ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Ahoniemi, Katja; Lanne, Markku

Working Paper

Realized volatility and overnight returns

Bank of Finland Research Discussion Papers, No. 19/2010

Provided in Cooperation with:
Bank of Finland, Helsinki

Suggested Citation: Ahoniemi, Katja; Lanne, Markku (2010) : Realized volatility and overnight returns, Bank of Finland Research Discussion Papers, No. 19/2010, ISBN 978-952-462-623-1, Bank of Finland, Helsinki,
https://nbn-resolving.de/urn:nbn:fi:bof-20140807555

This Version is available at:
https://hdl.handle.net/10419/212176

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Katja Ahoniemi – Markku Lanne

Realized volatility and overnight returns

SUOMEN PANKKI
FINLANDS BANK
EUROJÄRJESTELMÄ
EUROSYSTEMET

Bank of Finland Research
Discussion Papers
19 • 2010

Suomen Pankki
Bank of Finland
PO Box 160
FI-00101 HELSINKI
Finland
☎ +358 10 8311

http://www.bof.fi
E-mail: Research@bof.fi

Bank of Finland Research
Discussion Papers
19 • 2010

Katja Ahoniemi* – Markku Lanne**

Realized volatility and overnight returns

The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland.

* Corresponding author. Aalto University School of Economics. Email: katja.ahoniemi@hse.fi.
** University of Helsinki. Email: markku.lanne@helsinki.fi

The authors thank Andrew Patton for valuable comments. Katja Ahoniemi thanks the Finnish Foundation for Advancement of Securities Markets, and Markku Lanne thanks the Academy of Finland and the OP-Pohjola Group Research Foundation for financial support.

http://www.bof.fi

ISBN 978-952-462-622-4
ISSN 0785-3572
(print)

ISBN 978-952-462-623-1
ISSN 1456-6184
(online)

Helsinki 2010

Realized volatility and overnight returns

Bank of Finland Research
Discussion Papers 19/2010

Katja Ahoniemi – Markku Lanne
Monetary Policy and Research Department

Abstract

No consensus has emerged on how to deal with overnight returns when calculating realized volatility in markets where trading does not take place 24 hours a day. This paper explores several common volatility applications, investigating how the chosen treatment of overnight returns affects the results. For example, the selection of the best volatility forecasting model depends on the way overnight returns are incorporated into realized volatility. The evidence favours weighted estimators over those that have been more commonly used in the existing literature. The definition of overnight returns is particularly challenging for the S&P 500 index, and we propose two alternative measures for its overnight return.

Keywords: realized volatility, forecasting

JEL classification numbers: C14, C22, C52

Toteutunut volatiliteetti ja yliyön tuotot

Suomen Pankin keskustelualoitteita 19/2010

Katja Ahoniemi – Markku Lanne
Rahapolitiikka- ja tutkimusosasto

Tiivistelmä

Toistaiseksi ei ole löytynyt yksimielisyyttä siitä, miten tulisi käsitellä yön yli -tuottoja, kun lasketaan toteutunutta volatiliteettia markkinoilla, joilla ei käydä kauppaa ympäri vuorokauden. Tässä tutkimuksessa tarkastellaan useita yleisiä volatiliteettisovelluksia ja sitä, miten valittu yön yli -tuottojen käsittelytapa vaikuttaa tuloksiin. Parhaan ennustemallin valinta volatiliteetin laskemiseksi riippuu esimerkiksi siitä, miten yön yli -tuotot sisällytetään toteutuneen volatiliteetin laskentaan. Tulosten perusteella kannattaa suosia painotettuja volatiliteettiestimaattoreita ennemmin kuin sellaisia estimaattoreita, joita on käytetty yleisemmin olemassa olevassa kirjallisuudessa. Yön yli -tuottojen määritteleminen on erityisen haastavaa S & P 500 -osakeindeksin tapauksessa, ja tässä keskustelualoitteessa esitetään kaksi vaihtoehtoista mittaria tämän indeksin yön yli -tuotoille.

Avainsanat: toteutunut volatiliteetti, ennustaminen

JEL-luokittelu: C14, C22, C52

Contents

Abstract 3
Tiivistelmä (abstract in Finnish) 4

1 Introduction 7

2 Realized volatility estimators 8

3 Data 9

4 Data-based ranking test 14

5 Empirical comparisons 15

6 Conclusions 18

References 20

1 Introduction

The use of high-frequency data to calculate and sum intraday squared returns has become the prevalent method for estimating volatility in recent years. The early literature on realized volatility (RV) dealt with foreign exchange markets, where trading takes place around the clock (see eg Andersen and Bollerslev, 1998). However, the same approach of summing intraday squared returns has since been applied to data from other markets that are closed for at least a part of each 24-hour period. The best way to incorporate the information that arrives during the times of market closure is not obvious at the outset.

The existing literature on stock market realized volatility has adopted several approaches to dealing with the time period when the market is closed (in other words, the overnight period). The simplest approach is to ignore the overnight period, in other words, summing only the intraday squared returns (Wu, 2010; Corsi et al, 2008); Thomakos and Wang, 2003; Andersen et al, 2001). However, Hansen and Lunde (2006) argue that such an estimator is not a proper proxy of the true volatility because it does not span a full 24-hour period. Another solution in the literature is to calculate the overnight return by subtracting each day's close value from the next day's open, and to add this squared return as one of the factors in the sum of intraday returns (Bollerslev et al, 2009; de Pooter et al, 2008; Becker et al, 2007; Martens, 2002; Blair et al, 2001). A third method is to calculate realized volatility by ignoring the overnight period, but then scaling the resulting value upward so that the volatility estimate covers an entire 24-hour day (Koopman et al, 2005; Martens, 2002). Fourth, Hansen and Lunde (2005b) have derived optimal weights for the overnight return and the sum of intraday returns.[1]

In the absence of a consensus approach, this paper compares all of the existing solutions to dealing with returns from the overnight period. We also introduce an alternative, naïve weighting scheme to complement that of Hansen and Lunde (2005b, henceforth HL). As the true data-generating process is unobservable, we cannot directly compare the various measures to the true volatility. However, we run the test of Patton (2009), which allows for ranking various volatility estimators. For S&P 500 returns, this test procedure selects the HL and naïve volatility estimators over the other alternatives when using a mean squared error type loss function. We also show that, in a basic volatility forecasting framework, the selection of the forecasting model can depend on the treatment of overnight returns. Mean squared errors tend to be lowest, and R^2's highest, with the HL estimator. Additional analyses based on volatility feedback and leverage effect regressions provide more evidence in favor of the HL weighted estimator. Combining all the collected evidence, we recommend the use of the HL estimator.

