AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

On the estimation of the volatility-growth link

Andrey Launov, Olaf Posch and Klaus Wälde

CREATES Research Paper 2012-21

Department of Economics and Business
Aarhus University
Bartholins Allé 10
DK-8000 Aarhus C
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

On the estimation of the volatility-growth link

Andrey Launov[(a)], Olaf Posch[(b)] and Klaus Wälde[(a,c)1]

[(a)]University of Mainz, CESifo, and UCL Louvain la Neuve
[(b)]Aarhus University and CREATES, CESifo
[(c)]Center for Structural Estimation, University of Bristol

April 2012

It is common practice to estimate the volatility-growth link by specifying a standard growth equation such that the variance of the error term appears as an explanatory variable in this growth equation. The variance in turn is modelled by a second equation. Hardly any of existing applications of this framework includes exogenous controls in this second variance equation. Our theoretical findings suggest that the absence of relevant explanatory variables in the variance equation leads to a biased and inconsistent estimate of the volatility-growth link. Our simulations show that this effect is large. Once the appropriate controls are included in the variance equation consistency is restored. In short, we suggest that the variance equation must include relevant control variables to estimate the volatility-growth link.

JEL codes: E32, O47
Keywords: volatility and growth, growth regression, endogenous variance unbiased estimates

1 Introduction

The background – Understanding the link between volatility and growth is central to many empirical analyzes. A prominent approach includes the variance of the error term of a growth regression into the very same growth equation as an explanatory variable. This is the approach pioneered by Ramey and Ramey (1995), henceforth RR. This approach has been extremely influential and led to many valuable insights. Despite this huge success, endogeneity of growth volatility has been seldom discussed.

The problem – Although it has been established that growth volatility is endogenous to determinants of economic growth,[2] empirical modeling of such dependence has not been discussed in detail so far. We argue that failing to properly account for dependence of the error variance on exogenous factors in this type of modelling may substantially bias the parameter estimates. It has already been recognized by RR that the endogeneity of volatility is important. They

[1]Contact details of authors: Olaf Posch, Aarhus University, Department of Economics and Business, Bartholins Allé 10, 8000 Århus C, phone +45.8942-1938, e-mail oposch@econ.au.dk. Andrey Launov and Klaus Wälde, University of Mainz, Mainz School of Management and Economics, Jakob-Welder-Weg 4, 55131 Mainz, Germany, fax + 49.6131.39-23827, phone + 49.6131.39-22006, e-mail andrey.launov@uni-mainz.de, klaus.waelde@uni-mainz.de. The first author appreciates financial support from the Center for Research in Econometric Analysis of Time Series, CREATES, funded by The Danish National Research Foundation. We thank Florian Heiss for discussions and comments.

[2]For a discussion of the literature on economic growth see Temple (1999). An elaborated investigation on linking the endogeneity of macroeconomic volatility to weak institutions is in Acemoglu et al. (2003). Theoretical analysis of the joint endogeneity of long-run growth and short-run volatility are undertaken in the 'natural volatility' literature (see e.g. Matsuyama, 1999, Francois Lloyed-Ellis, 2003, Wälde 2005 and Posch and Wälde, 2011). Fernández-Villaverde et al. (2011) document strong influence of volatility shocks on real variables like output, consumption, investment and hours worked.

make the variance of the error term in the growth regression dependent on the squared residuals taken from forecasting regressions for government expenditures. Their idea was to use forecast errors as a measure of unobserved shocks. RR do not discuss, however, the correct specification of variance endogeneity or consequences of its misspecification, leaving the issue of endogenous volatility basically implicit.

The problem with the standard specification is the absence of explanatory variables in the equation for the growth volatility. This suggests that more explanatory variables are needed in the conditional variance equation than just forecast errors. Since the volatility term appears among explanatory variables in the growth equation, omitted variables in the conditional variance equation potentially lead to correlation between explanatory variables and the error term in the growth equation. This renders estimation of the feedback effect on economic growth captured by the variance term in the growth equation inconsistent. Inconsistence persists even if there are no omitted variables in the growth equation and if all variables are measured without error.

Our proposal – This note discusses an extension of the original RR model and specifies it as a model of conditional heteroscedasticity in mean (henceforward CH-M). As in RR, there is a growth equation that contains the volatility term and a variance equation. The extension consists in explicit allowing for explanatory variables in the variance equation. Using these additional explanatory variables, the RR approach will continue to remain a highly useful framework to investigate the volatility-growth link.

