Robert Engle is Director of the Volatility Institute of NYU’s Stern School of Business and Professor of Finance, rengle@stern.nyu.edu, 212 998 0702. He thanks many people for comments and suggestions on this paper, particularly participants at Oxford’s Launch of EMoD, SoFiE Singapore, EC², the San Francisco FED, and the University of Chicago Booth Seminar. Particular thanks go to Viral Acharya, Turan Bali, Niel Ericsson, Gene Fama, Eric Ghysels, Jim Hamilton, David Hendry, Alain Monfort, Matt Richardson, Jeff Wooldridge, Rob Capellini, Hahn Le, and Emil Siriwardane. Generous support from the Sloan Foundation, the Armellino Foundation, NASDAQ, Deutsche Bank, Blackrock and A.I.G. are gratefully acknowledged.
ABSTRACT

Dynamic Conditional Beta (DCB) is an approach to estimating regressions with time varying parameters. The conditional covariance matrices of the exogenous and dependent variable for each time period are used to formulate the dynamic beta. Joint estimation of the covariance matrices and other regression parameters is developed. Tests of the hypothesis that betas are constant are non-nested tests and several approaches are developed including a novel nested model. The methodology is applied to industry multifactor asset pricing and to global systemic risk estimation with non-synchronous prices.

Keywords: GARCH, DCC, Time Varying Parameters, Multivariate GARCH, Non-Nested Tests, Multi-factor Asset Pricing, Systemic Risk, SRISK

In empirical finance and in time series applied economics in general, the least squares model is the workhorse. In class there is much discussion of the assumptions of exogeneity, homoskedasticity and serial correlation. However in practice it may be unstable regression coefficients that are most troubling. Rarely is there a credible economic rationale for the assumption that the slope coefficients are time invariant.

Econometricians have developed a variety of statistical methodologies for dealing with time series regression models with time varying parameters. The three most common are rolling window estimates, interaction with trends, splines or economic variables, and state space models where the parameters are treated as a state variable to be estimated by some version of the Kalman Filter. Each approach makes very specific assumptions on the path of the unknown coefficients. The first approach specifies how fast the parameters can evolve, and by using least squares on each moving window, employs an inconsistent set of assumptions. The second specifies a family of deterministic paths for the coefficients that may have undesirable or inconsistent implications particularly when extrapolated. The third requires specifying a stochastic process for the latent vector of parameters which may include unit roots and stochastic
trends or regime shifts that are generally unmotivated and rarely based on any economic analysis.

This paper will propose an approach for a wide class of data generating processes associated with heteroskedastic processes. In addition, it will allow a test of the constancy of the parameter vector.

I. Dynamic Conditional Beta

Consider a vector of observables, \((y_t, x_t)', t = 1, \ldots, T\) where \(x_t' = (x_{t,1}, \ldots, x_{t,k})\). The objective is to characterize the conditional distribution of \(y\) given \(x\). Although we will think of \(y_t\) as a scalar time series, most of the formulae apply directly if it is also a vector. In a time series context this requires finding \(f_{t-1}(y_t | x_t)\) which is the conditional distribution of \(y\) given \(x\) and the past of \(y\) and \(x\). The regression is the mean of this distribution, denoted \(E_{t-1}(y_t | x_t)\). In the very special case we generally assume, this may turn out to be linear in \(x\) with parameters that do not depend upon the past information set. Much more generally, this regression may be linear with parameters that do depend upon the past information set or of course may be a general non-linear function of \(x\) and the past.

A natural formulation of this problem is in terms of the joint distribution of all the variables conditional on the past. Suppose

\[
\begin{pmatrix} y_t \\ x_t \end{pmatrix} \mathbb{P}_{t-1} \sim N\left( \mu_t, H_t \right)
\]

(1)

where \((\mu_t, H_t)\) are measurable with respect to \(\mathbb{P}_{t-1}\), the information set adapted to the past of \((y_t, x_t)\). Then the desired conditional distribution is simply
where subscripts represent natural partitions. Thus the time varying regression coefficients that are needed are

$$\beta_t = H_{xx,t}^{-1} H_{xy,t}$$  (3)

The betas from this expression will be called Dynamic Conditional Betas.