The data in the present paper is for the S&P 500 index, which is quoted from 9:30 AM to 4:00 PM Eastern time. In other words, trading only takes place for six and a half hours each day. In addition, the published opening quote of the index has, traditionally, been equal to the previous day's close

[1] There are also numerous studies with realized volatility applications that do not mention how the overnight return is treated, if accounted for at all. These include Giot and Laurent (2007) and Engle and Gallo (2006).

due to the fact that trading has not commenced immediately at 9:30 AM. This poses an additional challenge in determining the overnight returns for the index. To that end, we propose two alternative solutions: using the return from the previous close up until 9:35 AM (a five minute return), and using the difference between the previous close and the so-called special opening quote (SOQ). The five-minute proxy is favored by the Patton (2009) test, but the SOQ proxy dominates in the empirical applications.

The rest of the paper is organized as follows. Section 2 describes all of the RV measures that are compared in the paper. Section 3 describes the data set and the two proposed overnight return proxies. Section 4 provides details on the Patton (2009) test and shows which RV estimators it selects. Section 5 contains the empirical analyses: volatility forecasts as well as volatility feedback and leverage effect regressions. Section 6 concludes.

2 Realized volatility estimators

This section outlines the various competing realized volatility estimators that are compared in later sections. In what follows, the term realized volatility refers to both realized variance and its square root. The basic measure of RV is defined in the standard way

$$RV_t = \sum_{i=1}^{m} (p_{t,i} - p_{t,i-1})^2 \tag{2.1}$$

where p is the log price of an asset and m denotes the number of intraday returns to be summed. For example, in the US stock market, there are 78 five-minute returns in one trading day. The squared overnight return ON can be added to this simple intraday measure as a 79th factor in the sum. Alternatively, the RV estimator can be scaled in such a way that the six-and-a-half hour trading day is extended into a 24-hour day. We follow the scaling procedure of HL, and define the scaled estimator as

$$RV_t^{SC} = \hat{c} \cdot RV_t \tag{2.2}$$

where

$$\hat{c} = \frac{\sum_{t=1}^{n} (r_t - \bar{r})^2}{\sum_{t=1}^{n} RV_t} \tag{2.3}$$

and where r_t is the daily close-to-close return, and $\bar{r}$ is its sample average over the n-day sample period. HL show that this scaling factor is a consistent estimator of c, which is defined as $E(IV_t)/E(RV_t)$, where IV_t is the integrated variance.

HL also introduce a way to optimally weight the squared overnight return and the sum of intraday squared returns. Denoting the average overnight squared return with μ_1, the average of RV_t with μ_2, and the average of $RV_t + ON_t$ with μ, the linear combination of the two elements is constructed in such a way that $\omega_1\mu_1 + \omega_2\mu_2 = \mu$, where the weights ω_1 and ω_2 are the optimal weights of ON_t and RV_t, respectively.

The naïve RV estimator proposed in this paper is also a weighted sum of the squared overnight return and the intraday squared returns. The weights are selected so that $\omega_2/\omega_1 = \mu_2/\mu_1$, satisfying $\omega_1\mu_1 + \omega_2\mu_2 = \mu$. In other words, the weights are in the same proportion as the average overnight squared return and the average value of RV_t (the sum of intraday squared returns).

3 Data

The data set in this study consists of intraday observations of the S&P 500 index, the most important benchmark in the US stock market. The data is acquired from Tick Data, and the chosen data sample covers the period 1.1.1994–31.5.2007, containing 3,377 observations. The end date is chosen so that the financial crisis period is omitted from the analysis due to concerns that normal phenomena and behavior have been (temporarily) disturbed.[2] Although the S&P 500 index value is published several times each minute, we choose to use five-minute returns for our analysis. As the index consists of 500 share prices, observing the value very frequently would introduce a large number of stale prices. It is plausible to assume that nearly all component stocks of the S&P 500 index trade at least once within each five-minute interval. Also, Bollerslev et al (2009) note that for highly liquid assets, the five-minute frequency seems to be a reasonable choice, providing a balance between frequent sampling and market microstructure noise.[3]

To facilitate the out-of-sample forecasting exercise of Section 5, the data sample is further divided into an in-sample period of 2,771 observations, corresponding to 1.1.1994–31.12.2004, and an out-of-sample period of 1.1.2005–31.5.2007 for forecast evaluation. The in-sample period is used to determine the scaling factor $\hat{c}$ and the HL and naïve weights ω_1 and ω_2.

The estimation of S&P 500 volatility faces two challenges. The first is the issue of how to incorporate overnight returns into the estimate of realized volatility.[4] The second challenge involves how to calculate the overnight return in the case of the S&P 500 index, as the previous day's close and next day's open quotes have traditionally been exactly the same. This stems from the fact that it takes some time to open stocks for trading on the New York Stock Exchange (NYSE). Figure 1 shows the daily squared close-to-open returns for the time period of this study. Prior to 2006, the squared return is non-zero on some days in the sample period, but this is most likely due to clerical errors or

[2] The Financial Turmoil Timeline maintained by the Federal Reserve Bank of New York at http://www.newyorkfed.org/research/global_economy/crisis_timeline.pdf begins in June 2007.

[3] We choose not to make corrections for possible market microstructure effects (methods for dealing with market microstructure noise are introduced in eg Zhang et al, 2005, and Barndorff-Nielsen et al, 2008). The five-minute observation frequency should take care of most noise considerations arising from e.g. infrequent trading and price discreteness. Other equity market studies such as Andersen et al (2001) and Engle and Gallo (2006) also use the five-minute observation frequency, whereas Giot and Laurent (2007) use a 20-minute frequency, and Becker et al (2007) a 30-minute frequency.