We demonstrate theoretically that a bias arises in the CH-M model of output growth if relevant control variables are omitted from the conditional variance equation. Thus, for example, neglecting the RR prediction error of government expenditure shocks potentially leads to a systematic bias in the estimated parameters of interest, particularly the one that links volatility and growth. In a simulation based on an example borrowed from the literature (Posch 2011), we show that the above bias is of economic importance for the volatility-growth nexus.

The literature – The most recent literature that follows the empirical setup of RR includes Dawson et al. (2001), Imbs (2007), Edwards and Yang (2009), Ponomareva and Katayama (2010), Posch (2011) and Posch and Wälde (2011), among others. Of these only Edwards and Yang (2009), Posch (2011) and Posch and Wälde (2011) explicitly consider the conditional volatility. Edwards and Yang (2009) analyze spatial differences in the influence of volatility on growth, Posch (2011) and Posch and Wälde (2011) include tax rates and further controls. None of these papers, however, addresses the source of the potential bias in the estimates of the volatility-growth link and its quantitative importance. It is somewhat unfortunate that in the rest of the literature modelling conditional variance has passed unnoticed, whereas exactly this gives rise to the mentioned systematic bias of the estimated effect of volatility on growth.

Remarkably, Dawson et al. (2001) and Ponomareva and Katayama (2010) discuss a related bias which appears in the empirical RR model if some explanatory variables in the growth equation are measured with error. We show that the bias induced by omitted variables in the conditional variance equation can be alternatively represented as an errors-in-variables bias, where volatility term could be considered as a regressor measured with error. Thus both types of biases have similar manifestation. Still an important difference in our case is that if some relevant controls are omitted form the volatility equation the bias will arise even if all other variables included in the growth regression are measured without error.

Our model may be viewed as a special case of the original ARCH-M model of Engle et al. (1987), where the coefficients in front of autoregressive terms in the variance equation are set to zero. As a consequence, only explanatory variables of a current period matter for the variance (see Engle et al., 1987, equation 9, with $\alpha = 0$). Without emphasizing the role of explanatory variables in the variance equation explicitly, Engle et al. (1987) provide the framework that accounts for the bias discussed here.

The outline – The next section presents an augmented CH-M model. Section 3 provides theoretical insights into the existence and the source of the bias. It also conducts Monte-Carlo simulations in order to show the quantitative importance of the bias for the RR estimate of the link between volatility and growth. Section 4 concludes this note.

2 A volatility-growth regression with controls in the conditional variance equation

2.1 The regression setup

Consider the following extension of Ramey and Ramey (1995) borrowed from Posch (2011). We specify the following growth equation and conditional variance equation,

$$\Delta y_{it} = \nu\sigma_{it} + \theta X_{it} + \varepsilon_{it}, \quad \text{where} \quad \varepsilon_{it} \sim N(0, \sigma_{it}^2), \tag{1a}$$

$$\log(\sigma_{it}) = \alpha_i + \mu_t + \beta Z_{it}. \tag{1b}$$

In these equations, Δy_{it} is the growth rate of output for country i in year t, σ_{it} is the standard deviation of the error term in the growth equation; X_{it} is a vector of control variables (e.g., the Levine-Renelt variables); Z_{it} is a vector of control variables (e.g., a subset of X_{it}); α_i and μ_t are country and time fixed effects; θ and β are vectors of coefficients. The key parameter of interest in a volatility-growth analysis is ν, which links growth to volatility.

We will now show the importance of including additional controls in the conditional variance equation in two ways. First, we will demonstrate analytically that omitted control variables induce systematic bias into the maximum likelihood (ML) estimator of the volatility-growth link ν. Second, we will confirm by simulation that this bias is large quantitatively.

