cc}-\beta & 1\end{array}\right)\Sigma^{AMZ}\left(\begin{array}{c}-\beta\\ 1\end{array}\right).
$$

The exact expression of Σ^{AMZ} is rather complicated, and the reader can find it in Zhang (2011). We do not need the exact expression for estimation because we will use subsampling to estimate Σ^{AMZ}, see Section 4.2. The above method is used for every period $[i-1,i)$, $i=1,\ldots,k$. Asymptotic distributions of $\sqrt{V^{AMZ}}n_i^{1/6}\left(\widehat{\beta}_i^{AMZ}-\beta_i\right)$ are again independent across periods i, as in the no noise case in Section 3.1.

The next section proves robustness of the subsampling method of Kalnina (2011) against long memory volatility and extends the method to the multivariate case. Section 5 then builds on the two preceeding sections to construct methods for investigation of time variation in betas.

4 Construction of robust confidence intervals by subsampling

Our aim in this section is to estimate Σ_i^{AMZ} in an automatic way, i.e., doing so without using the exact expression for Σ_i^{AMZ}. Our main motivation is to obtain confidence intervals that are robust to different specifications of the underlying processes. In particular, Kalnina (2011) shows that her univariate subsampling method is not only robust to autocorrelation but also heteroscedasticity of the market microstructure noise process. All these features show up in the expression for Σ_i^{AMZ} and hence make element by element estimation of Σ_i^{AMZ} prone to misspecification. However, one first has to extend the subsampling method of Kalnina (2011) to the multivariate case before using it on betas. This is done in Section 4.2 below.

In the same spirit of robust confidence intervals, we note that in general, the assumption of Brownian Semimartingale for the spot volatility (see Assumption A1 below) precludes long memory. This is rather unfortunate given that long memory feature in stock price volatility is a well documented stylized fact. Brownian Semimartingale structure for spot volatility is otherwise rather general and has been used in a series of papers, including that of Kalnina (2011).[10] Therefore, we start by demonstrating robustness of her subsampling method to long memory in the spot volatility process.

For simplicity of notation, we omit reference to multiple intervals in this section and work with a generic interval $[0, T]$.

4.1 Subsampling with long memory volatility

The subsampling method of Kalnina (2011) rests on marriage of two ideas. The first is the subsampling method of Politis and Romano (1994), the second is time localization. To see how the first is meant to work, suppose there is some general estimator $\widehat{\theta}_n$ (think of i.i.d. $Y_i's$, a parameter of interest $\theta = E(Y)$, and $\widehat{\theta}_n = \frac{1}{n}\sum Y_i$). Suppose we know its asymptotic distribution

$$\sqrt{n}(\widehat{\theta}_n - \theta) \Longrightarrow N\left(0, V\right),$$

but we would like to estimate V, in order to be able to construct confidence intervals for $\widehat{\theta}_n$. Construct K different subsamples of $m = m(n)$ consecutive observations, starting at different values (whether they are overlapping or not is irrelevant here), where $m \to \infty$, $m/n \to 0$. Denote by $\widehat{\theta}_{n,m,l}$ the estimator $\widehat{\theta}_n$ calculated using the l^{th} block of m observations. Then, under the stationarity assumption the asymptotic distribution of $\sqrt{m}(\widehat{\theta}_{n,m,l} - \theta)$ is the same, i.e.,

$$\sqrt{m}\left(\widehat{\theta}_{n,m,l} - \theta\right) \Longrightarrow N\left(0, V\right).$$

Moreover, one could replace the unobservable θ by $\widehat{\theta}_n$ and above result would still hold. Therefore, $m\left(\widehat{\theta}_{n,m,l} - \widehat{\theta}_n\right)^2$ is one noisy observation on V. By averaging over many such observations we can estimate V consistently.

As it stands this idea does not work in our context. Realized Volatility (RV), for example, on the whole interval does not approximate the same object as RV on some small subsample. The second idea is therefore to use above construction locally in time, provided V is additive in time, say if it can be written as

$$V = \int_0^T f(\sigma(t))dt, \tag{10}$$

[10] See e.g. Christensen, Oomen and Podolskij (2009). Examples of papers with a similar assumption that also allows for jumps in volatility are Jacod (2008), Jacod, Podolskij, and Vetter (2010), and Todorov and Tauchen (2011).

for some function f. We cannot use $\widehat{\theta}_n$ instead of θ this time, so an additional set of subsamples has to be constructed. One way of obtaining the second set of subsamples is to use longer subsamples, see Figure 1 for a graphical illustration.[11] Then, $m\left(\widehat{\theta}_l^{short}-\widehat{\theta}_l^{long}\right)^2$ becomes a noisy observation on $f(\sigma(t))$ for some t, and averaging over subsamples delivers a consistent estimate of V.[12]

Two things follow from this discussion. First, additivity in time of V, (see e.g. equation 10) is crucial. Asymptotic variances of beta estimators do not satisfy this condition. Therefore, beta cannot be subsampled directly. However, if one extends these ideas to the multivariate framework, asymptotic variance of beta can be estimated by subsampling estimators of $\langle X,Y\rangle$ and $\langle X,X\rangle$ jointly and using a Delta method. Therefore, Section 4.2 below proposes the multivariate subsampling method.

Second, some smoothness assumption on volatility is needed for this subsampling scheme to work. This is evident by looking at the Figure 1. The subsample that $\widehat{\theta}_1^{long}$ is calculated on does not cover the same interval as $\widehat{\theta}_1^{short}$. For the subsampling idea to work, these estimators have to approximate the same quantity. Suppose a jump occurs in volatility of returns that are used in $\widehat{\theta}_1^{long}$, but not in $\widehat{\theta}_1^{short}$. Then, the two do not approximate the same quantity and $\widehat{\theta}_1^{long}$ cannot be used to demean $\widehat{\theta}_1^{short}$ and extract its variance. To ensure sufficient smoothness, Kalnina (2011) assumed volatility follows a diffusion:

Assumption A1. *The volatility process $\{\sigma(t), t\in[0,T]\}$ is a Brownian semimartingale of the form*

$$d\sigma(t)=\tilde{\mu}(t)dt+\tilde{\sigma}(t)d\widetilde{W}(t)$$

where $\widetilde{W}(t)$ is standard Brownian motion, the stochastic process $\widetilde{\mu}(t)$ is locally bounded and the stochastic process $\widetilde{\sigma}(t)$ is càdlàg.

This assumption includes many parametric models used in practice, but it excludes long memory dynamics such as[13]

Assumption A2. *The logarithm of the volatility process $x(t)=\ln\sigma(t)$ solves the following first-order fractional SDE:*

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

where

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

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

We now show that the subsampling of Kalnina (2011) is robust to the above long memory in volatility:

[11] See Section 2.3 of Kalnina (2011) for an alternative construction of the subsamples.

[12] Some scaling factor might be needed to ensure this object is of order 1, see equation (12) for exact expression.

[13] This is because fractional Brownian motion with $\alpha\neq 1/2$ is not a semimartingale, see e.g. Compte (2005).

Proposition 1. *Suppose log-price X satisfies equation (1), and the noise ϵ_{t_i} satisfies Assumption N. Let $\widehat{\theta}_n$ be the TSRV estimator $\widehat{\langle X, X\rangle}^{AMZ}$ defined by (7), with parameters G_1 and G_2 satisfying $G_1 = \lfloor cn^{2/3} \rfloor$ for some constant c, G_2 is such that $\operatorname{Cov}(\epsilon_1, \epsilon_{G_2}) = o\left(n^{-1/2}\right)$, $G_2 \to \infty$, and $G_2/G_1 \to 0$. Let V be defined by V^{AMZ} in (9), and $\widehat{V}_{sub}$ be defined by $\widehat{\Sigma}_{sub}$ in (12) with below with $s = m$. Let $J \to \infty$, $m \to \infty$, $J/m \to \infty$, $m/n \to \infty$, $G_1/J \to 0$. Suppose one of the two:*

(a) σ satisfies Assumption A1 and $Jmn^{-5/3} \to 0$, or

(b) σ satisfies Assumption A2 and $Jm^{\alpha}n^{-2/3-\alpha} \to 0$.

Then,

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

The proof of Proposition 1b can be found in the appendix. Proposition 1a is equivalent to Theorem 4 of Kalnina (2011) and is stated here for ease of comparison. Note that the long memory parameter α enters the conditions for Proposition 1b. That is because some smoothness condition is needed on σ due to the time localization used in the construction of $\widehat{V}$, and more persistent memory gives more smooth sample paths. A typical estimate in the literature is $\alpha = 0.4$, see for example Andersen, Bollerslev, Diebold, and Labys (2002), Bandi and Perron (2006), and Ray and Tsay (2000).

The above proposition can easily be extended to accomodate heteroscedastic noise as in Kalnina and Linton (2008), provided the heteroscedasticity does not destroy the additive bias correction and hence consistency of the Two Scale estimator. In the latter case, one can always use the *jittered* Two Scale estimator proposed in Kalnina and Linton (2008) to restore the consistency of the Two Scale estimator.

We now turn to the multivariate extension of the subsampling procedure of Kalnina (2011).

4.2 Multivariate subsampling

The aim of this section is to generalize the ideas of Kalnina (2011) to the multivariate framework, thus producing an automatic, positive semi-definite estimator of the asymptotic variance-covariance matrix.[14] This will be applied to tick data to estimate Σ^{AMZ} in (8). In general, we seek to estimate matrix Σ in

$$\tau_n \left(\widehat{\theta} - \theta\right) \Rightarrow MN\left(0, \Sigma\right) \tag{11}$$

where τ_n is a known rate of convergence when n observations are used. The rate of convergence is assumed to be the same for all elements of $\widehat{\theta}$. For our purposes

$$\theta = \begin{pmatrix} \langle X, X\rangle \\ \langle X, Y\rangle \end{pmatrix},$$

[14]The sense in which this method is automatic is that it does not rely on the information about the expression for the asymptotic variance V.

and $\widehat{\theta}$ is later taken to be the Two Scale estimator. Construct a series of longer blocks of observations, m returns in each block, as well as a series of shorter blocks of observations, J returns in each block, $J < m < n$, see Figure 1. Denote $\mathcal{A}$ to be a set containing some observation times, and $\widehat{\theta}(\mathcal{A})$ to be an estimator $\widehat{\theta}$ calculated using observations at times in $\mathcal{A}$. Using this notation, the subsampling estimator of the asymptotic variance-covariance matrix Σ is

$$\widehat{\Sigma}_{sub} = \frac{J}{n}\frac{1}{K}\sum_{l=1}^{K}\tau_n^2\left(\frac{n}{J}\widehat{\theta}_l^{short} - \frac{n}{m}\widehat{\theta}_l^{long}\right)\left(\frac{n}{J}\widehat{\theta}_l^{short} - \frac{n}{m}\widehat{\theta}_l^{long}\right)' \tag{12}$$

where

$$\begin{aligned}
\widehat{\theta}_l^{long} &= \widehat{\theta}\left(\left\{t_{(l-1)s+1}, t_{(l-1)s+2}, ..., t_{(l-1)s+m+1}\right\}\right) \\
\widehat{\theta}_l^{short} &= \widehat{\theta}\left(\left\{t_{(l-1)s+1}, t_{(l-1)s+2}, ..., t_{(l-1)s+J+1}\right\}\right) \\
K &= \left\lceil \frac{n-m}{s} + 1 \right\rceil,
\end{aligned}$$

K is the number of subsamples, and s stands for 'shift", i.e., by how many observations to roll the window to obtain the next subsample. Thus, it controls the amount of overlap between the subsamples. The smallest s is 1 and it corresponds to the maximum overlap and largest number of subsamples; Figure 1 is drawn for this case. This choice also gives the smallest asymptotic variance. However, it can be very computationally intensive in practice, so a larger s can also be used at the expense of less efficient, but nevertheless consistent $\widehat{\Sigma}_{sub}$. From the definitions of $\widehat{\theta}_l^{short}$ and $\widehat{\theta}_l^{long}$ above, we can see that longer and shorter subsamples start at the same time. This case is less involved to write down, but in practice the case drawn in Figure 1 is slightly better, i.e., both subsamples are centered at the same time. For this case, shorter subsample should start at $t_{(l-1)s+\lfloor(m-J)/2\rfloor+1}$ and not $t_{(l-1)s+1}$. Also, notice that the formula in (12) simplifies to one in Kalnina (2011) for univariate estimator and no overlap case (i.e., $s = m$ and $K = \lceil n/m \rceil$).