[4] This issue is circumvented by eg Martens (2002) by using S&P 500 futures instead of the cash index, as futures trading takes place around the clock.

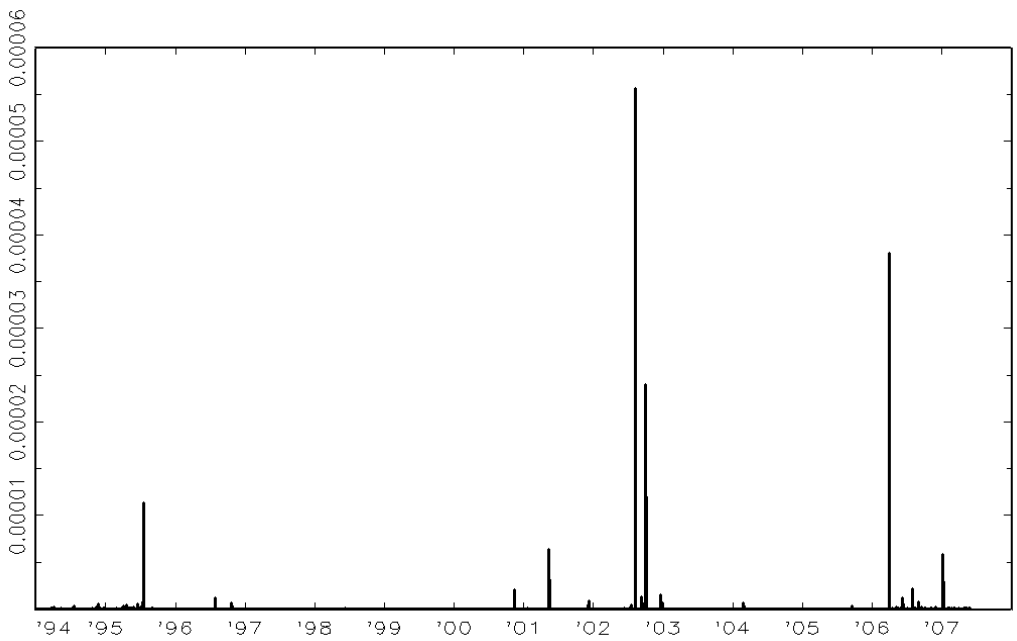

Figure 1: Close-to-open squared return of the S&P 500 index 1.1.1994–31.5.2007 (3377 observations)

other unusual circumstances. As Stoll (2000) notes, the consecutive close and open quotes of the S&P 500 index are identical. A change occurs gradually in 2006, near the end of the data set, whereafter the close-to-open return is no longer zero. This coincides with the adoption of an electronic trading system at the New York Stock Exchange, and with changes that facilitated the opening of stocks for trading at the start of the day.

Given the above evidence, it is not feasible to use the difference between the open and close quotes to calculate the overnight return in the case of the S&P 500 index. As a first remedy, we propose that the return from the previous day's close up until 9:35 AM, or up until five minutes of trading have elapsed, be used as a proxy of the overnight return. This same measure of overnight returns is used in Chan et al (1991). Although perhaps seemingly arbitrary, there is evidence suggesting that this choice is logical. First, when calculating the returns for each five-minute period of the day, the average squared return for the first five minutes differs markedly from the remaining 77 five-minute periods of the trading day. The standard deviation of the squared returns is also higher for the squared returns of the first five minutes (see Figure 2). In fact, this phenomenon prompted Stoll and Whaley (1990b) to omit the first two five-minute returns of each trading day from their analysis of S&P 500 cash index and index futures returns, and Lin et al (1994) wait for 30 minutes of trading to elapse on the NYSE before observing their opening quote for the S&P 500 index. Stoll and Whaley (1990a) report that in 1986, it took large stocks around five to six minutes to open for trading on the NYSE. Spurlin et al (2008), in a sample of S&P 500 stocks covering the years 1996–2001, provide evidence that the average opening time had fallen to around 3.5 minutes. As a consequence of our decision to use the first five minutes of each day as an

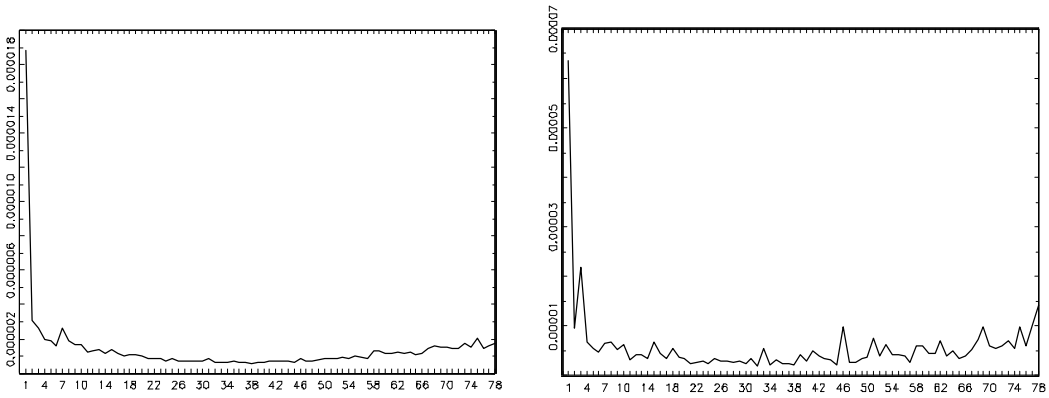

Figure 2: Average squared return of five-minute trading periods (left panel) and standard deviation of five-minute squared returns (right panel). There are 78 five-minute periods within one trading day at the NYSE.

overnight return proxy, in what follows we calculate RV_t beginning at 9:35 AM, or with 77 daily five-minute intervals.