2.2 Analytical result

- ML estimation with neglected controls in variance equation

Consider the model in (1a)-(1b), where for simplicity we drop the subscript i and country/time fixed effects, as they do not alter the argument. Let the standard deviation of the error term in the correctly specified growth regression, σ_t, depend on explanatory variables as: $\sigma_t = \exp\{\alpha + \beta Z_t\}$. Once dependence on Z_t is neglected, the same standard deviation in the misspecified model, say $\tilde{\sigma}_t$, will be given just by: $\tilde{\sigma}_t = \exp\{\alpha\}$. Keeping this in mind, growth equation (1a) can be written as

$$\begin{aligned}\Delta y_t &= \nu\sigma_t \pm \nu\tilde{\sigma}_t + \theta X_t + \varepsilon_t \\ &= \nu\tilde{\sigma}_t + \theta X_t + (\varepsilon_t + \nu[\sigma_t - \tilde{\sigma}_t]) \\ &= \nu\tilde{\sigma}_t + \theta X_t + \tilde{\varepsilon}_t,\end{aligned}$$

where $\tilde{\varepsilon}_t \equiv \varepsilon_t + \nu[\sigma_t - \tilde{\sigma}_t]$ and ε_t is not correlated with X_t and Z_t by assumption. Inserting for both σ_t and $\tilde{\sigma}_t$ in this new error term $\tilde{\varepsilon}_t$ we get

$$\tilde{\varepsilon}_t = \varepsilon_t + \nu e^{\alpha}[e^{\beta Z_t} - 1].$$

Consider now estimation of the equation

$$\Delta y_t = \nu\tilde{\sigma}_t + \theta X_t + \tilde{\varepsilon}_t$$

where explicit dependence on Z_t in the variance equation is omitted. Omitting the dependence on Z_t amounts to specifying the error term in this equation identically to that of the original equation (1a), i.e.

$$\Delta y_t = \nu\tilde{\sigma}_t + \theta X_t + u_t, \quad \text{where} \quad u_t \sim N(0, \tilde{\sigma}_t), \tag{2a}$$
$$\log(\tilde{\sigma}_t) = \alpha. \tag{2b}$$

It is straightforward to show (see Appendix) that the maximum likelihood estimator of the parameter ν in the misspecified model (2a)-(2b) has a form

$$\hat{\nu} = T^{-1}\sum_{t=1}^{T}\frac{\Delta y_t - \theta X_t}{\tilde{\sigma}_t}$$

Taking the expected value of $\hat{\nu}$ with respect to the distribution of the dependent variable in the correctly specified model we obtain

$$E(\hat{\nu}) = \nu\, T^{-1}\sum_{t=1}^{T} e^{\beta Z_t}$$

as shown in the Appendix. This expected value is not equal to the true parameter ν unless $\beta = 0$, meaning that $\hat{\nu}$ is biased. Furthermore, the bias does not disappear asymptotically, as

$$\text{plim}\ \hat{\nu} = \nu\, E e^{\beta Z},$$

meaning that unless the true β is equal to zero $\hat{\nu}$ is inconsistent.[3]

Clearly, in a correctly specified model which explicitly considers Z_t in the variance equation the ML estimator of ν has all the standard properties.

- Alternative look at the source of the bias

The above demonstrated bias can also be interpreted as an errors-in-variables bias, where growth volatility could be seen as a regressor measured with error. Assume that by some chance we are able measure the volatility term in the data (e.g. via collecting multiple proxy variables for growth volatility and creating a composite index). Though, our measure can be only imperfect. Once the true measure, σ_t, in the growth equation is substituted by the available imperfect measure, $\tilde{\sigma}_t$, the error term immediately adjusts by the difference between the two, where the difference is completely attributed to the measurement error. Since this difference is a function of Z_t, as the true volatility is the function of Z_t, the new error term will be correlated with X_t, namely

$$Cov\left(X_t, \tilde{\varepsilon}_t\right) = Cov\left(X_t, \varepsilon_t\right) + Cov\left(X_t, \nu e^{\alpha}[e^{\beta Z_t} - 1]\right) = \nu e^{\alpha} Cov\left(X_t, e^{\beta Z_t}\right).$$

Whenever $\beta = 0$ or X_t and Z_t are not stochastically independent it follows that $Cov\left(X_t, \tilde{\varepsilon}_t\right) \neq 0$. Correlation between the error term and the regressors is a common source of the endogeneity bias.

Also note that except of σ_t all other variables, namely X_t and Z_t, are implicitly assumed to be measured correctly, which is different from the analysis of Dawson et al. (2001).

[3] Repeating the steps outlined in the Appendix it is likewise possible to show that the ML estimate of θ in the misspecified model is also biased and inconsistent.