This expression for DCB is very familiar if the time subscript is dropped. In this case we simply have ordinary least squares because conditional moments are the same as unconditional moments. The three other approaches to time varying parameters discussed in the introduction can all be interpreted as forecasts of covariance matrices as in equation (1).

For example if we use rolling regressions, then we are forecasting covariance matrices and betas with “historical” volatilities and correlation. From (1) it is clear that the rolling covariance matrices should not include information from date t. The time width of the window is a central feature of historical volatilities as it regulates the volatility of volatility. In this context the goodness of fit of the model with predicted betas can serve as a criterion for selecting this width. See Engle(2014) for an example. Naturally, the criticisms of historical volatility forecasts apply here as well. That is, a window of k days is the correct choice if the joint vector of returns on the next day is equally likely to come from any of the previous k days but could not be from returns more than k days in the past. Because of the unrealistic nature of this specification, exponential smoothing is a better alternative for most cases.

The normality assumption is obviously restrictive but can be generalized. For example, if the joint distribution in (1) is in the elliptical family of distributions, then (2) holds with the appropriate elliptical distribution. Even without elliptical distributions, (2) can be interpreted as the conditional linear projection. The projection may be linear but with a non-normal error
which will support estimation as in Brownlees and Engle(2011). The more difficult issue is that we do not observe directly either the vector of conditional means or conditional covariances. A model is required for each of these. For many financial applications where the data are returns, the mean is relatively unimportant and attention can be focused on the conditional covariance matrix. Think for a moment of the estimate of the beta of a stock or portfolio, or multiple betas in factor models, style models of portfolios, or the pricing kernel, and many other examples. In these cases, a natural approach is to use a general covariance matrix estimator and then calculate the betas.

It is important to recognize that equation (3) is a natural extension of estimators that have long been used in time series. Probably the first application was Bollerslev, Engle and Wooldridge(1988) where the beta was computed as the ratio of the conditional covariance to the conditional variance. One of the leading applications of multivariate GARCH models has always been estimation of hedge ratios. This is discussed in surveys such as Bollerslev, Chou and Kroner(1994), Bollerslev, Engle and Nelson(2004), Bera and Higgins(1993) and Lien and Tse(2002). Specific hedging models are proposed by Kroner and Sultan(1993), Cecchetti, Cumby and Figlewski(1998), Bera, Garcia and Roh(1997) and Braun, Nelson and Sunier(1995). More recently, hedging performance is used as a test of the multivariate garch specification in Engle(2009), and Lien, Tse and Tsui(2002) among others. Recent applications to asset pricing and systemic risk are in Bali, Engle and Tang(2012) and Brownlees and Engle(2014). When data are available at a higher frequency than the regression analysis, estimates of realized variances, realized covariances and realized betas can be constructed. These realized betas are time varying and can be directly forecast and modeled. See for example Barndorff-Nielsen, O.

All of these examples would be considered as Dynamic Conditional Betas by the description above. Almost all of these applications involve only bivariate models where there is just one covariance to be estimated and the hedge ratio is simply the ratio of the conditional covariance to the conditional variance. In no case is there a statistical test of the null hypothesis that the beta is truly constant. There is no discussion of maximum likelihood estimation in the context of a multivariate regression with time varying coefficients.

This paper will extend the analysis of dynamic conditional betas into a multivariate context and then explicitly consider the estimation and testing of the model from a classical perspective. The tools developed here would have been useful for many of these prior studies.

II. Maximum Likelihood Estimation of DCB
A. All betas are varying

Maximum likelihood estimation of the Dynamic Conditional Beta model requires specification of both the covariance matrix and the mean of the data vector. In general, the covariance matrix and mean of these data will include unknown parameters. Extending equation (1) to include a jx1 vector of unknown parameters in the mean and variance equations, we get

\[
\begin{pmatrix}
    y_t \\
    x_t
\end{pmatrix}_{t-1} \sim N\left( \mu_t(\theta_0), H_t(\theta_0) \right), \quad \theta_0 \in \Theta
\]

(4)

Clearly, this specification includes the wide range of multivariate volatility models that are used in practice such as VEC, BEKK, DCC and general classes of mean functions. In most cases, the model does not satisfy assumptions for weak exogeneity as in Engle Hendry and Richard(1985) because the distribution of the x variables has information about the dynamic conditional betas.
In special cases where the covariance matrix has no unknown parameters or these can be partitioned in special ways, weak exogeneity can be established.