The proof of the consistency of $\widehat{\Sigma}_{sub}$ follows exactly the same steps as in Theorem 4 of Kalnina (2011), and hence is omitted. The assumptions needed are essentially the same. The latter theorem proves the consistency of $\widehat{\Sigma}_{sub}$ for the univariate case when $\theta = \langle X, X \rangle$ and when θ is estimated by the Two Scale estimator. The assumptions on the lengths of subsamples (m and J) and the parameters of the Two Scale estimator (G_1 and G_2) are the same as in the univariate case, $J \to \infty$, $m \to \infty$, $J/m \to 0$, $G_1/J \to 0$, and $Jmn^{-5/3} \to 0$. The latent process X should follow multivariate brownian semimartingale as in (16). Also, the spot volatility should also follow a multivariate Brownian semimartingale. Theorem 4 of Kalnina (2011) assumes equidistant observations. This can be extended to irregular and endogenous (but synchronous) observations case as follows. Suppose that $t_n = F(i/n)$ where F is a smooth random process which does not depend on n. In particular $F(t) = \int_0^t \eta^2(u)du$ where $\eta^2(u)$ has strictly positive, càdlàg sample paths. Then all

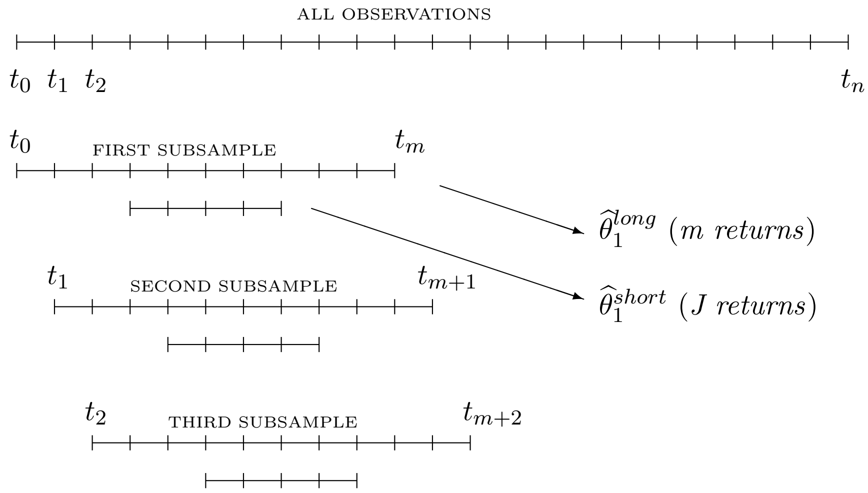


Figure 1: *The Subsampling Scheme of Kalnina (2011)*

calculations can be implemented in transaction times, sampling say every 5 transactions where the formulas call for equal distances in time. This is theoretically equivalent to working with a Brownian semimartingale with drift $\mu_{F(u)}\eta(u)$ and spot volatility $\sigma_{F(u)}\eta(u)$ instead of $\mu(u)$ and $\sigma(u)$, see Proposition 2 of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008a). We note that settings with much stronger endogeneity in sampling times are possible, which distort the asymptotic distribution more, up to the extent that prevents construction of any confidence intervals due to lack of mixed asymptotic normality, see Li, Mykland, Renault, Zhang, and Zheng (2009). The last issue is asynchronous observations. Zhang (2011) shows that the bias correction for the noise of the Two Scale estimator also corrects for the effect of asynchronous observations, so that there is no effect on the asymptotic distribution on the Two Scale estimator. It is therefore intuitive that subsampled Two Scale estimator will not have any effect from asynchronous observations, but the proof is highly technical and so we do not pursue it here. Monte Carlo section demonstrates this in finite sample for several designs.

It is easy to see that $\widehat{\Sigma}_{sub}$ is positive semi-definite by construction. This avoids any risk of length of estimated confidence intervals for $\widehat{\beta}$ (or other continuous functionals of elements of $\widehat{\theta}$) being negative.

With these tools in hand, we turn to constructing different tests for time variation in beta. We then use these in section 5.3 to propose a nonparametric method of search for breaks in beta.

5 Testing for constant betas with high frequency data

5.1 The test statistic for the beta of a single asset

Using the joint asymptotic distribution of betas across k time periods, we can construct a Chi-square test for the true betas being constant across these time periods. The number of time periods k is any fixed positive integer. The construction will be based on $\widehat{\beta}^{\text{AMZ}}$ and $\widehat{\beta}^{RV}$, so we use a generic $\widehat{\beta}$ to denote any of them. Define

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

Given that standardized asymptotic distributions of $\widehat{\beta}_i - \beta_i$ are independent across time periods $i = 1, 2, \ldots, k$ (see e.g. Mykland and Zhang, 2006), we obtain the joint asymptotic distribution from the marginal ones. The asymptotic variance covariance matrix will be diagonal. We have

$$\Phi\left(\widehat{\beta} - \beta\right) \Rightarrow MN\left(0, V\right) \tag{13}$$

where

$$\begin{aligned} V &= diag\left(V_1, V_2, ..., V_k\right) \\ \Phi &= diag\left(\tau_{n_1}, \tau_{n_2}, \ldots, \tau_{n_k}\right). \end{aligned}$$

In above, τ_{n_i} is the rate of convergence of the estimator, so that $\tau_{n_i} = n_i^{1/2}$ for $\widehat{\beta}^{RV}$ and $\tau_{n_i} = n_i^{1/6}$ for $\widehat{\beta}^{\text{AMZ}}$. The observations can be different across days, but they should be of the same magnitude asymptotically, see Proposition 2 below.

We are interested in testing the hypothesis that true beta is constant over time,

$$H_0 : \beta_1 = \ldots = \beta_k, \text{ versus } H_1 : \beta_i \neq \beta_j \text{ for some } i \text{ and } j. \tag{14}$$

Our test statistic is a sum of squared differences $\widehat{\beta}_i - \widehat{\beta}_1$ for $i = 2, \ldots, k$, properly standardized. For this purpose, introduce the following $k-1$ by k matrix

$$\underset{(k-1)\times k}{\Delta} = \begin{pmatrix} -1 & 1 & 0 & 0 & ... & 0 \\ -1 & 0 & 1 & 0 & ... & 0 \\ -1 & 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... & ... \\ -1 & 0 & 0 & 0 & ... & 1 \end{pmatrix}. \tag{15}$$

We use this matrix to obtain a vector of length $k-1$ containing all differences in betas,

$$\Delta\left(\widehat{\beta}-\beta\right)=\begin{pmatrix}\widehat{\beta}_2-\widehat{\beta}_1-(\beta_2-\beta_1)\\ \widehat{\beta}_3-\widehat{\beta}_1-(\beta_3-\beta_1)\\ \dots\\ \widehat{\beta}_k-\widehat{\beta}_1-(\beta_k-\beta_1)\end{pmatrix}.$$

Proposition 2. *Let $\widehat{\beta}$ satisfy the asymptotic distribution in (13) with $\tau_{n_i}=\tau_{n_1}(c_i+o(1))$ for some positive constants $c_i, i=2,\dots,k$. Then,*

$$\left(\widehat{\beta}-\beta\right)'\Delta'\left(\Delta\Phi^{-2}V\Delta'\right)^{-1}\Delta\left(\widehat{\beta}-\beta\right)\Rightarrow\mathcal{X}^2_{k-1}$$

holds conditionally on the volatility path $(\sigma_t)_{t\geq 0}$, as well as unconditionally. Moreover, under the H_0 in (14), $\Delta\left(\widehat{\beta}-\beta\right)=\Delta\widehat{\beta}$ and hence

$$T\equiv\left(\widehat{\beta}\right)'\Delta'\left(\Delta\Phi^{-2}\widehat{V}\Delta'\right)^{-1}\Delta\widehat{\beta}\Rightarrow\mathcal{X}^2_{k-1}$$

where $\widehat{V}$ is any consistent estimator of V.

There are several choices for the estimation of V. We recommend to estimate V^{RV} by methods of Barndorff-Nielsen and Shephard (2004) for moderate frequency and V^{AMZ} by the subsampling method of Section 4.2 for high frequency data. These are the choices used in the simulation section and empirical analysis.

Under the alternative H_1, this test statistic T diverges to infinity, meaning that the test is consistent. We use the estimate of the first beta for centering, but it can be shown that the resulting test statistic is invariant to the choice of centering.

5.2 Joint tests across several assets

The above analysis can be extended to joint tests across several assets. Suppose we have N assets where N is a fixed number. These assets plus the market factor give a $(N+1)$-dimensional vector valued stochastic process X, which we assume to be multivariate Brownian semimartingale,

$$dX_t=\mu_t dt+\sigma_t dW_t \tag{16}$$

where $\{\mu_t\}_{t\geq 0}$ is a $(N+1)$-dimensional predictable locally bounded drift vector, $\{\sigma_t\}_{t\geq 0}$ is a $(N+1)\times(N+1)$-dimensional adapted càdlàg covolatility matrix process, and $\{W_t\}_{t\geq 0}$ is a $(N+1)$-dimensional standard Brownian motion. Define vector valued integrated betas over $i-th$ interval as

$$\underset{(N\times 1)}{\beta_i}=\left(\begin{array}{ccc}\frac{\langle X^{(1)},X^{(N+1)}\rangle_i}{\langle X^{(N+1)},X^{(N+1)}\rangle_i} & \frac{\langle X^{(2)},X^{(N+1)}\rangle_i}{\langle X^{(N+1)},X^{(N+1)}\rangle_i} & \dots & \frac{\langle X^{(N)},X^{(N+1)}\rangle_i}{\langle X^{(N+1)},X^{(N+1)}\rangle_i}\end{array}\right)', i=1,2,\dots,k.$$

The null hypothesis we aim to test is

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

where $\beta_1^{(i)}$ means beta of the i^{th} stock on the 1^{st} time period etc.