2.3 Monte-Carlo simulation

To provide quantitative support for the above demonstrated bias we simulate our model. As in the analytical discussion we consider the model in (1a)-(1b) suppressing country and time fixed effects for simplicity. Furthermore, again to simplify the simulation, we assume that $X_t = Z_t$, i.e. both growth rate of output and variance of this growth rate are determined by the same set of explanatory variables. We assume Z_t to follow a structure displaying time-variation of the kind shown in Figure 1.[4]

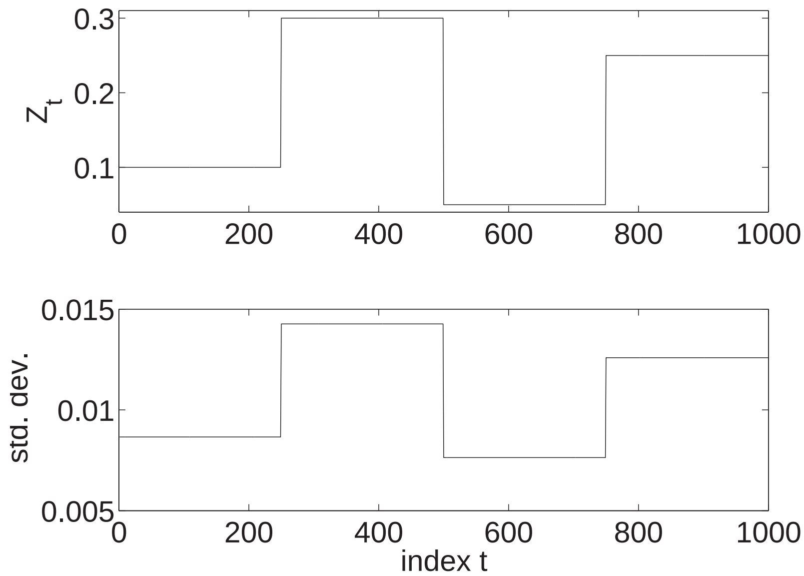

Figure 1 *Our control variable Z_t and the implied standard deviation σ_t of the residuals*

Under these assumptions our model for the simulation becomes

$$\Delta y_t = \nu\sigma_t + \theta Z_t + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim N(0, \sigma_t^2),$$
$$\log(\sigma_t) = \alpha + \beta Z_t.$$

For a set of predetermined parameters ν, θ, α, and β, conditional on the tax vector Z_t and the vector of variances σ_t^2 at any time t, we draw a sample of $T = 1.000$ errors ε_t from the normal distribution $N(0, \sigma_t^2)$. This allows us computing T values for Δy_t. Having done so, we estimate the parameters of the above model by maximum likelihood using the simulated data. The resulting estimates constitute the estimates from a correctly specified model, of which we record the estimated value of ν. Next we consider the misspecified model ignoring Z_t in the variance equation. Estimating by maximum likelihood the misspecified model, we obtain what we call biased parameters. Among these we again record the estimated value of the parameter ν. We repeat this procedure $N = 10.000$ times, which results in N pairs of estimates of ν, first element of this pair being the estimate from the correctly specified and second element - from the misspecified model. After that we plot these estimates of ν against the true values of ν chosen for the simulation. We do not vary the parameters α and β.

[4]This structure was originally motivated by understanding the effect of taxes and tax reforms on growth and volatility. Such a tax vector could reflect three tax reforms over the length of time for which data is available. Tax rates are constant between reforms. The resulting standard deviations in the lower panel show that values are quantitatively reasonable. From a cross-sectional perspective, Z_t (that would then be denoted Z_i) could reflect differences in tax rates accross countries i with tax rates that are time-invariant. Neglecting the cross-sectional variation would then also bias estimates.