As a second alternative proxy for the S&P 500 overnight returns, we propose using the so-called special opening quote. This index value is calculated daily by Standard & Poor's using the opening value of each of the 500 component stocks, and is used as the final settlement price of S&P 500 index-related options that trade on the CBOE and CME. In other words, it is calculated at a different time each day – once the first trade of the day takes place for the 500th stock. However, given the above evidence that large stocks tend to open during the first few minutes of each trading day, there should be little risk of the new day's news contaminating the open quotes. Thus, the special opening quote (SOQ) should reflect the information that accumulates during the overnight period quite well. Figure 3 shows the time series of the close-to-9:35 squared returns and the close-to-SOQ squared returns. The y-axis is equal in both panels of the figure, highlighting that the close-to-SOQ return is larger on average.

The following list summarizes the various RV measures that will be compared in later phases of this study:

- RV_t – sum of intraday squared returns with no overnight return
- RV_t^{SC} – RV_t scaled to a full 24 hours; scaling factor $\hat{c} = 1.6380$
- $RV_t^{+9:35}$ – RV_t plus the overnight return, defined as the close-to-9:35 return
- RV_t^{+SOQ} – RV_t plus the overnight return, defined as the close-to-SOQ return
- $RV_t^{HL,9:35}$ – RV_t and the overnight return, defined as the close-to-9:35 return, HL weights
- $RV_t^{HL,SOQ}$ – RV_t and the overnight return, defined as the close-to-SOQ return, HL weights

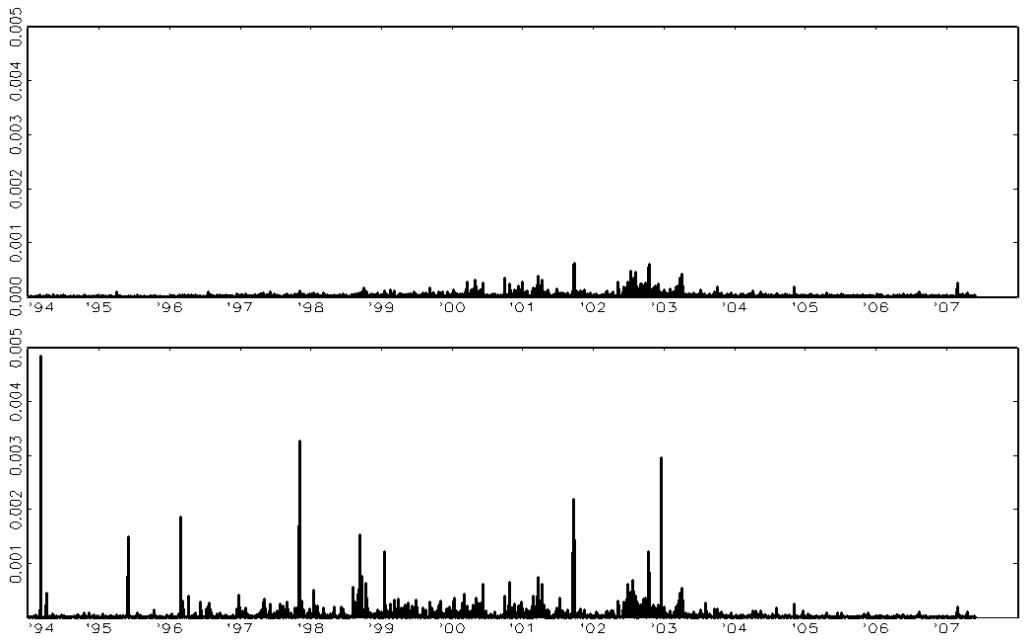

Figure 3: Close-to-9:35 squared return of the S&P 500 index (upper panel) and close-to-SOQ squared return (lower panel) 1.1.1994–31.5.2007.

- $RV_t^{N,9:35}$ – RV_t and the overnight return, defined as the close-to-9:35 return, naïve weights
- $RV_t^{N,SOQ}$ – RV_t and the overnight return, defined as the close-to-SOQ return, naïve weights

As a representative example of the various RV series, Figure 4 shows the time series of RV_t for the full sample period. Table 1 provides descriptive statistics on the eight RV estimators and the two overnight return series, and Table 2 contains the HL and naïve weights, as well as ω_2/ω_1. The descriptive statistics show that RV_t^{SC} is highest, on average. The value for RV with naïve or HL weights depends on the chosen measure of the overnight return – using the SOQ leads to a higher RV estimate than using the five-minute overnight return proxy. The averages of $RV_t^{+9:35}$, $RV_t^{HL,9:35}$, and $RV_t^{N,9:35}$ are very close to one another, as are the averages of RV_t^{+SOQ}, $RV_t^{HL,SOQ}$, and $RV_t^{N,SOQ}$. This is because the weights are calculated so that $\omega_1\mu_1 + \omega_2\mu_2 = \mu$. In fact, this causes the means to be exactly equal within the sample that is used to determine the weights. In our case, the use of the in-sample period to find the weights causes the averages to differ slightly when looking at the full sample. The mean of the SOQ overnight return proxy is over double that of the five-minute return proxy.

Despite the closeness in means of the weighted estimators, Table 2 shows that the four sets of weights (HL and naïve weights for two overnight return proxies) differ clearly from one another. Looking first at the five-minute return proxy, the HL weights place two times as much weight on the day-time squared returns than on the overnight squared return. For the naïve weights, in turn, ω_2/ω_1 is close to four, or the in-sample value of μ_2/μ_1. The picture is reversed