Collect all betas into one $Nk-$dimensional row vector, $\beta = (\beta_1', \beta_2', \ldots, \beta_k')'$. As before, we use generic $\widehat{\beta}$ to mean either $\widehat{\beta}^{RV}$ or $\widehat{\beta}^{AMZ}$. For every interval we have the joint asymptotic distribution,

$$\tau_{n_i}\left(\widehat{\beta}_i - \beta_i\right) \Rightarrow MN\left(0, V_i\right), \; i = 1, 2, \ldots, k.$$

The rate of convergence τ_{n_i} is $n_i^{1/2}$ for $\widehat{\beta}^{RV}$ and $n_i^{1/6}$ for $\widehat{\beta}^{AMZ}$ where n_i is the number of observations for every stock on i^{th} interval. Since after standardizing by $V_i^{-1/2}$, the asymptotic estimation errors are again independent across intervals, we have

$$\tau_n\left(\widehat{\beta} - \beta\right) \Rightarrow MN\left(0, V\right) \tag{17}$$

where $V = diag\left(V_1, V_2, \ldots, V_k\right)$ and $\tau_n = diag\left(\tau_{n_1} I_N, \tau_{n_2} I_N, \ldots, \tau_{n_k} I_N\right)$. Exactly as in Proposition 2, assume $\tau_{n_i} = \tau_{n_1}\left(c_i + o(1)\right)$ for some positive constants $c_i, i = 2, \ldots, k$. We now want to consider the same type of differences as in the Section 5.1, so define $\underline{\Delta}$ to be the transformation matrix that produces them,

$$\underline{\Delta}\left(\widehat{\beta} - \beta\right) = \begin{pmatrix} \Delta\left(\widehat{\beta}^{(1)} - \beta^{(1)}\right) \\ \Delta\left(\widehat{\beta}^{(2)} - \beta^{(2)}\right) \\ \ldots \\ \Delta\left(\widehat{\beta}^{(N)} - \beta^{(N)}\right) \end{pmatrix}$$

where Δ is as in equation (15) and $\beta^{(j)}$ is a $k \times 1$ vector of betas of the j^{th} stock across all time periods. Then, similarly to Proposition 2,

$$\left(\widehat{\beta} - \beta\right)' \underline{\Delta}' \left(\underline{\Delta}\tau_n^{-2} V \underline{\Delta}'\right)^{-1} \underline{\Delta}\left(\widehat{\beta} - \beta\right) \Rightarrow \mathcal{X}^2_{(k-1)N}.$$

Under the null hypothesis, $\underline{\Delta}\left(\widehat{\beta}^{AMZ} - \beta\right) = \underline{\Delta}\widehat{\beta}^{AMZ}$, so we can define the test statistic

$$T = \widehat{\beta}' \underline{\Delta}' \left(\underline{\Delta}\tau_n^{-2}\widehat{V}\underline{\Delta}'\right)^{-1} \underline{\Delta}\tau_n\widehat{\beta},$$

which under the null is asymptotically distributed as $\mathcal{X}^2_{(k-1)N}$. In above, $\widehat{V}$ is any consistent estimator of V. We recommend to estimate V by methods of Barndorff-Nielsen and Shephard (2004) for moderate frequency and by method of Section 4.2 for high frequency data. Besides general time series dynamics, this test is also robust to very rich interactions between the assets considered.

5.3 Nonparametric search for breaks in betas

Instead of finding out whether beta is constant over a particular time period or not, we might be interested in finding the time of breakpoints in beta nonparametrically. Performing many tests at the same time is subject to multiple testing problem, meaning that some hypotheses will be rejected by chance alone (a chance that is not controlled by the significance level of individual tests).

We propose to search for breaks in betas nonparametrically by considering each possible break time as a hypothesis and accounting for multiple testing. Several aspects of our procedure are important. First, we account for multiple testing by controlling family-wise error rate. Second, we fully account for the dependence structure of the test statistics while preserving their nonparametric structure. Third, we use a stepwise procedure first initiated by Holm (1979) to increase the power. The procedure below is for one asset, but can be extended easily for search for simultaneous breaks across several assets.

Suppose we have k time intervals, and we are interested in finding the breaks in beta across these time intervals. Instead of one null hypothesis, we have $k-1$ null hypotheses. Label the null hypothesis by intervals,

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

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

$$FWE = P\left\{\text{Reject at least one true null hypothesis}\right\}, \tag{18}$$

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

Denote as before $\Phi = diag(\tau_{n_1}, \tau_{n_2}, \ldots, \tau_{n_k})$. Assume $\Phi = \tau_{n_1} C$ where C is a diagonal matrix containing positive constants.[15] Define the following $k-1$ test statistics $w_1, w_2, \ldots, w_{k-1}$ based on the differences in beta estimates across intervals,

$$w = \begin{pmatrix} w_1 & w_2 & \ldots & w_{k-1} \end{pmatrix}' = \tau_{n_1} \Delta \widehat{\beta}.$$

If all null hypotheses are true,

$$w = \tau_{n_1} \Delta \widehat{\beta} = \tau_{n_1} \Delta \left(\widehat{\beta} - \beta \right) \Rightarrow MN\left(0, \Delta C^{-2} V \Delta'\right).$$

The testing procedure is implemented in a stepwise manner. Order elements of w according to their absolute values, from largest to smallest, $|w_{r_1}| \geq |w_{r_2}| \geq \ldots \geq |w_{r_{k-1}}|$ (r_1 is the index of the largest test statistic and so on). For the first step, the ideal critical value is the $1-\alpha/2$ quantile of the sampling distribution of

[15] If we assume the slightly weaker $\Phi = \tau_{n_1}\left(C + o(1)\right)$, the critical values described below should be calculated using Φ/τ_{n_1} instead of C, and the first-order properties of the testing procedure remain exactly the same.

$\max_j |w_j|$,

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

Since we can estimate the joint distribution of $w_1, \ldots, w_{k-1}$, the estimate of the above, $\hat{c}_1$, can be obtained by simulation. The test procedure is then to reject those null hypotheses, for which the individual confidence interval $[w_{r_s} \pm \hat{c}_1]$ does not contain zero.

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

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

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

$$c_j = c_j(1-\alpha) = \inf\left\{x : P\left\{\max_{R_j+1\le s\le k-1} |w_{r_s}| \le x\right\} \ge 1 - \frac{\alpha}{2}\right\},$$

where R_{j-1} is the number of hypotheses rejected in the first $j-1$ steps ($R_0 = 0$). The procedure is continued until no new hypotheses can be rejected. While detailed analysis of the data is deferred to Section 7, the reader can see the importance of subsequent steps in real data in Table 8.

For implementation we standardize test statistics. This is advocated by Romano and Wolf (2005) to distribute the p-values more evenly across the hypotheses. The standardization is not reflected in above description to facilitate notation.[16]

6 Simulation Study

This section has two objectives. First, we verify the finite sample properties of the multivariate subsampling with irregular and asynchronous data. Second, we verify the size properties of the test for constant betas. In particular, we replicate with simulated data the setting of Table 6 of the empirical section. As in the empirical section, length of intervals $[i-1, i)$ is taken to be one week.

For the first objective, we simulate data for one week. Efficient log-price of each of the two stocks follows a univariate Heston (1993) model:

$$\begin{aligned} dX_t^{(i)} &= \left(\alpha_1 - v_t^{(i)}/2\right) dt + \sigma_t^{(i)} dW_t^{(i)} \\ dv_t^{(i)} &= \alpha_2 \left(\alpha_3 - v_t^{(i)}\right) dt + \alpha_4 \left(v_t^{(i)}\right)^{1/2} dB_t^{(i)}, \ i = 1, 2 \end{aligned}$$

[16] In other words, if we denote by $\mathcal{S}$ the diagonal matrix containing the diagonal elements of $\Delta C^{-2} V \Delta'$, we work with test statistics $\widehat{\mathcal{S}}^{-1/2} w$ instead of w.

	$\varrho = 0$	$\varrho = 0.25$	$\varrho = 0.5$	$\varrho = 0.75$
$\xi^2 = 0$	0.943	0.926	0.932	0.914
$\xi^2 = 0.001$	0.939	0.927	0.925	0.924
$\xi^2 = 0.01$	0.938	0.925	0.921	0.922

Table 1: *Empirical coverage probabilities of* 95% *confidence interval for beta, where beta is estimated using Two Scale estimators as in (6). Confidence intervals of the Two Scale estimator are calculated using the plug-in approach with asymptotic variance estimated by subsampling.* ξ^2 *is the noise to signal ratio.* ϱ *controls the realizations of beta, see (19). Number of simulations is 2000.*

where $v_t^{(i)} = \left(\sigma_t^{(i)}\right)^2$, $W_t^{(i)}$ and $B_t^{(i)}$ are independent Brownian Motions. The parameters of the univariate efficient log-price process are chosen to be the same as in Zhang et al. (2005). They are $\alpha_1 = 0.05$, $\alpha_2 = 5$, $\alpha_3 = 0.04$, and $\alpha_4 = 0.5$ (the same for $i = 1, 2$). Correlation of the two processes is obtained by setting $Corr\left(W_t^{(1)}, W_t^{(2)}\right) = \varrho$, with ϱ taking values 0, 0.25 , 0.50, and 0.75 across experiments. In this model, the beta for the i^{th} period is

$$\beta_i = \varrho \int_{i-1}^{i} \sigma_u^{(1)}\sigma_u^{(2)} du \Bigg/ \int_{i-1}^{i} \left(\sigma_u^{(1)}\right)^2 du. \tag{19}$$

Microstructure noise is simulated as a normally distributed white noise with variance $\xi^2 IQ^{(1)}$, where ξ^2 is a noise-to signal ratio taking values 0, 0.001, and 0.01, and $IQ^{(1)}$ is the weekly integrated quarticity of the first stock (approximated as a Riemann sum of simulated 1 second values of σ_t^2). Since volatility paths are different across simulations, noise variance also varies across simulations and increases with higher volatility of the efficient price. Observed prices are efficient log-prices plus noise.

Asynchronous data is simulated as follows. As a first step, we simulate one week of 1 second synchronous observations (simulation is done via an Euler scheme with one year as a unit of time and one second step length). From these, we take 35,000 irregular and asynchronous observations for each stock as follows. We draw a random permutation of all observation times in a week, take the first 35,000 of them, and sort them. Observation times are independent across stocks. Observations are then synchronized using the Refresh Time method, resulting in a random number of observations (usually somewhere around 25,000).

The Two Scale estimator is implemented with exactly the same parameters as in the empirical analysis (Section 7). In other words, G_1 is taken so as to correspond to 5 minutes on average (typically around 70), and $G_2 = 3$ (see Section 3.2 for the meaning of these parameters). Subsampling parameters are also taken as in section 7, i.e., $J = 5G_1$ and $m = 20G_1$.

Table 1 shows the resulting coverage probabilities for betas. There is some undercoverage, but otherwise subsampling method seems to work reasonably well in finite samples.

For the second objective, we replicate the above setup for each week 52 times. To investigate the size of the test we need to simulate a process that satisfies the null hypothesis of constant betas. The simplest way to achieve this is to keep the volatility path the same across weeks. The Brownian Motion driving the prices, market microstructure noises, and the trading process is simulated anew for each week.

Table 4 in the appendix shows the resulting empirical rejection probabilities with 5% level of significance of the test of constant betas (described in Section 5.1). Simulation designs include different noise to signal ratios $\xi^2 = 0, 0.001$, and 0.01, different $\varrho = 0, 0.25, 0.50$ and 0.75, and different number of weeks, mimicking the test with real data within quarters and a year of the Table 6. Both high frequency and moderate frequency methods perform reasonably well, although there is some overrejection for the high frequency method and some underrejection for the moderate frequency methods.

7 Empirical Analysis

This section implements above methods with real data. For both moderate and high frequency data, we first implement the test of constant betas, then search for the exact time of breaks nonparametrically. All figures and tables of this section, with the exception of Table 2, can be found in the appendix.

To implement these testing procedures, we need estimated variances of $\widehat{\beta}^{RV}$ and $\widehat{\beta}^{AMZ}$. To this end, we use Barndorff-Nielsen and Shephard (2004) estimator of the variance-covariance matrix of $\widehat{\beta}^{RV}$, and the estimator of variance-covariance matrix of $\widehat{\beta}^{AMZ}$ obtained by subsampling as in Section 4.2. We calculate $\widehat{\beta}^{AMZ}$ using all data, but for $\widehat{\beta}^{RV}$ we need to choose some lower sampling frequency. We choose three frequencies that are popular in practice, 5, 15, and 20 minutes, and denote the resulting estimators as $\widehat{\beta}^{RV}_{5min}$, $\widehat{\beta}^{RV}_{15min}$, and $\widehat{\beta}^{RV}_{20min}$. In all that follows, length of intervals $[i-1, i)$ is taken to be one week.