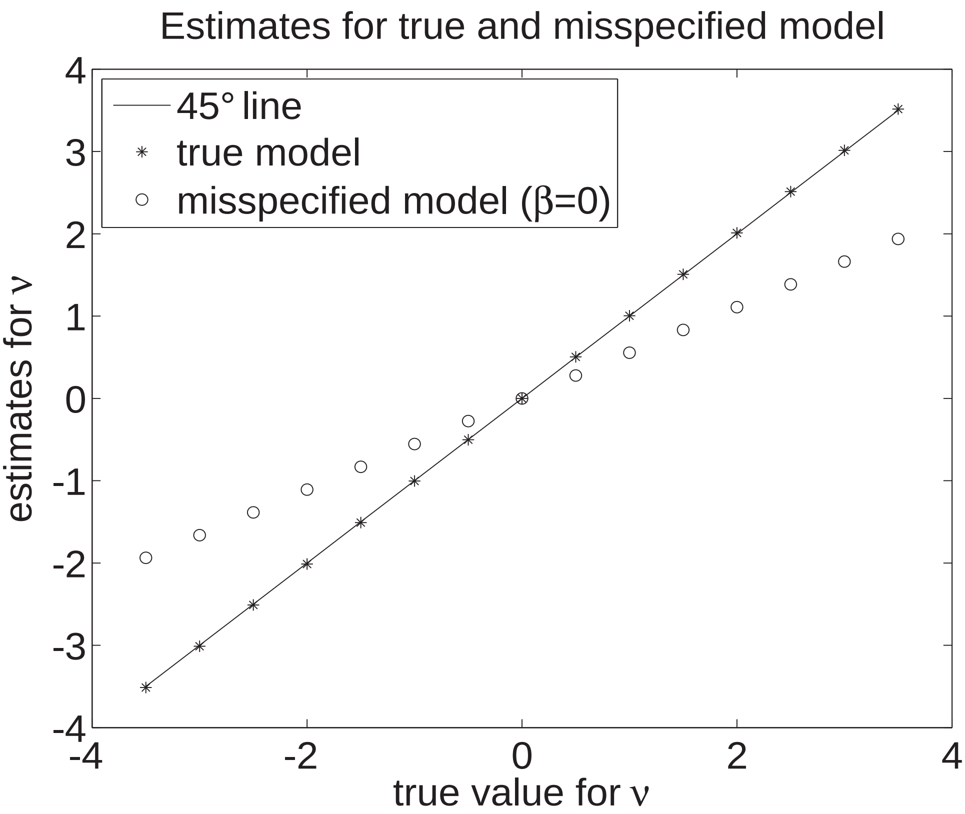

Figure 2 *Estimates for true and misspecified model for various v*

Simulation results are summarized in fig. 2. The horizontal axis of this figure shows the true values for ν used in the simulation. On the vertical axis we plot the estimates from the correctly specified (asterisks) and misspecified (circles) models. In addition, we draw a 45° line to illustrate the equality to the true values of ν. We see that estimates from the correctly specified model are all stretching along the 45° line, whereas estimates from the misspecified model fail to replicate the 45° line by wide margin. The number of replications N has been chosen such that confidence intervals around the estimated values are narrow enough to be neglected. A similar picture emerges though with smaller sample sizes, say $T = 100$. Thus, fig. 2 eloquently tells that if Z is omitted from the variance equation, the estimates of the feedback effect of the volatility on growth can be substantially biased, confirming our analytical result.[5]

3 Conclusion

Economic theory suggests that the degree of volatility of an economy is endogenous. Empirical frameworks that do not account for this endogeneity imply that the estimate for the volatility-growth link is biased. We show this both theoretically and by Monte-Carlo simulations. We suggest that the growth-volatility link should only be estimated if the endogeneity of volatility is sufficiently controlled for by including explanatory variables also in the variance equation.

[5]Although lying beyond the scope of present discussion, we also find that estimates of θ in the misspecified model are biased even more than those of ν.

4 Appendix - ML estimation of the growth-volatility parameter in the misspecified model

Derivation of the estimator – Consider the misspecified model (2a)-(2b). The individual contribution to the likelihood is

$$\ell_t = \frac{1}{\tilde{\sigma}_t\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{\Delta y_t - (\nu\tilde{\sigma}_t + \theta X_t)}{\tilde{\sigma}_t} \right)^2 \right\},$$

and the total log-likelihood reads

$$\log \mathcal{L} = -\frac{T}{2}\log(2\pi) - T\log(\tilde{\sigma}_t) - \frac{1}{2}\sum_{t=1}^{T} \left(\frac{\Delta y_t - \theta X_t}{\tilde{\sigma}_t} - \nu \right)^2 .$$

Taking first order condition with respect to ν we get

$$\frac{\partial \log \mathcal{L}}{\partial \nu} = -\frac{1}{2}\sum_{t=1}^{T} \frac{\partial}{\partial \nu} \left[\left(\frac{\Delta y_t - \theta X_t}{\tilde{\sigma}_t} - \nu \right)^2 \right] = \sum_{t=1}^{T} \left(\frac{\Delta y_t - \theta X_t}{\tilde{\sigma}_t} - \nu \right).$$

Setting this result to zero the ML estimate $\hat{\nu}$ of the true parameter ν in the misspecified model immediately follows

$$\hat{\nu} = T^{-1}\sum_{t=1}^{T} \frac{\Delta y_t - \theta X_t}{\tilde{\sigma}_t}.$$

In this result, the rest of the parameters are for the moment kept as their true unknown values.