The MLE of $\theta$ can be computed for the data vector $z_t = (y_t, x_t)'$. Letting $\hat{z} = \{z_1, ..., z_T\}$,

$$
\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta; \hat{z}), \quad L(\theta; \hat{z}) = \frac{1}{2} \sum_{t=1}^{T} \log |H(\theta)_t| + (z_t - \mu(\theta)_t)' H(\theta)_t^{-1} (z_t - \mu(\theta)_t)
$$

(5)

Then as $\beta_t$ is a function of the unknown parameters and data, the MLE is simply

$$
\hat{\beta}_t = H(\hat{\theta})^{-1}_x H(\hat{\theta})^{-1}_{xy}
$$

(6)

The limiting distribution of such estimators can be established under certain conditions. In Bollerslev and Wooldridge(1992), Theorem 2.1 establishes conditions for the limiting distribution of $\hat{\theta}$, to be

$$
\sqrt{T} \left( \hat{\theta} - \theta_0 \right) \rightarrow N\left( 0, A^{-1} \right)
$$

(7)

where $A$ could be the information matrix if the normality assumption is valid or a robust covariance matrix of the sandwich form if it is not. This theorem requires correct specification of the first two moments at the true parameter vector, and a variety of regularity conditions. The theorem requires $\theta_0 \in \text{interior } \Theta$ and identifiably unique. $(\mu(\theta), H(\theta))$ must be twice continuously differentiable in $\theta$ for all possible values of $z$, the log likelihood must satisfy a uniform weak law of large numbers, the score vector must satisfy a central limit theorem and the Hessian must satisfy a law of large numbers and converge to a non-singular matrix. Some of these conditions can be weakened, particularly in the context of stationary ergodic random variables.

Define $\beta_{0,t}$ as the true, conditional beta at time $t$ from (6) when $\hat{\theta} = \theta_0$. Applying the mean value theorem to (6) at time $t$,
\[ \hat{\beta}_i - \beta_{0,j} = G_t(\overline{\theta}_i)(\hat{\theta} - \theta_0), \quad G_t(\overline{\theta}_i) = \frac{\partial [H(\theta)^{-1}]_{xy,t} H(\theta)_{yx,t}}{\partial \theta'} \bigg|_{\theta = \overline{\theta}_i} \]  \hspace{1cm} (8)

where \( \overline{\theta}_i \in (\theta_0, \hat{\theta}) \) with possibly different values for different rows and different \( t \). Then

\[ \sqrt{T} (\hat{\beta}_i - \beta_{0,j}) \xrightarrow{D} \mathcal{N} \left( 0, G_t^{-1} G_t' \right) \]  \hspace{1cm} (9)

and \( G_t \) can be estimated as \( G_t(\hat{\theta}) \). Notice that this distribution applies at each point within the sample as the sample size goes to infinity. The large sample result delivers consistent estimates of the unknown fixed parameters \( \theta \) which then imply time varying standard errors for the betas.

A useful expression for the log likelihood can be found by changing variables. Convert the likelihood from \( (y_t, x_t') \) to \( (w_t, x_t') \) where \( w_t = y_t - \mu_t - H_{x,t}^{-1} \left( x_t - \mu_x \right) \). Here \( w \) can be interpreted as the residuals from a DCB regression and \( h_{w,t} \) is the conditional variance of the residuals. As the Jacobian of the transformation is 1, the likelihood simply has a recalculated covariance matrix which is now diagonal giving:

\[ L(\theta; \bar{y}, \bar{x}) = L(\theta; \bar{w}, \bar{x}) = -\frac{1}{2} \sum_{t=1}^{T} \log(h_{w,t}) + \frac{w_t^2}{h_{w,t}} - \frac{1}{2} \sum_{t=1}^{T} \log[H(\theta)^{-1}]_{xx,t} + x_t' H(\theta)^{-1} x_t \]  \hspace{1cm} (10)