7.1 The Data

We use high frequency transactions data on six individual stocks. They are American International Group, Inc. (listed under the ticker symbol AIG), General Electric Co. (GE), International Business Machines Co. (IBM), Intel Co. (INTC), Minnesota Mining and Manufacturing Co. (MMM), and Microsoft Co. (MSFT). To proxy for a market portfolio, we use Standard and Poors Depositary Receipts (SPIDERS for short, ticker symbol SPY), which are an Exchange Traded Fund set up to mimic the movements of the Standard and Poor's 500 Composite Stock Price Index. Our data covers the whole year 2006 and is obtained from the NYSE Trade and Quote database.

We clean the data as follows. We apply time filter 9:30 - 16:00. We retain all satellite markets. Where multiple transactions per second are recorded, we take the first one. Where repeated times are recorded, we take the average. Next, we delete bounce backs, jumps, as well as gradual jumps as follows. Bounce backs are most likely to result from data mistakes, such as incorrect time record, so as a first step we identify bounce backs among prices and delete them. We define bounce backs as two consecutive price changes of the opposite sign, where each of the two price changes is larger, in absolute value, than five standard deviations of the observed price over a moving window of 500 transactions. Next, we remove jumps using the thresholding methodology of Mancini (2004). In other words, we set those returns that are larger than some threshold to zero. The threshold for this purpose is defined as five standard deviations of the observed price, and is calculated over a moving window of one day. Finally, we remove gradual jumps. Barndorff-Nielsen et al. (2008b) discuss the fact that Realized Kernels do not behave well when price only rises (falls) over some period of time. Two Scale estimator is similarly not robust to gradual jumps, so we also have to deal with them. Barndorff-Nielsen et al. (2008b) define gradual jumps as relatively long periods containing only price increases or only price decreases. They then replace the returns of this period with one single jump. We define gradual jumps as at least 5 minutes long interval containing only price increases (or decreases), provided the total price change exceeds a threshold of five standard deviations of the observed returns. Gradual jumps are replaced with a zero return. The threshold is recalculated over a moving window of 5 days. All window lengths mentioned in the cleaning procedure are average ones; windows are fixed in terms of number of transactions so as to achieve the target calendar time period on average over the year.

In order to calculate realized betas, we need to synchronize the data (see Section 3.2). When constructing tests on individual assets, we synchronize data in pairs to maximize the information used. For example, when testing time variability of beta of INTC we only require to synchronize INTC with SPY; to estimate beta of MMM, we synchronize MMM with SPY. Therefore, different transformations of the original SPY data is used to calculate different betas. When implementing the joint tests we synchronize data across all assets. The next section discusses the effect of these choices.

7.2 Results

We start by analyzing the high frequency data. Table 3 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 bid-ask bounce. In fact, if we removed all zero returns, the remaining data would display very persistent autocorrelation with alternating signs (see figure with autocorrelations in Kalnina (2011), this has been also noted in e.g. Griffin and Oomen 2005). In full data set with zero returns, this effect is attenuated because switching times of bid and ask are random. Third autocorrelations are of different signs and small. The estimates of the noise variance (columns 2 and 3 in Table 3) are very small, and in fact several orders of magnitude smaller than Hansen and Lunde (2006) estimates for year 2004. For example, the simplest estimator of the noise variance is

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

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

Table 5 contains the same summary statistics for the cleaned data. As long as there is any asynchronicity in the observations, number of synchronized observations is always smaller. We can see the reduction of the transactions per week by comparing first columns of Table 3 and 5. Joint synchronization across assets gives the smallest number of synchronized observations (last 7 rows of Table 5). Noise variances are larger as measured by $\widehat{\omega}^2$, but we can easily verify this is purely due to larger finite-sample bias caused by smaller number of observations. In particular, the bias-adjusted estimators

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

are the same with and without synchronization. Autocorrelations are smaller, which is due to frequency being lower.

Figure 2 contains volatility signature plots for each individual stock (plots of realized variance against the frequency used in its calculation), as well correlation signature plots (plots of realized correlation against the frequency).[17] Volatility signature plots show a large increase for highest frequencies, consistent with the additive noise model where bias explodes as we sample more and more frequently. On the other hand, realized covariances display the so-called Epps effect due to Epps (1979), i.e., they tend to zero as the frequency increases, so that the realized correlations are also driven to zero.[18] Not surprisingly, realized beta signature plots in Figure 3 show a clear bias towards zero for highest frequencies. Therefore, neither Realized Variance, nor Realized Covariance should be calculated using the highest frequencies. On the other hand, the Two Scale estimator, while using all the synchronized data, cancels both the effect of noise and asynchronous observations and is consistent (see Zhang, 2011).

[17] Realized correlation is defined as

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

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

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

Figures 4 - 6 show plots of estimated betas using $\widehat{\beta}^{RV}_{5min}$ and $\widehat{\beta}^{AMZ}$ together with 95% confidence intervals, which are based on equation (4) and subsampling, respectively. In fact, similar series of confidence intervals for $\widehat{\beta}^{RV}_{5min}$ was also graphed by Andersen et al. (2004) in their Figures 13 - 15, except they used 15 minute and daily data to calculate estimated betas over intervals of one quarter. They did not however formally test variability of betas across time or tried to adjust for the multiple testing that is implicit in such graphs. The emphasis of their paper was parametric modelling of betas over time. In figures 4 - 6, we see that beta is estimated much more precisely using full record transaction prices. The two parameters in $\widehat{\beta}^{AMZ}$ were chosen as follows. G_1 was set to the number of ticks as to correspond to 5 minutes on average. G_2 was set to 3, given that there is no evidence of autocorrelations at larger lags. The two parameters of the subsampling scheme were set to $m = 20G_1$ and $J = 5G_1$. Estimates and confidence intervals for $\widehat{\beta}^{RV}_{15min}$ and $\widehat{\beta}^{RV}_{20min}$ are not shown, but they have much longer confidence intervals than $\widehat{\beta}^{RV}_{5min}$.

Table 6 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 14. Four different tests are implemented based on four estimators: $\widehat{\beta}^{RV}_{5min}, \widehat{\beta}^{RV}_{15min}, \widehat{\beta}^{RV}_{20min}$ and $\widehat{\beta}^{AMZ}$. The reader should be careful when interpreting the p-values since at this stage they are not adjusted to reflect multiple testing. However, a general idea of results can be anticipated by looking at the figures with point estimates and their confidence intervals. 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 $\widehat{\beta}^{AMZ}$ can reject the null, at 5% level of significance, for any of scenarios considered, except it has a p-value of 0.057 for GE Q1. The test based on $\widehat{\beta}^{RV}_{5min}$ rejects the null for fewer cases. The test based on $\widehat{\beta}^{RV}_{15min}$ fails to reject the null for roughly half of quarters-based cases, and the test based on $\widehat{\beta}^{RV}_{20min}$ fails to reject the null for most of quarters-based cases.

Table 7 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; for asynchronous data, the increase of number of observations is smaller due to data loss arising from synchronization of more stocks). Indeed, we see that even the moderate frequency estimators now reject most null hypotheses.

The benefit of using more data is mitigated to some extent by the data loss due to synchronising more stocks than before. Overall, we see that the first effect dominates, with largest differences in 5 minute data. Unsurprisingly, all hypotheses are again rejected with 1 tick data.

When considering the whole year 2006, the results across frequencies are similar in the sense that we can reject the null hypothesis of constant betas in all cases. The test statistics are however much higher for the 1 tick data. This difference becomes crucial when we attempt to determine the exact time of the break. Table 2 below shows the number of breaks found among 52 weeks in 2006 (note that highest possible number of breaks is 51).

		AIG	*GE*	*IBM*	*INTC*	*MMM*	*MSFT*
	α = *0.01*	0	0	0	1	0	7
5 min	α = *0.05*	0	0	0	2	0	8
	α = *0.10*	0	0	0	2	0	10
	α = *0.01*	0	0	0	0	0	1
15 min	α = *0.05*	0	0	0	1	0	2
	α = *0.10*	0	0	0	1	0	2
	α = *0.01*	0	0	0	0	0	1
20 min	α = *0.05*	0	0	0	1	0	3
	α = *0.10*	0	0	0	1	0	3
	α = *0.01*	48	46	48	49	45	49
1 tick	α = *0.05*	50	48	48	49	47	50
	α = *0.10*	50	48	51	49	47	51

Table 2: *Number of detected breaks in weekly betas during 2006, i.e., number of rejected null hypotheses out of 51. See section* 5.3 *for exact description of the test.* α *is the level of significance.*

Differences across frequencies are striking. We can see that tick data reveals significant breaks in beta for virtually every week, while moderate frequency data does not find any for most of the stocks. In similar circumstances when no null hypotheses can be rejected after accounting for multiple testing by controlling FWE, a strategy that can be considered is to change the definition of the size of the test, so as to make the test more lenient to the possibility of rejecting true null hypotheses. One prominent example is control of False Discovery Rate (see Benjamini and Hockberg, 1995). With probability α (the size of the test), it allows, on average, some fixed percentage such as 10% of the total number of hypotheses to be falsely rejected. Therefore, it rejects more hypotheses by construction and is more powerful against false hypotheses. However, even one break in betas is important for applications, so this approach cannot be justified in our setting. Fortunately, we see that with the amount of precision tick data gives one can determine breaks in betas without a recourse to more liberal definitions of the size of the test.

Additional stages of the test increase finite sample power of the test while not changing its asymptotic consistency and size. Table 8 in the appendix gives a more detailed view by breaking down the result of Table 2 at every stage of the test. We see that additional steps indeed sometimes find additional breaks, though usually very few.

We conclude that while moderate frequency data can give evidence about a break in beta taking place somewhere in a period of time such as a year, it cannot in general detect the exact time of the break nonparametrically. On the other hand, due to the much higher precision, the nonparametric search for breaks with high frequency data can detect many breaks.

8 Conclusion

This paper studies the question of time variability in equity betas. Recent developments in high frequency econometrics allows us to estimate quadratic variation version of the betas in a model-free framework with ultra high frequency and asynchronously observed data from NYSE. Due to the market microstructure noise in this data, estimators of beta can have complicated expressions of the asymptotic variance, in which case it is convenient to use an automatic inference method to implement tests on betas. We show how the multivariate version of the subsampling method of Kalnina (2011) can be used to achieve this. We demonstrate the robustness of the inference method of Kalnina (2011) to long memory and extend it to the multivariate case. Using this inference method, we construct tests for no time variation in betas. Next, we build on these to arrive at a nonparametric procedure for search for breaks in betas by treating every possible time of break in betas as a hypothesis and accounting for multiple testing. We implement these methods with six stocks on the NYSE over year 2006 with Standard and Poors Depositary Receipts as a proxy for the market factor. We find the strongest advantage of the high frequency data when searching for breaks in beta. In particular, we find that moderate frequency data does not have enough power to identify a single break in betas in 2006 for most of the stocks, while highest frequency data reveals breaks in most weeks in 2006 for each stock considered. An interesting application that we leave for future research is implementation of the nonparmetric search for breaks over a longer period of time and then relating these to the dynamics of macroeconomic fundamentals in a second step.

References

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

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

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

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

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

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

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

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

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

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

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

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

[13] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and N. Shephard (2008b). Realised Kernels in Practice: Trades and Quotes. Econometrics Journal, forthcoming.

[14] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and N. Shephard (2010). Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-synchronous Trading. Journal of Econometrics, forthcoming.