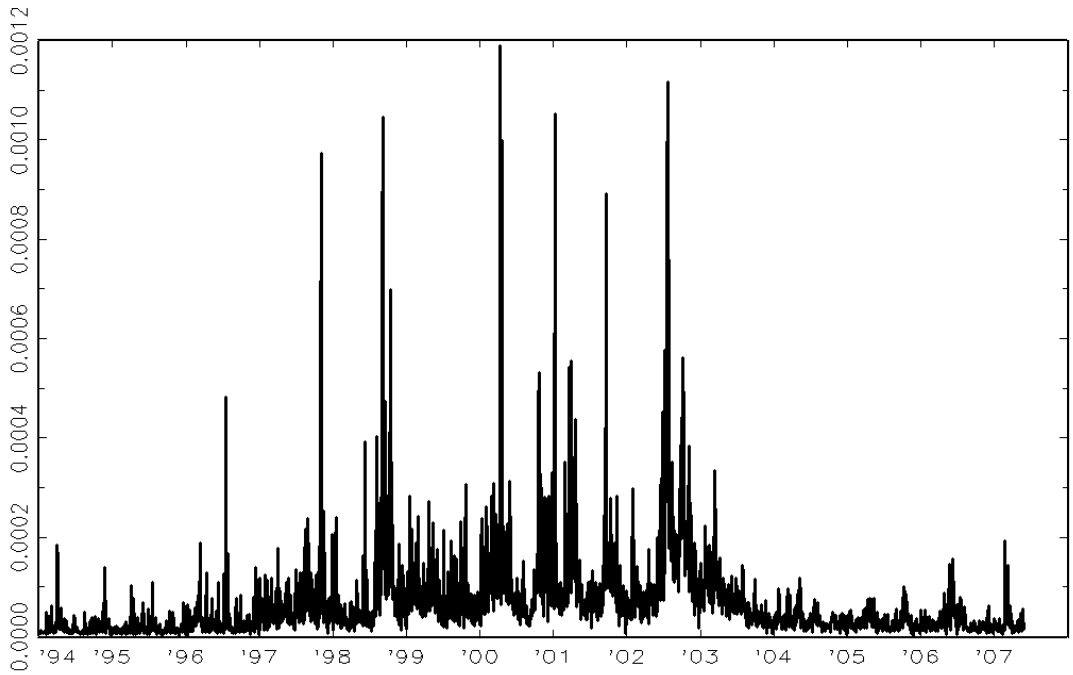

Figure 4: RV_t, or the sum of intraday squared returns, 1.1.1994–31.5.2007.

for the SOQ, which produced a much higher average overnight squared return. The HL weights now award much less weight to the overnight return, whereas the naïve weights react to the higher overnight values by awarding it more weight. With the naïve scheme, the daytime weight is now less than double the night-time weight, whereas the relation ω_2/ω_1 is over 39 with HL weights. Note that the weights for the two weighting schemes do not need to sum to the same value.

Table 1. Descriptive statistics for realized volatility measures and overnight squared returns, multiplied by 1000.

	Mean	St.dev.
RV_t	0.0667	0.0909
RV_t^{SC}	0.1092	0.1489
$RV_t^{+9:35}$	0.0833	0.1128
RV_t^{+SOQ}	0.1027	0.1973
$RV_t^{HL,9:35}$	0.0833	0.1114
$RV_t^{HL,SOQ}$	0.1040	0.1417
$RV_t^{N,9:35}$	0.0832	0.1118
$RV_t^{N,SOQ}$	0.1031	0.1677
ON 9:35	0.0167	0.0411
ON SOQ	0.0360	0.1496

Table 2. HL and naïve weights with both proxies of S&P 500 overnight returns

	9:35	
	HL	naïve
ω_1	0.5483	0.2909
ω_2	1.1118	1.1755
ω_2/ω_1	2.0277	4.0408
	SOQ	
	HL	naïve
ω_1	0.0390	0.6656
ω_2	1.5386	1.1874
ω_2/ω_1	39.4175	1.7841

4 Data-based ranking test

The data-generating process, and thus the true volatility, are unobservable. However, the test procedure introduced by Patton (2009) provides a straightforward way for ranking various volatility estimators with the help of the daily close-to-close return. We first perform this test in order to gauge the differences in the RV estimators, and to find an unconditional ranking of the estimators. Patton (2009) provides conditions under which loss functions based on any unbiased proxy of the RV can be used to compare competing RV estimators by means of standard tests, such as those of Diebold and Mariano (1995) and the Model Confidence Set (MCS) test of Hansen et al (2009). This proxy does not have to be very precise, and following Patton (2009), we employ the readily available daily squared return. However, standard tests cannot be directly used to compare RV estimators to the daily squared return because the estimation error in the RV will, in general, be correlated with the error in the proxy. To overcome this difficulty, Patton (2009) suggests using the first lead of the squared return as an instrument in the loss function (an alternative would be a weighted average of multiple leads of the proxy). Furthermore, we adopt his assumption that the RV follows a random walk, which has often proved to be a good approximation. These assumptions, coupled with a number of mild regularity conditions, are sufficient to consistently estimate the difference in average accuracy of any two competing estimators. Hence, it is justified to use the Diebold-Mariano test and the MCS test for comparisons of RV estimator accuracy based on commonly used loss functions. We employ the MCS test in our application, as it allows for comparing the full range of RV estimators simultaneously. The best estimator is included in the model confidence set with a given confidence level. The MCS test does not require knowledge of the true data-generating process, in other words, knowledge of the true realized volatility. We use the mean squared error loss function

$$L_t = (r_{t+1}^2 - V_t)^2 \tag{4.1}$$

where V_t is an RV estimator. The MCS test results are provided in Table 3. The p-values indicate that with this loss function, $RV_t^{N,9:35}$ and $RV_t^{HL,9:35}$ are the

most accurate measures. These two measures belong to the model confidence set with a confidence level of ten percent, but they cannot be distinguished from one another in a statistical sense. Based on the results from the test, it appears to be favorable to weight the overnight and intraday returns, and both of the proposed weighting schemes fare well. Also, the 9:35 AM overnight return proxy dominates the special opening quote.

Table 3. P-values from model confidence set test

RV estimator	MCS p-value
RV_t	0.0000
RV_t^{SC}	0.0000
$RV_t^{+9:35}$	0.0000
RV_t^{+SOQ}	0.0000
$RV_t^{HL,9:35}$	0.1735
$RV_t^{HL,SOQ}$	0.0168
$RV_t^{N,9:35}$	1.0000

5 Empirical comparisons

We now turn our attention to several empirical volatility applications in order to further investigate which RV estimator should be favored when analyzing the S&P 500 index. First, we compare the RV estimators through a volatility forecasting exercise. As mentioned previously, we divide the data into an in-sample period (1.1.1994–31.12.2004) and an out-of-sample period (1.1.2005–31.5.2007). The in-sample and out-of-sample periods contain 2,771 and 606 observations, respectively.