Properties of the estimator – Taking the expected value of $\hat{\nu}$ with respect to the distribution of the dependent variable in the true model

$$\Delta y_t = \nu\sigma_t + \theta X_t + \varepsilon_t$$

we obtain

$$\begin{aligned} E(\hat{\nu}) &= T^{-1}\sum_{t=1}^{T} \frac{E(\Delta y_t) - \theta X_t}{\tilde{\sigma}_t} = T^{-1}\sum_{t=1}^{T} \frac{E(\nu\sigma_t + \theta X_t + \varepsilon_t) - \theta X_t}{\tilde{\sigma}_t} \\ &= T^{-1}\sum_{t=1}^{T} \left[\frac{\nu\sigma_t}{\tilde{\sigma}_t} + \frac{E(\varepsilon_t)}{\tilde{\sigma}_t} \right] = \nu T^{-1}\sum_{t=1}^{T} \frac{\exp\{\alpha + \beta Z_t\}}{\exp\{\alpha\}} = \nu T^{-1}\sum_{t=1}^{T} e^{\beta Z_t} \end{aligned}$$

This implies that $E(\hat{\nu}) \neq \nu$ unless $\beta = 0$ in the true model.

Furthermore, for any sequence of random variables $\{Z_t\}_{t=1}^{T}$ with appropriate conditions on the moments (and possibly distribution) of Z_t a corresponding law of large numbers applies and

$$T^{-1}\sum_{t=1}^{T} e^{\beta Z_t} \xrightarrow{p} E\left(e^{\beta Z}\right).$$

as $T \to \infty$. From this follows that

$$\text{plim } \hat{\nu} = \nu \, E e^{\beta Z} \neq \nu$$

unless $\beta = 0$ in the true model.

Research Papers 2012

CREATES
Center for Research in Econometric Analysis of Time Series

2012-04: Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: The Value of Multivariate Model Sophistication: An Application to pricing Dow Jones Industrial Average options

2012-05: Anders Bredahl Kock: On the Oracle Property of the Adaptive LASSO in Stationary and Nonstationary Autoregressions

2012-06: Christian Bach and Matt P. Dziubinski: Commodity derivatives pricing with inventory effects

2012-07: Cristina Amado and Timo Teräsvirta: Modelling Changes in the Unconditional Variance of Long Stock Return Series

2012-08: Anne Opschoor, Michel van der Wel, Dick van Dijk and Nick Taylor: On the Effects of Private Information on Volatility

2012-09: Annastiina Silvennoinen and Timo Teräsvirta: Modelling conditional correlations of asset returns: A smooth transition approach

2012-10: Peter Exterkate: Model Selection in Kernel Ridge Regression

2012-11: Torben G. Andersen, Nicola Fusari and Viktor Todorov: Parametric Inference and Dynamic State Recovery from Option Panels

2012-12: Mark Podolskij and Katrin Wasmuth: Goodness-of-fit testing for fractional diffusions

2012-13: Almut E. D. Veraart and Luitgard A. M. Veraart: Modelling electricity day-ahead prices by multivariate Lévy

2012-14: Niels Haldrup, Robinson Kruse, Timo Teräsvirta and Rasmus T. Varneskov: Unit roots, nonlinearities and structural breaks

2012-15: Matt P. Dziubinski and Stefano Grassi: Heterogeneous Computing in Economics: A Simplified Approach

2012-16: Anders Bredahl Kock and Laurent A.F. Callot: Oracle Inequalities for High Dimensional Vector Autoregressions

2012-17: Eric Hillebrand, Huiyu Huang, Tae-Hwy Lee and Canlin Li: Using the Yield Curve in Forecasting Output Growth and Inflation

2012-18: Eric Hillebrand and Tae-Hwy Lee: Stein-Rule Estimation and Generalized Shrinkage Methods for Forecasting Using Many Predictors

2012-19: Bent Jesper Christensen, Morten Ørregaard Nielsen and Jie Zhu: The impact of financial crises on the risk-return tradeoff and the leverage effect

2012-20: Hendrik Kaufmann, Robinson Kruse and Philipp Sibbertsen: On tests for linearity against STAR models with deterministic trends

2912-21: Andrey Launov, Olaf Posch and Klaus Wälde: On the estimation of the volatility-growth link