[15] Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and Bipower Variation with Stochastic volatility and Jumps. Journal of Financial Econometrics, Vol. 2, No. 1, 1-37.

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

[17] Benjamini, Y. and Y. Hochberg, (2005). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Teting. J. R. Statist. Soc. B 57, No. 1, pp. 289-300.

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

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

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

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

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

[23] Comte, F. (2005). Long range dependence in financial markets. Fractals in engineering, 4, 159-179.

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

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

[26] Donovon, P., S. Gonçalves, and N. Meddahi (2008). Bootstrapping realized multivariate volatility measures. Journal of Econometrics, forthcoming.

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

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

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

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

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

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

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

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

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

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

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

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

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

[40] Mancini, C. (2004). Estimation of the characteristics of the jumps of a general Poisson-diffusion model, Scandinavian Actuarial Journal 1, 42-52.

[41] Li, Y., Mykland, P.A., Renault, E., Zhang, L., and Zheng, X. (2009). Realized volatility when sampling times are possibly endogenous. Working paper.

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

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

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

[45] Patton, A. and M. Verardo (2011). Does beta move with news? Firm-specific information flows and learning about profitability. Working Paper.

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

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

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

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

[50] Rossi, E. and P. S. de Magistris (2011). Estimation of long memory in integrated variance. CREATES Research Paper 2011-11.

[51] Sizova, N. (2011). Long-Horizon Return Regressions with Historical Volatility. Working paper.

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

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

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

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

[56] ZHANG, L. (2011). Estimating Covariation: Epps Effect, Microstructure Noise. Journal of Econometrics 160, 33-47.

[57] ZHANG, L., P. MYKLAND, AND Y. AÏT-SAHALIA (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association, 100, 1394–1411.

[58] ZU, Y., AND P. BOSWIJK (2009). Estimating Realized Spot Volatility with Noisy High-Frequency Data. Working Paper.

A Proofs

A.1 Proof of Proposition 1

In what follows we use C to denote a generic constant, the meaning of which changes from line to line. We start with the following lemma, see Section A.2 for a proof.

Lemma 3. *Suppose Assumption A2 holds. Then for any $k \in \mathbb{N}$ and $\alpha \in (0, 0.5)$ we have*

$$E\left[\left|\sigma^k(t_2) - \sigma^k(t_1)\right|\right] \leq C|t_2 - t_1|^{\alpha/2}.$$

To prove Proposition 1, we first introduce some notation (same as in Kalnina, 2011),

$$\begin{aligned}
V &= \frac{4}{3}c\int_0^1 \sigma_u^4 du + 8c^{-2}\mathrm{Var}\left(\epsilon\right)^2 + 16c^{-2}\lim_{n\to\infty}\sum_{i=1}^{n}\mathrm{Cov}\left(\epsilon_0, \epsilon_{i/n}\right)^2 \\
V_l^{short} &= \frac{4}{3}c\int_{(l-1)m/n}^{[(l-1)m+J]/n} \sigma_u^4 du + 8c^{-2}\frac{J}{n}\mathrm{Var}\left(\epsilon\right)^2 + 16\frac{J}{n}c^{-2}\lim_{n\to\infty}\sum_{i=1}^{n}\mathrm{Cov}\left(\epsilon_0, \epsilon_{i/n}\right)^2 \\
\theta_l^{long} &= \int_{(l-1)m/n}^{m/n} \sigma_u^2 du, \qquad \theta_l^{short} = \int_{(l-1)m/n}^{[(l-1)m+J]/n} \sigma_u^2 du,
\end{aligned}$$

and V_l the same as V_l^{short} except with m instead of J.

Note that we are using the non-overlapping version of the estimator (i.e., $s=m$) and hence the number of subsamples is $K=\lceil \frac{n}{m} \rceil$. In the proof of Theorem 4 of Kalnina (2011), Assumption A1 is used for two statements[19]:

$$V-\sum_{l=1}^{K}\frac{m}{J}V_l^{short}=o_p(1) \tag{20}$$

and

$$\frac{m}{J}\sum_{j=1}^{K}n^{1/3}\left(\theta_l^{short}-\frac{J}{m}\theta_l^{long}\right)^2=o_p(1). \tag{21}$$

We now show they both remain true under Assumption A2 instead. To prove equation (20), notice that

$$\begin{aligned}
\mathrm{E}\left|V_l-\frac{m}{J}V_l^{short}\right| &= \frac{4c}{3}\mathrm{E}\left|\int_{(l-1)m/n}^{lm/n}\sigma_u^4du-\frac{m}{J}\int_{(l-1)m/n}^{[(l-1)m+J]/n}\sigma_u^4du\right| \\
&\leq \frac{4c}{3}\mathrm{E}\left|\int_{(l-1)m/n}^{lm/n}\left(\sigma_u^4-\sigma_{(l-1)m/n}^4\right)du\right|+\frac{4c}{3}\mathrm{E}\left|\frac{m}{J}\int_{(l-1)m/n}^{[(l-1)m+J]/n}\left(\sigma_u^4-\sigma_{(l-1)m/n}^4\right)du\right| \\
&= \frac{4c}{3}\mathrm{E}\left|\int_{(l-1)m/n}^{lm/n}\left(\sigma_u^4-\sigma_{\lfloor Ku\rfloor/K}^4\right)du\right|+\frac{4c}{3}\mathrm{E}\left|\frac{m}{J}\int_{(l-1)m/n}^{[(l-1)m+J]/n}\left(\sigma_u^4-\sigma_{\lfloor Ku\rfloor/K}^4\right)du\right| \\
&= \frac{4}{3}c\mathrm{E}\left|\frac{m}{n}s_l\right|+\mathrm{E}\left|\frac{m}{n}s_l'\right| \\
&= o(m/n)
\end{aligned}$$

where s_l and s_l' with

$$\inf_{\frac{(l-1)m}{n}\leq u\leq\frac{lm}{n}}\left(\sigma_u^2-\sigma_{\frac{\lfloor Ku\rfloor}{K}}^2\right)\leq s_l\leq\sup_{\frac{(l-1)m}{n}\leq u\leq\frac{lm}{n}}\left(\sigma_u^2-\sigma_{\frac{\lfloor Ku\rfloor}{K}}^2\right)$$

and

$$\inf_{\frac{(l-1)m}{n}\leq u\leq\frac{(l-1)m+J}{n}}\left(\sigma_u^2-\sigma_{\frac{\lfloor Ku\rfloor}{K}}^2\right)\leq s_l'\leq\sup_{\frac{(l-1)m}{n}\leq u\leq\frac{(l-1)m+J}{n}}\left(\sigma_u^2-\sigma_{\frac{\lfloor Ku\rfloor}{K}}^2\right)$$

for each sample path of σ, are random variables such that the third equality holds. Their existence for each sample path is guaranteed by Mean Value Theorem. Last equality follows because s_l and s_l' converge to zero a.s. by right-continuity and boundedness of σ. Notice that neither Assumption A1 nor A2 are needed here.

[19]Both these statements are from the main body of the proof of Theorem 4 of Kalnina (2011). Lemma 7 of Kalnina (2011) does not use Assumption A1.

To prove equation (21), we have

$$
\begin{aligned}
& \mathrm{E}\left(\theta_l^{short} - \frac{J}{m}\theta_l^{long}\right)^2 \\
= \; & \mathrm{E}\left(\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \sigma_u^2 du - \frac{J}{m}\int\limits_{(l-1)m/n}^{lm/n} \sigma_u^2 du\right)^2 \\
= \; & \mathrm{E}\left(\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \left(\sigma_u^2 - \sigma_{(l-1)m/n}^2\right) du - \frac{J}{m}\int\limits_{(l-1)m/n}^{lm/n} \left(\sigma_u^2 - \sigma_{(l-1)m/n}^2\right) du\right)^2 \\
\leq \; & 2\mathrm{E}\left(\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \left(\sigma_u^2 - \sigma_{(l-1)m/n}^2\right) du\right)^2 + 2\mathrm{E}\left(\frac{J}{m}\int\limits_{(l-1)m/n}^{lm/n} \left(\sigma_u^2 - \sigma_{(l-1)m/n}^2\right) du\right)^2 \\
= \; & 2\mathrm{E}\left(\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right) du\right)^2 + 2\mathrm{E}\left(\frac{J}{m}\int\limits_{(l-1)m/n}^{lm/n} \left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right) du\right)^2 \\
\leq \; & 2\mathrm{E}\left(\frac{J}{n}\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right)^2 du\right) + 2\mathrm{E}\left(\frac{J^2}{m^2}\frac{m}{n}\int\limits_{(l-1)m/n}^{lm/n} \left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right)^2 du\right) \\
= \; & 2\frac{J}{n}\int\limits_{(l-1)m/n}^{[(l-1)m+J]/n} \mathrm{E}\left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right)^2 du + 2\frac{J}{n}\frac{J}{m}\int\limits_{(l-1)m/n}^{lm/n} \mathrm{E}\left(\sigma_u^2 - \sigma_{\lfloor Ku\rfloor/K}^2\right)^2 du \\
= \; & O\left(\frac{J}{n}\frac{J}{n}\left(\frac{J}{n}\right)^\alpha\right) + O\left(\frac{J}{n}\frac{J}{m}\frac{m}{n}\left(\frac{m}{n}\right)^\alpha\right) \\
= \; & O\left(\frac{J^2}{n^2}\left(\frac{m}{n}\right)^\alpha\right)
\end{aligned}
$$

where 5th equality follows by Lemma 3. This (together with $K = \lceil \frac{n}{m} \rceil$) proves equation (21) as long as condition $Jm^\alpha n^{-2/3-\alpha} \to 0$ holds.

A.2 Proof of Lemma 3

Proof. If we prove

$$
E\left[\left(\sigma^k(t_2) - \sigma^k(t_1)\right)^2\right] \leq C|t_2 - t_1|^\alpha, \tag{22}
$$

Lemma 3 follows by the Cauchy-Schwarz inequality. For $k = 1$, equation (22) is a statement from the Appendix of Comte and Renault (1998), p. 314. The process $x(t) = \ln \sigma(t)$ can also be written as

$$\int_0^t a(t-s)dW(s)$$

with

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

and $W(s)$ a standard Brownian Motion, see Comte and Renault (1998). Let $t_1 \leq t_2$. We have $\sigma(t) = exp(x(t))$ and hence $\sigma^k(t) = exp(kx(t))$. Next,

$$\begin{aligned}
& E\left(\sigma^k(t_2) - \sigma^k(t_1)\right)^2 \\
= \; & E\left(\exp(kx(t_2)) - \exp(kx(t_1))\right)^2 \\
= \; & E\left(e^{2kx(t_2)} + e^{2kx(t_1)} - 2e^{kx(t_1)+kx(t_2)}\right) \\
= \; & e^{2k^2 \int_0^{t_1} a^2(x)dx} + e^{2k^2 \int_0^{t_2} a^2(x)dx} - 2e^{\frac{1}{2}k^2 \int_0^{t_1} a^2(x)dx + \frac{1}{2}k^2 \int_0^{t_2} a^2(x)dx + k^2 \int_0^{t_1} a(x)(a(t_2-t_1+x))dx} \\
= \; & e^{2k^2 \int_0^{t_2} a^2(x)dx} \left(1 + e^{-2k^2 \int_{t_1}^{t_2} a^2(x)dx} - 2e^{-\frac{3}{2}k^2 \int_0^{t_1} a^2(x)dx - k^2 \int_0^{t_1} a(x)(a(x)-a(t_2-t_1+x))dx}\right) \\
\leq \; & 2e^{2k^2 \int_0^{t_2} a^2(x)dx} \left(1 - e^{-\frac{3}{2}k^2 \int_{t_1}^{t_2} a^2(x)dx - k^2 \int_0^{t_1} a(x)(a(x)-a(t_2-t_1+x))dx}\right).
\end{aligned}$$

The term inside the last parenthesis being necessarily nonnegative, the term in the last great 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 α-Hölder,

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

which implies

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

Then using that $\forall u \leq 0, \; 0 \leq 1 - e^u \leq |u|$, equation (22) follows. This concludes the proof of Lemma 3.