There is, naturally, a virtually unlimited number of forecasting models to choose from. To keep the analysis compact, we focus on only three commonly used models: the GARCH(1,1), GJR-GARCH(1,1), and APARCH(1,1). We include the basic GARCH model as it is a natural benchmark in a forecasting application, and it was found to perform very well as a forecaster in Hansen and Lunde (2005a) for an exchange rate series. The popular GJR model of Glosten et al (1993) takes the asymmetric market reaction to positive and negative shocks into account. Hansen and Lunde (2005a) find that when forecasting the volatility of a stock return series, taking such a leverage effect into account improves forecasts. The GJR model is also the chosen specification for S&P 100 returns in Blair et al (2001). The APARCH model, introduced in Ding et al (1993), is the best volatility forecaster in an application similar to ours, Hansen and Lunde (2006).[5] This group of models is sufficient to illustrate our point that the treatment of overnight returns can affect the conclusions a researcher would draw concerning the best forecasting model. We evaluate the forecasts

[5] Hansen and Lunde (2006) perform a similar forecasting exercise with three RV estimators: RV_t^{SC}, $RV + ON$ (equivalent to $RV_t^{+9:35}$ and RV_t^{+SOQ} in this study), and the daily squared return. Their focus, however, is not on the comparison of RV estimators, but on which forecast evaluation criteria to use.

with two standard criteria, mean squared error (MSE) and the R^2 from a Mincer-Zarnowitz regression of the RV estimator on the forecast. Patton (2008) shows that ranking forecasts based on MSE is robust to noise in the volatility proxy, and Meddahi (2002) shows the same for the Mincer-Zarnowitz R^2. However, Pagan and Schwert (1990) and Engle and Patton (2001) note that the R^2 criterion can be sensitive to outliers in the volatility proxies.

Table 4 summarizes the results. The R^2 criterion is very clear: the best forecasting model is GJR-GARCH, irrespective of the RV estimator used. However, based on the MSE, the model choice varies depending on the RV estimator, and would fall either on the GARCH model or the APARCH model. Looking at both criteria simultaneously, a researcher would choose from the basic GARCH model and the GJR model if using RV_t, $RV_t^{+9:35}$, RV_t^{+SOQ}, $RV_t^{HL,9:35}$, $RV_t^{N,9:35}$ or $RV_t^{N,9:35}$, but the choice would fall on GJR or APARCH when using RV_t^{SC} or $RV_t^{HL,SOQ}$. The highest R^2 is provided by the GJR forecasts for $RV_t^{HL,9:35}$, and the lowest MSE comes with APARCH forecasts for $RV_t^{HL,SOQ}$. Going row by row, $RV_t^{HL,SOQ}$ delivers the best value (lowest MSE or highest R^2) three times, $RV_t^{HL,9:35}$ twice, and $RV_t^{N,SOQ}$ once. Although we are cautious to draw any direct conclusions from this, it is an additional indication that the weighted estimators are the most precise. In light of the forecast results, the SOQ now receives support as an overnight return proxy, although the Patton (2009) test favored the five-minute proxy.

To further explore the differences in the RV estimators, we follow Bollerslev and Zhou (2006) and run volatility feedback and leverage effect regressions with a daily sampling frequency. The volatility feedback effect, which underlies, for example, the ARCH-M model of Engle et al (1987), implies that returns and volatility should be positively linked. To investigate the volatility feedback effect, the daily log return is regressed on the contemporaneous RV estimator

$$r_t = \alpha + \beta V_t + \varepsilon_t \tag{5.1}$$

The leverage effect, on the other hand, links the previous period's return with the next period's volatility. The relation is assumed to be negative, so that a negative (positive) return raises (lowers) volatility in the next period. In the model for the leverage effect, the RV estimator is regressed on the previous period's log return

$$V_t = \alpha + \beta r_{t-1} + \varepsilon_t \tag{5.2}$$

The adjusted R^2's from the two regressions are provided in Table 5, along with the significance levels of the explanatory variables with Newey-West HAC standard errors. The results are more interesting for the volatility feedback effect. The highest adjusted R^2 is achieved with RV_t^{SC} and RV_t, which yield the same value, as $RV_t^{SC} = \hat{c} \cdot RV_t$. The second-highest R^2 comes with $RV_t^{HL,SOQ}$. If the analysis of the volatility feedback effect were the only issue on the research agenda, the results would be affected by the choice of the RV estimator. In particular, if using the popular estimator of type $RV + ON$, here either $RV_t^{+9:35}$ or RV_t^{+SOQ}, a researcher would easily conclude that there is no strong volatility feedback effect. The evidence in favor of the effect is much stronger with some of the other estimators. For the leverage effect

Table 4. Out-of-sample forecast results 1.1.2005–31.5.2007: mean squared errors and R^2s. MSEs multiplied by 10^8. Forecasting model that would be selected within each column and each criterion in boldface.

	RV_t	RV_t^{SC}	$RV_t^{+9:35}$	RV_t^{+SOQ}	$RV_t^{HL,9:35}$	$RV_t^{HL,SOQ}$	$RV_t^{N,9:35}$	$RV_t^{N,SOQ}$
	Mean squared errors							
GARCH (1,1)	**0.1159**	0.1026	**0.1058**	**0.1062**	**0.0976**	0.0979	**0.0966**	**0.0962**
GJR-GARCH(1,1)	0.1515	0.1061	0.1234	0.1264	0.1177	0.1058	0.1182	0.1121
APARCH(1,1)	0.1227	**0.0938**	0.1113	0.1115	0.1009	**0.0915**	0.0988	0.0974
	R^2							
GARCH (1,1)	0.2093	0.2093	0.2111	0.1961	0.2245	0.2108	0.2223	0.2150
GJR-GARCH(1,1)	**0.3118**	**0.3118**	**0.3175**	**0.2894**	**0.3364**	**0.3140**	**0.3322**	**0.3185**
APARCH(1,1)	0.3047	0.3047	0.2529	0.2402	0.2928	0.3053	0.3050	0.2813

regressions, there are no differences in statistical significance. However, the highest adjusted R^2 is associated with $RV_t^{HL,SOQ}$, and RV_t^{SC} and RV_t are now second-best. To summarize, this part of the analysis points toward $RV_t^{HL,SOQ}$ as the estimator of choice. This same estimator fared best in the row-by-row analysis of the forecast results, followed by $RV_t^{HL,9:35}$.