A.3 Proof of Proposition 2

From $\tau_{n_i} = \tau_{n_1}(c_i + o(1))$, we have

$$\Phi = \tau_{n_1} diag\left(c_1 + o(1), c_2 + o(1), \ldots, c_k + o(1)\right).$$

Define $C = diag(1, c_2, c_3, \ldots, c_k)$. Equation (13) becomes

$$\tau_{n_1} diag\left(c_1 + o(1), c_2 + o(1), \ldots, c_k + o(1)\right) \left(\widehat{\beta} - \beta\right) \Rightarrow MN(0, V),$$

where convergence is stable in law. Thus,

$$\begin{aligned}
\tau_{n_1} C\left(\widehat{\beta}-\beta\right) &\Rightarrow MN(0, V) \\
\tau_{n_1}\left(\widehat{\beta}-\beta\right) &\Rightarrow MN(0, C^{-2} V) \\
\tau_{n_1} \Delta\left(\widehat{\beta}-\beta\right) &\Rightarrow MN(0, \Delta C^{-2} V \Delta') \\
\tau_{n_1}\left(\Delta C^{-2} V \Delta'\right)^{-1/2} \Delta\left(\widehat{\beta}-\beta\right) &\Rightarrow N(0, I_{k-1}) \\
\tau_{n_1}^2\left(\widehat{\beta}-\beta\right)'\left(\Delta C^{-2} V \Delta'\right)^{-1} \Delta\left(\widehat{\beta}-\beta\right) &\Rightarrow \mathcal{X}_{k-1}^2. \qquad (23)
\end{aligned}$$

Since $\Phi = \tau_{n_1} C$ plus smaller order terms, we obtain the first equation of the Proposition 2, without conditioning on the volatility path. Next, since the right-hand-side random variable $\mathcal{X}_{k-1}^2$ is independent of the volatility path and the convergence is stable in law, equation (23) also holds, conditional on the volatility path. Second equation of the Proposition 2 follows immediately from the first.

B Figures and Tables

	trans./week	$\widehat{\omega}^2 \cdot 10^7$	$\widetilde{\omega}^2 \cdot 10^7$	$\widehat{\xi}^2 \cdot 10^5$	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 3: *Summary statistics of data before synchronization. First column contains average number of transactions per week. Second and third columns contains variance of the noise estimates over the whole year 2006,* $\widehat{\omega}^2 = RV/2n$, $\widetilde{\omega}^2 = \left(RV - \widehat{IV}\right)/2n$ *where IV is estimated by the TSRV; n is the total number of transactions in 2006 for the corresponding stock. Fourth column contains estimated noise-to-signal ratio,* $\widehat{\xi}^2 = \widehat{\omega}^2/\widehat{IV}$. *Last three columns contain autocorrelation functions of returns at first, second, and third lag.*

	$\xi^2 = 0$					$\xi^2 = 0.001$					$\xi^2 = 0.01$				
	$Q1$	$Q2$	$Q3$	$Q4$	Y	$Q1$	$Q2$	$Q3$	$Q4$	Y	$Q1$	$Q2$	$Q3$	$Q4$	Y
$\varrho = 0$															
1 tick	6.30	6.30	6.85	7.25	7.70	6.00	6.80	6.95	6.65	5.80	6.25	5.95	5.35	5.80	6.05
5 min	1.90	1.85	1.95	2.60	2.15	2.05	2.25	3.00	2.85	2.55	2.85	2.55	2.35	2.40	2.70
15 min	1.95	2.80	2.45	2.65	2.60	2.80	3.30	2.30	2.70	2.30	2.80	2.45	2.05	2.20	2.95
20 min	2.30	2.30	3.10	2.30	2.90	2.20	3.00	2.75	2.70	2.40	1.75	2.60	2.80	3.805	2.55
$\varrho = 0.25$															
1 tick	6.40	6.15	6.75	7.45	6.95	6.50	6.55	6.60	6.15	6.55	6.40	5.75	6.35	6.75	5.90
5 min	2.65	2.20	2.20	2.95	2.25	2.25	2.00	2.05	1.80	1.60	2.95	2.30	2.40	1.85	2.95
15 min	2.20	2.85	2.25	2.40	2.70	2.55	2.60	2.80	2.70	2.75	2.40	2.05	2.60	2.40	2.75
20 min	2.15	2.75	2.60	2.60	2.85	2.70	2.20	2.65	2.60	2.95	2.25	2.00	3.15	2.60	3.00
$\varrho = 0.5$															
1 tick	6.15	7.75	6.35	6.25	7.60	6.50	6.55	6.60	6.15	6.55	6.40	5.75	6.35	6.75	5.90
5 min	2.70	2.00	2.10	2.05	2.10	2.25	2.00	2.05	1.80	1.60	2.95	2.30	2.40	1.85	2.95
15 min	2.20	2.20	2.30	1.90	2.85	2.55	2.60	2.80	2.70	2.75	2.40	2.05	2.60	2.40	2.75
20 min	2.00	2.35	2.00	2.55	2.75	2.70	2.20	2.65	2.60	2.95	2.25	2.00	3.15	2.60	3.00
$\rho = 0.75$															
1 tick	7.10	6.60	5.95	6.50	8.45	6.85	6.90	6.75	6.50	7.40	7.10	6.60	5.95	6.50	8.45
5 min	2.60	2.25	2.40	2.40	2.40	2.05	2.60	2.35	1.80	2.15	2.60	2.25	2.40	2.40	2.40
15 min	2.50	1.95	2.80	1.80	2.95	2.50	2.15	2.30	2.25	2.40	2.50	1.95	2.80	1.80	2.95
20 min	2.405	2.15	2.90	2.50	2.95	2.05	2.05	2.55	2.70	2.25	2.45	2.15	2.90	2.50	2.95

Table 4: *Empirical rejection probabilities with 5% level of significance. 2000 simulations.*

	trans./week	$\widehat{\omega}^2 \cdot 10^7$	$\widetilde{\omega}^2 \cdot 10^7$	$\widehat{\xi}^2 \cdot 10^5$	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 5: *Summary statistics of data after synchronization. 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 other 6 series. See Table 1 annotation for meaning of other column entries.*

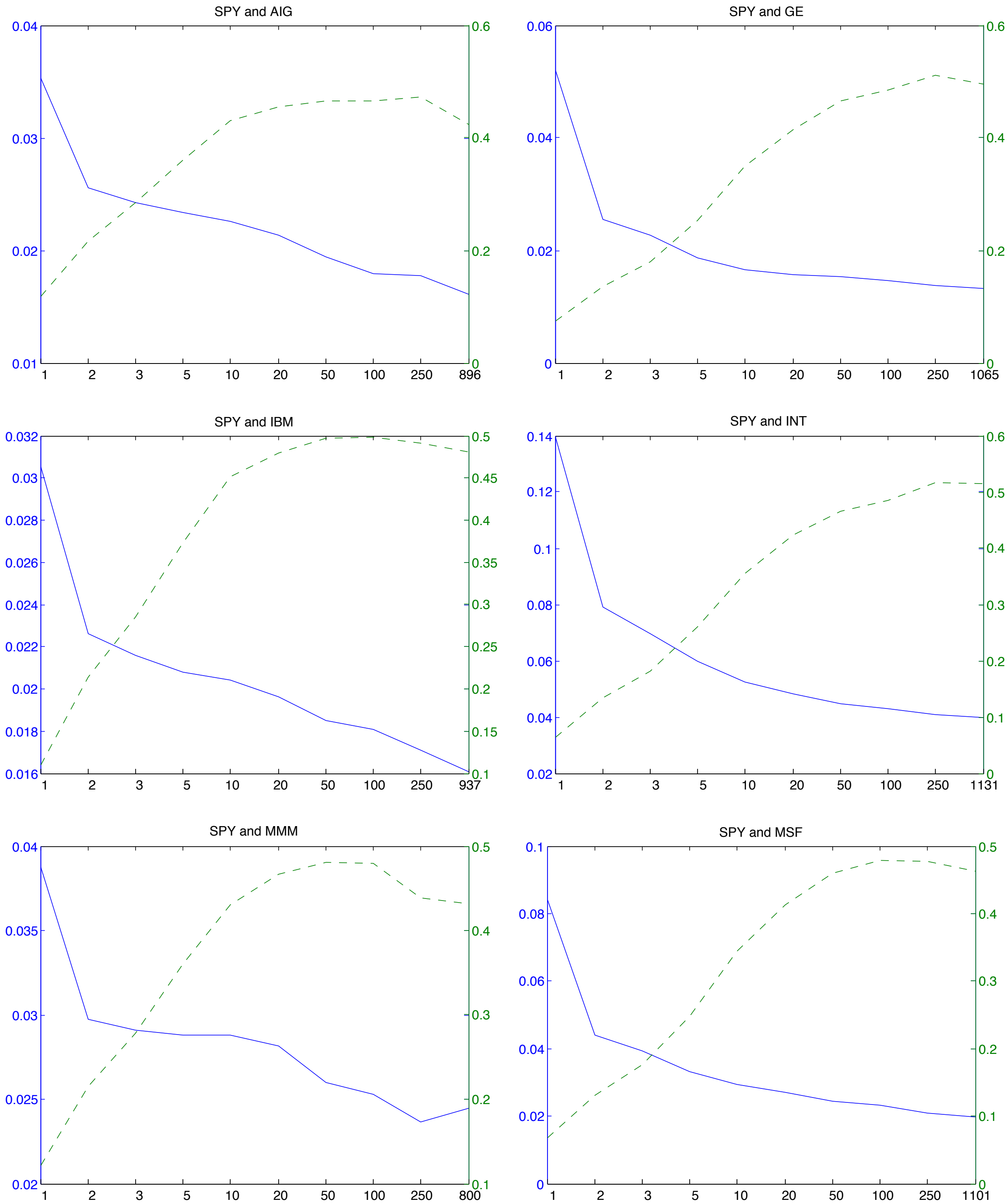


Figure 2: *The solid lines (left y-axis) are the volatility signature plots, i.e., realized variance plotted against the frequency (in ticks) used in its calculation. Dashed lines (right y-axis) are the realized correlation plots against the frequency (in ticks). Data covers the whole year 2006.*