Table 5. Adjusted R^2's from volatility feedback and leverage effect regressions. *** denotes significance at the one-percent level, level, ** at the five-percent level, and * at the ten-percent level for the explanatory variable in each equation: for the volatility feedback effect, V_t from the model $r_t = \alpha + \beta V_t + \varepsilon_t$ and for the leverage effect, r_{t-1} from the model $V_t = \alpha + \beta r_{t-1} + \varepsilon_t$.

	Vol. feedback	Leverage
RV_t	0.0132***	0.0524***
RV_t^{SC}	0.0132***	0.0524***
$RV_t^{+9:35}$	0.0045*	0.0434***
RV_t^{+SOQ}	0.0000	0.0324***
$RV_t^{HL,9:35}$	0.0081**	0.0488***
$RV_t^{HL,SOQ}$	0.0124***	0.0529***
$RV_t^{N,9:35}$	0.0105***	0.0509***
$RV_t^{N,SOQ}$	0.0015	0.0423***

6 Conclusions

When calculating realized volatility in a market that is not open for continuous trading, a choice must be made on how to treat the returns generated during the periods of market closure. Several alternative ways to deal with these overnight returns have become prevalent in the related literature. This paper shows, using a formal statistical test due to Patton (2009), that weighting the squared overnight return and the sum of intraday squared returns is the most accurate measure of realized volatility. The importance of the choice regarding overnight return treatment is underscored in the volatility forecasting exercise, which shows that a researcher would select a different forecasting model with different estimators of realized volatility. Based on the combined evidence from the statistical test and the forecasts, the weighted RV due to Hansen and Lunde (2005b) appears to be the best RV estimator. The HL weighted estimator performs better than the intuitive but naïvely weighted estimator introduced in the paper.

We also explore how to proxy for the overnight return of the S&P 500 index, as its close and open quotes have traditionally been equal. The test of Patton (2009) favors the five-minute overnight return proxy, whereas the special opening quote emerges as the top alternative in the empirical applications. The volatility feedback effect analysis again shows that the treatment of overnight returns can affect conclusions drawn by a researcher: statistical

significance is not achieved with all RV estimators. Importantly, we can conclude that two commonly-used RV estimators, one that adds overnight returns as one equal factor into the sum of intraday returns, and one that includes no overnight returns, cannot be recommended based on our analysis. The combined evidence is most in favor of the Hansen and Lunde (2005b) weighted realized volatility estimator.

References

Andersen, T G – Bollerslev, T (1998) **Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts.** International Economic Review, 39, 885–905.

Andersen, T G – Bollerslev, T – Diebold, F X – Ebens, H (2001) **The Distribution of Realized Stock Return Volatility.** Journal of Financial Economics, 61, 43–76.

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

Becker, R – Clements, A E – White, S I (2007) **Does Implied Volatility Provide Any Information Beyond That Captured in Model-Based Volatility Forecasts?** Journal of Banking & Finance, 31, 2535–2549.

Blair, B J – Poon, S-H – Taylor, S J (2001) **Forecasting S&P 100 Volatility: the Incremental Information Content of Implied Volatilities and High-frequency Index Returns.** Journal of Econometrics, 105, 5–26.

Bollerslev, T – Zhou, H (2006) **Volatility Puzzles: a Simple Framework for Gauging Return-Volatility Regressions.** Journal of Econometrics, 131, 123–150.

Bollerslev, T – Tauchen, G – Zhou, H (2009) **Expected Stock Returns and Variance Risk Premia.** Review of Financial Studies, 22, 4463–4492.

Chan, K – Chan, K C – Karolyi, G A (1991) **Intraday Volatility in the Stock Index and Stock Index Futures Market.** Review of Financial Studies, 4, 657–684.

Corsi, F – Mittnik, S – Pigorsch, C – Pigorsch, U (2008) **The Volatility of Realized Volatility.** Econometric Reviews, 27, 46–78.

de Pooter, M – Martens, M – van Dijk, D (2008) **Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data – But Which Frequency to Use?** Econometric Reviews, 27, 199–229.

Diebold, F X – Mariano, R S (1995) **Comparing Predictive Accuracy.** Journal of Business & Economic Statistics, 13, 253–263.

Ding, Z – Granger, C W J – Engle, R F (1993) **A Long Memory Property of Stock Market Returns and a New Model.** Journal of Empirical Finance, 1, 83–106.

Engle, R F – Gallo, G M (2006) **A Multiple Indicators Model for Volatility Using Intra-daily Data.** Journal of Econometrics, 131, 3–27.

Engle, R F – Lilien, D M – Robins, R P (1987) **Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model.** Econometrica, 55, 391–407.

Engle, R F – Patton, A J (2001) **What Good is a Volatility Model?** Quantitative Finance, 1, 237–245.

Giot, P – Laurent, S (2007) **The Information Content of Implied Volatility in Light of the Jump/Continuous Decomposition of Realized Volatility.** Journal of Futures Markets, 27, 337–359.

Glosten, L R – Jagannathan, R – Runkle, D E (1993) **On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks.** Journal of Finance, 48, 1779–1801.

Hansen, P R – Lunde, A (2005a) **A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?** Journal of Applied Econometrics, 20, 873–889.

Hansen, P R – Lunde, A (2005b) **A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data.** Journal of Financial Econometrics, 3, 525–554.

Hansen, P R – Lunde, A (2006) **Consistent Ranking of Volatility Models.** Journal of Econometrics, 131, 97–121.

Hansen, P R – Lunde, A – Nason, J M (2009) **The Model Confidence Set.** Federal Reserve Bank of Atlanta Working Paper 2005-7a.

Koopman, S J – Jungbacker, B – Hol, E (2005) **Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised, and Implied Volatility Measurements.** Journal of Empirical Finance, 12, 445–475.

Lin, W-L – Engle, R F – Ito, T (1994) **Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility.** Review of Financial Studies, 7, 507–538.

Martens, M (2002) **Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-frequency Data.** Journal of Futures Markets, 22, 497–518.

Meddahi, N (2002) **A Theoretical Comparison Between Integrated and Realized Volatility.** Journal of Applied Econometrics, 17, 479–508.

Pagan, A R – Schwert, G W (1990) **Alternative Models for Conditional Volatility.** Journal of Econometrics, 45, 267–290.

Patton, A J (2008) **Volatility Forecast Comparison Using Imperfect Volatility Proxies.** Journal of Econometrics, forthcoming.

Patton, A J (2009) **Data-Based Ranking of Realised Volatility Estimators.** Unpublished working paper.

Spurlin, W P – Van Ness, B F – Van Ness, R A (2008) **Open Volume and Time to Open on Option Expiration Days.** International Review of Economics and Finance, 17, 245–257.

Stoll, H R (2000) **Friction.** Journal of Finance, 55, 1479–1514.

Stoll, H R – Whaley, R E (1990a) **Stock Market Structure and Volatility.** Review of Financial Studies, 3, 37–71.

Stoll, H R – Whaley, R E (1990b) **The Dynamics of Stock Index and Stock Index Futures Returns.** Journal of Financial and Quantitative Analysis, 25, 441–467.

Thomakos, D D – Wang, T (2003) **Realized Volatility in the Futures Markets.** Journal of Empirical Finance, 10, 321–353.

Wu, L (2010) **Variance Dynamics: Joint Evidence from Options and High-Frequency Returns.** Journal of Econometrics, forthcoming.

Zhang, L – Mykland, P A – Aït-Sahalia, Y (2005) **A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data.** Journal of the American Statistical Association, 100, 1394–1411.

BANK OF FINLAND RESEARCH
DISCUSSION PAPERS

ISSN 0785-3572, print; ISSN 1456-6184, online

1/2010 Timo Korkeamäki – Elina Rainio – Tuomas Takalo **Law and stock markets: evidence from an emerging market.** 2010. 34 p. ISBN 978-952-462-568-5, print; ISBN 978-952-462-569-2, online.

2/2010 Iftekhar Hasan – Heiko Schmiedel – Liang Song **Growth strategies and value creation: what works best for stock exchanges?** 2010. 48 p. ISBN 978-952-462-570-8, print; ISBN 978-952-462-571-5, online.

3/2010 Iftekhar Hasan – Heiko Schmiedel – Liang Song **Return from retail banking and payments.** 2010. 53 p. ISBN 978-952-462-572-2, print; ISBN 978-952-462-573-9, online.

4/2010 Sungho Choi – Bill B Francis – Iftekhar Hasan **Cross-border bank M&As and risk: evidence from the bond market.** 2010. 47 p. ISBN 978-952-462-574-6, print; ISBN 978-952-462-575-3, online.

5/2010 Hanna Freystätter **Financial market disturbances as sources of business cycle fluctuations in Finland.** 2010. 46 p. ISBN 978-952-462-576-0, print; ISBN 978-952-462-577-7, online.

6/2010 Patrick M Crowley **Long cycles in growth: explorations using new frequency domain techniques with US data.** 2010. 49 p. ISBN 978-952-462-586-9, print; ISBN 978-952-462-587-6, online.

7/2010 Paavo Miettinen **Information acquisition during a Dutch auction.** 2010. 30 p. ISBN 978-952-462-588-3, print; ISBN 978-952-462-589-0, online.

8/2010 Syed Mujahid Hussain **Simultaneous monetary policy announcements and international stock markets response: an intraday analysis.** 2010. 32 p. ISBN 978-952-462-590-6, print; ISBN 978-952-462-591-3, online.

9/2010 Antti J Tanskanen – Petri Niininen – Kari Vatanen **Risk-based classification of financial instruments in the Finnish statutory pension scheme TyEL.** 2010. 35 p. ISBN 978-952-462-592-0, print; ISBN 978-952-462-593-7, online.

10/2010 Hanna Putkuri **Housing loan rate margins in Finland.** 2010. 39 p. ISBN 978-952-462-594-4, print; ISBN 978-952-462-595-1, online.

11/2010 Efrem Castelnuovo – Salvatore Nisticò **Stock market conditions and monetary policy in a DSGE model for the US.**2010. 62 p. ISBN 978-952-462-596-8, print; ISBN 978-952-462-597-5, online.

12/2010 Juha Kilponen – Torsten Santavirta **New evidence on implicit contracts from linked employer-employee data.** 2010. 25 p. ISBN 978-952-462-600-2, print; ISBN 978-952-462-601-9, online.

13/2010 George W Evans – Seppo Honkapohja – Kaushik Mitra **Does Ricardian Equivalence hold when expectations are not rational?** 2010. 26 p. ISBN 978-952-462-602-6, print; ISBN 978-952-462-603-3, online.

14/2010 Karlo Kauko **The feasibility of through-the-cycle ratings.** 2010. 32 p. ISBN 978-952-462-604-0, print; ISBN 978-952-462-605-7, online.

15/2010 Risto Herrala **Credit constraints and durable consumption. A new empirical approach.** 2010. 19 p. ISBN 978-952-462-614-9, print; ISBN 978-952-462-615-6, online.

16/2010 Patrick M Crowley – Aaron Schultz **A new approach to analyzing convergence and synchronicity in growth and business cycles: cross recurrence plots and quantification analysis.** 2010. 53 p. ISBN 978-952-462-616-3, print; ISBN 978-952-462-617-0, online.

17/2010 Esa Jokivuolle – Ilkka Kiema – Timo Vesala **Credit allocation, capital requirements and output.** 2010. 35 p. ISBN 978-952-462-618-7, print; ISBN 978-952-462-619-4, online.

18/2010 Seppo Orjasniemi **The effect of openness in a small open monetary union.** 2010. 31 p. ISBN 978-952-462-620-0, print; ISBN 978-952-462-621-7, online.

19/2010 Katja Ahoniemi – Markku Lanne **Realized volatility and overnight returns.** 2010. 24 p. ISBN 978-952-462-622-4, print; ISBN 978-952-462-623-1, online.

Suomen Pankki
Bank of Finland
P.O.Box 160
FI-00101 HELSINKI
Finland

. 2343