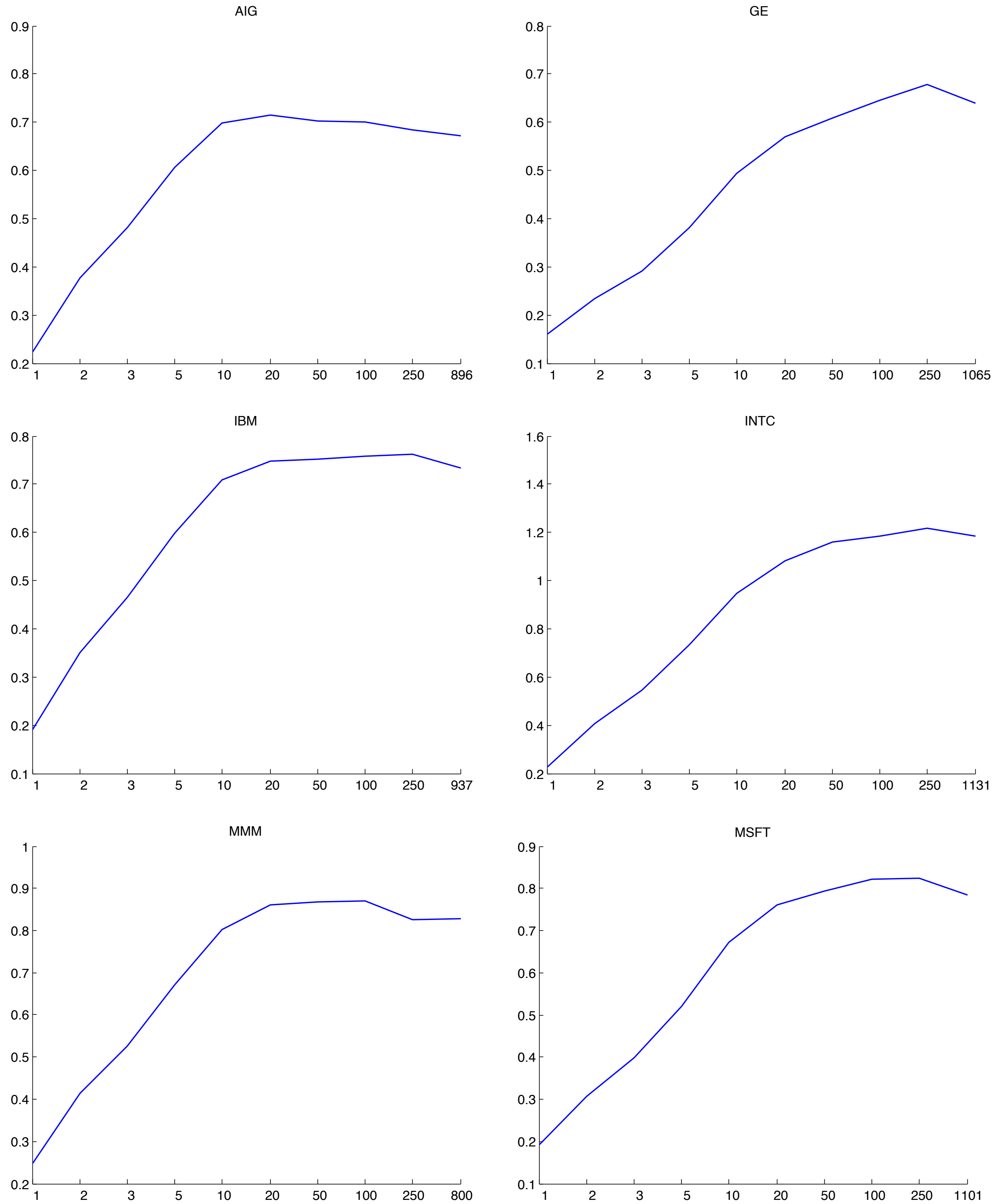


Figure 3: *Average weekly betas of individual stocks against the frequency (in ticks) used in their calculation. Data year 2006.*

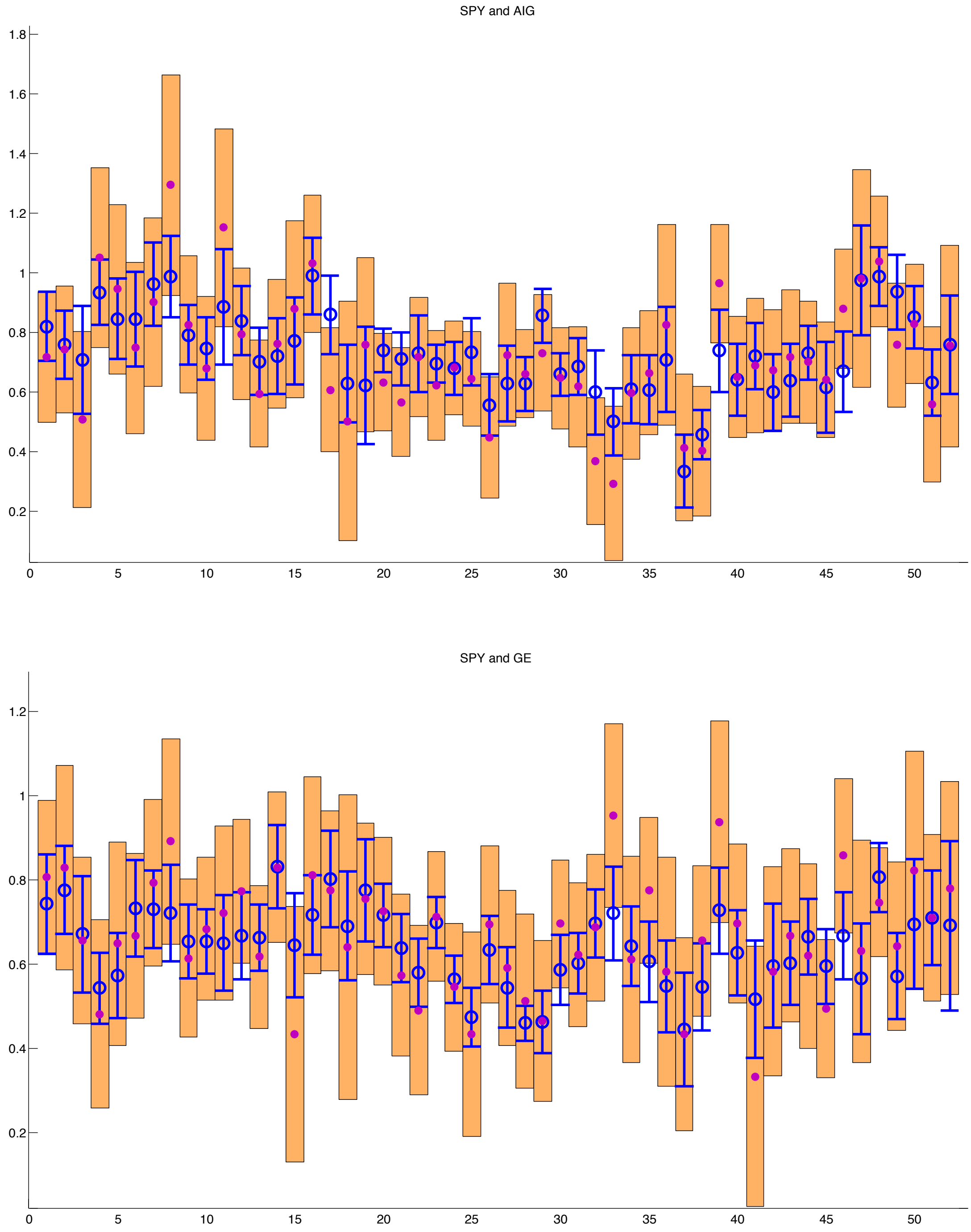
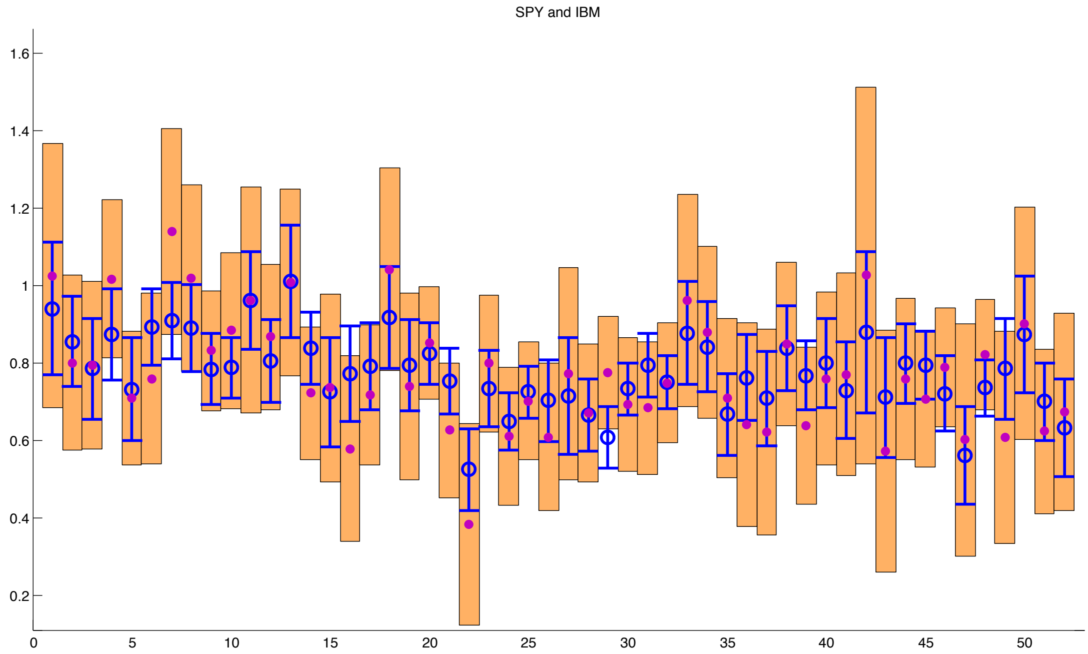


Figure 4: *Estimated betas for AIG and GE with 95% confidence intervals. Filled dots with rectangular CIs correspond to* $\widehat{\beta}^{RV}_{5min}$ *, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{AMZ}$ *. Weeks on the x-axis.*



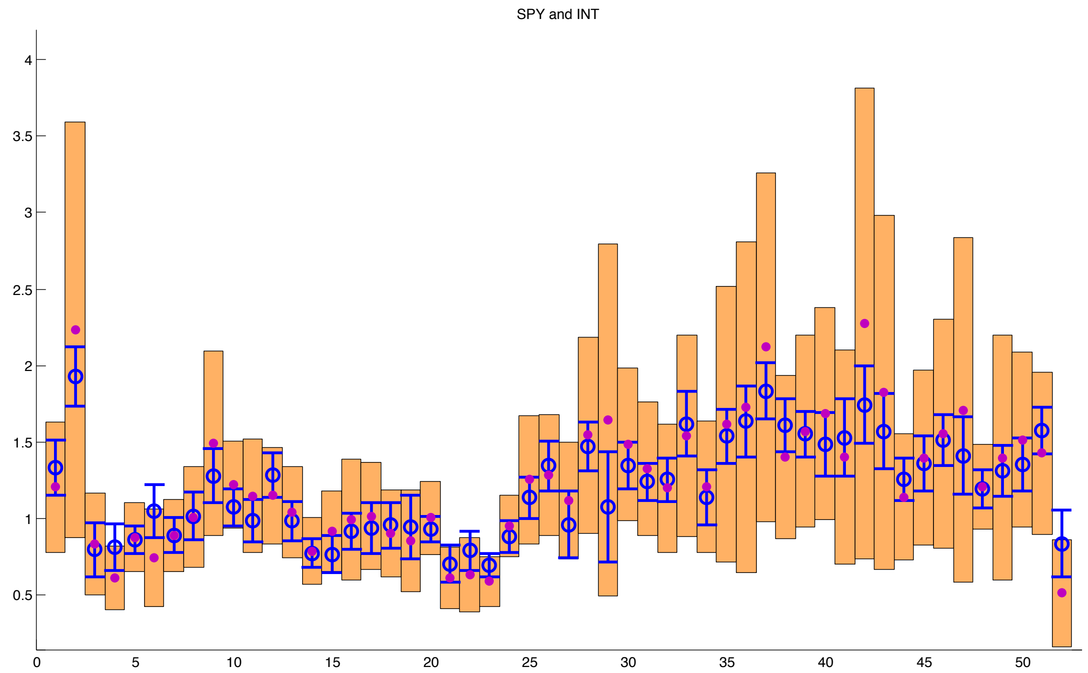
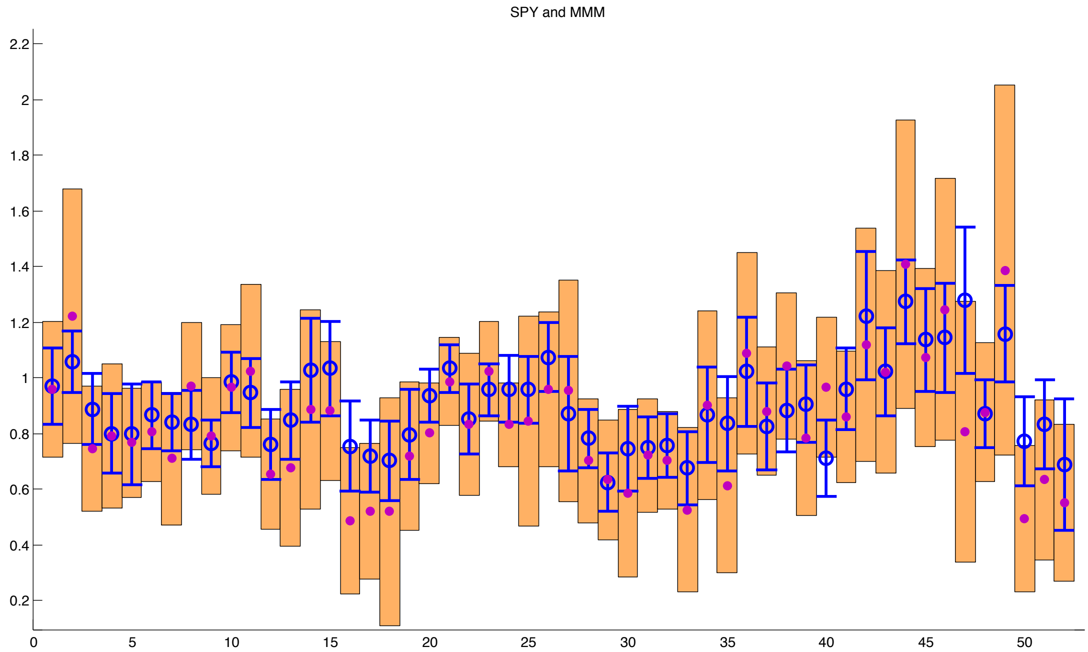


Figure 5: *Estimated betas for IBM and INTC with 95% confidence intervals. Filled dots with rectangular CIs correspond to* $\widehat{\beta}^{RV}_{5min}$ *, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{AMZ}$ *. Weeks on the x-axis.*



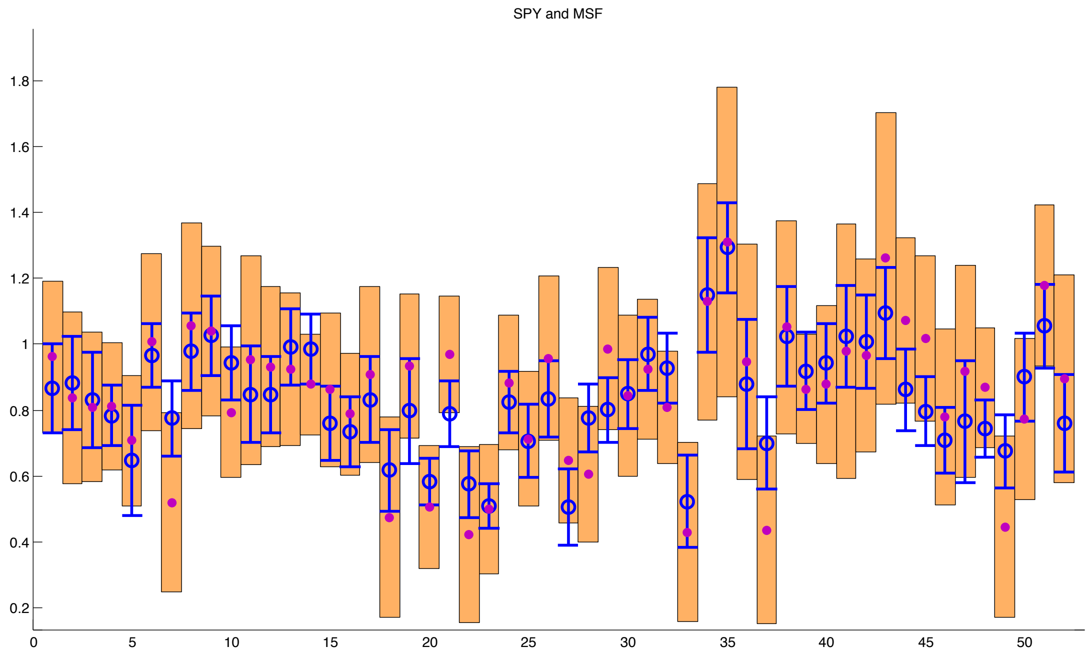


Figure 6: *Estimated betas for MMM and MSFT with* 95% *confidence intervals. Filled dots with rectangular CIs correspond to* $\widehat{\beta}^{RV}_{5min}$ *, empty dots with error-bar-type CIs correspond to* $\widehat{\beta}^{AMZ}$ *. Weeks on the x-axis.*

	2006		Q1		Q2		Q3		Q4	
Test based on $\widehat{\beta}^{RV}_{5min}$										
AIG	196.50	(0)	23.47	(0.024)	35.63	(0)	52.83	(0)	56.28	(0)
GE	101.03	(0)	15.26	(0.228)	15.55	(0.213)	27.75	(0.006)	26.00	(0.011)
IBM	110.82	(0)	14.58	(0.265)	30.02	(0.003)	26.59	(0.009)	17.41	(0.135)
INTC	248.96	(0)	33.84	(0.001)	40.96	(0)	36.15	(0)	43.65	(0)
MMM	132.45	(0)	27.48	(0.007)	25.10	(0.014)	21.83	(0.039)	37.48	(0)
MSFT	213.30	(0)	22.74	(0.03)	71.43	(0)	64.97	(0)	42.23	(0)
Test based on $\widehat{\beta}^{RV}_{15min}$										
AIG	112.70	(0)	26.03	(0.011)	19.99	(0.067)	33.69	(0.001)	15.78	(0.202)
GE	75.14	(0.012)	12.23	(0.427)	19.81	(0.071)	25.06	(0.015)	17.17	(0.143)
IBM	73.48	(0.017)	15.09	(0.236)	21.33	(0.046)	8.94	(0.708)	8.57	(0.739)
INTC	134.91	(0)	27.62	(0.006)	30.37	(0.002)	6.62	(0.882)	24.29	(0.019)
MMM	85.34	(0.001)	14.52	(0.269)	25.03	(0.015)	16.88	(0.154)	28.24	(0.005)
MSFT	121.52	(0)	15.49	(0.216)	39.48	(0)	34.18	(0.001)	22.80	(0.029)
Test based on $\widehat{\beta}^{RV}_{20min}$										
AIG	74.54	(0.014)	24.37	(0.018)	9.43	(0.666)	9.19	(0.687)	22.80	(0.03)
GE	85.89	(0.001)	16.26	(0.18)	15.16	(0.233)	17.38	(0.136)	22.89	(0.029)
IBM	68.38	(0.043)	10.93	(0.535)	17.80	(0.122)	11.85	(0.458)	7.36	(0.833)
INTC	107.53	(0)	20.81	(0.053)	29.48	(0.003)	3.36	(0.992)	19.16	(0.085)
MMM	73.00	(0.019)	14.54	(0.268)	14.30	(0.282)	16.99	(0.15)	20.90	(0.052)
MSFT	87.09	(0.001)	13.61	(0.326)	19.62	(0.075)	28.56	(0.005)	15.33	(0.224)
Test based on $\widehat{\beta}^{AMZ}$										
AIG	269.04	(0)	22.44	(0.033)	34.97	(0)	71.28	(0)	57.31	(0)
GE	224.40	(0)	20.56	(0.057)	65.13	(0)	63.42	(0)	26.08	(0.01)
IBM	136.70	(0)	18.93	(0.09)	37.52	(0)	26.76	(0.008)	18.96	(0.09)
INTC	845.32	(0)	143.45	(0)	83.10	(0)	79.34	(0)	55.97	(0)
MMM	200.38	(0)	28.75	(0.004)	40.19	(0)	21.96	(0.038)	63.09	(0)
MSFT	403.21	(0)	30.24	(0.003)	96.47	(0)	124.55	(0)	60.15	(0)

Table 6: *Values of the Chi-square test; corresponding p-values in parenthesis. The null hypothesis is that true betas are constant over the some time interval. The top row indicates the corresponding time interval.*

	$\widehat{\beta}^{RV}_{5min}$		$\widehat{\beta}^{RV}_{15min}$		$\widehat{\beta}^{RV}_{20min}$		$\widehat{\beta}^{AMZ}$	
2006	937.6	(0)	510.3	(0)	545.2	(0)	3530.7	(0)
Q1	107.2	(0.005)	64.3	(0.729)	85.2	(0.138)	458.4	(0)
Q2	147.1	(0)	96.1	(0.03)	106.9	(0.005)	510.4	(0)
Q3	296.8	(0)	168.9	(0)	140.6	(0)	745.7	(0)
Q4	172.4	(0)	103	(0.01)	154.7	(0)	741.1	(0)

Table 7: *Values of the joint Chi-square test (see section 5.2); corresponding p-values in parenthesis. The null hypothesis is that true betas for all 6 stocks are constant over the some time interval. The first row indicates the corresponding time interval. First three methods (labelled $\widehat{\beta}^{RV}_{5min}$, $\widehat{\beta}^{RV}_{15min}$, and $\widehat{\beta}^{RV}_{20min}$) are based on noise-free theory described in section 3.1; for the last method, TSRV method is used for point estimates of betas, and subsampling method is used to estimate their asymptotic variance-covariance matrix.*

		$\alpha = 0.01$			$\alpha = 0.05$			$\alpha = 0.10$		
		Step 1	*Step 2*	*Step 3*	*Step 1*	*Step 2*	*Step 3*	*Step 1*	*Step 2*	*Step 3*
AIG										
	5 min	0	0	0	0	0	0	0	0	0
	15 min	0	0	0	0	0	0	0	0	0
	20 min	0	0	0	0	0	0	0	0	0
	1 tick	47	1	0	48	1	1	48	2	0
GE										
	5 min	0	0	0	0	0	0	0	0	0
	15 min	0	0	0	0	0	0	0	0	0
	20 min	0	0	0	0	0	0	0	0	0
	1 tick	46	0	0	46	2	0	46	2	0
IBM										
	5 min	0	0	0	0	0	0	0	0	0
	15 min	0	0	0	0	0	0	0	0	0
	20 min	0	0	0	0	0	0	0	0	0
	1 tick	48	0	0	48	0	0	48	1	2
INTC										
	5 min	1	0	0	2	0	0	2	0	0
	15 min	0	0	0	1	0	0	1	0	0
	20 min	0	0	0	1	0	0	1	0	0
	1 tick	49	0	0	49	0	0	49	0	0
MMM										
	5 min	0	0	0	0	0	0	0	0	0
	15 min	0	0	0	0	0	0	0	0	0
	20 min	0	0	0	0	0	0	0	0	0
	1 tick	45	0	0	45	2	0	46	1	0
MSFT										
	5 min	4	1	2	8	0	0	8	2	0
	15 min	1	0	0	1	1	0	2	0	0
	20 min	1	0	0	2	1	0	3	0	0
	1 tick	49	0	0	49	1	0	49	1	1

Table 8: *Number of rejected null hypotheses out of 51 for each step. See section* 5.3 *for exact description of the test.* α *is the level of significance.*