B. All betas are constant

In this classical case, the likelihood function can again be evaluated based on equation (1) where the specification includes the assumption that the betas are all constant. However we do not assume that the data or the residuals are homoskedastic. Now we write the log likelihood as the sum of the log of the conditional density of \( y \) given \( x \) plus the log of the marginal density of \( x \). Ignoring irrelevant constants gives:
\[
L(\theta; \hat{y}, \hat{x}) = \sum_{t=1}^{T} \left[ \log f_{y|x}(y_t|x, z_{t-1}, \ldots; \theta) + \log f_{x}(x_t|z_{t-1}, \ldots; \theta) \right]
\]
\[
= -\frac{1}{2} \sum_{t=1}^{T} \left\{ \log \left( h_{u,t} \right) + \frac{u_t^2}{h_{u,t}} \right\} - \frac{1}{2} \sum_{t=1}^{T} \left[ \log \left| H(\theta)_{xx,t} \right| + \left( x_t - \mu_{x,t} \right)' H(\theta)^{-1}_{xx,t} \left( x_t - \mu_{x,t} \right) \right] \tag{11}
\]

In this equation
\[
u_t = y_t - \mu_{y,t} - \left( x_t - \mu_{x,t} \right)' \beta, \quad V_{t-1}(\nu_t) = h_{u,t} \tag{12}
\]

If \( h_u \) is a general volatility process with no common parameters from \( H_{xx} \) then maximum likelihood estimators of beta are simply given by maximizing the first part of the likelihood.

This is just least squares with a heteroskedastic error process.

The asymptotic distribution of beta is simply given by the conventional expression for heteroskedastic regression as in Engle(1982) and more generally Bollerslev and Wooldridge(1992).

C. Some betas are constant and some are varying

If it is known that a subset of the regression parameters are constant, then MLE can be adapted to estimate jointly the fixed and varying parameters of the model. Suppose there are \( k \) regressors and the first \( k_1 \) are time varying and the second \( k_2 \) are constant. In this case the dynamic conditional betas of the first \( k_1 \) elements will be the first \( k_1 \) elements of the expression in (6). Partitioning \( x = (x_1', x_2')' \) and then defining \( x_1^0 = \left( (x_1 - \mu_{x,t})', 0 \right)' \), the model can be expressed as
\[
y_t - \mu_{y,t} - H_{xx,t} H^{-1}_{xx,t} x_1^0 = \beta_2 \left( x_{2,t} - \mu_{x,t} \right) + v_t \tag{13}
\]
Thus the conditional log likelihood can be evaluated by regressing y minus the first terms on the fixed regressors and allowing heteroskedasticity in v. Full maximum likelihood when \((x_1, x_2)\) are correlated may require an iterative algorithm. The result is:

$$L(\theta; \hat{y}, \hat{x}) = -\frac{1}{2} \sum_{t=1}^{T} \left[ \log\left(h_{v,t}\right) + \frac{v_{t}^2}{h_{v,t}} \right] - \frac{1}{2} \sum_{t=1}^{T} \left[ \log\left| H(\theta)_{x,t} \right| + \left( x_t - \mu_{x,t} \right)^T H(\theta)^{-1}_{x,t} \left( x_t - \mu_{x,t} \right) \right]$$

(14)

Notice that the term in curly brackets has the same form in (10) (11) and (14). For many specifications of the volatility process, the estimates of the covariance matrix of \(x\) will be identical for these three specifications, so any comparison of log likelihoods can be done simply by comparing the marginal log likelihoods. The test becomes a test of which residual vector, \((w, u, v)\) has the highest likelihood. Interestingly, the regression models with time varying parameters do not necessarily have the best fit. Remember, the coefficients are the predicted coefficients for each data point.

III. Testing for constant regression coefficients

In this section a set of approaches will be developed for testing that one or more parameters are constant. These hypotheses are basically non-nested although they more precisely are overlapping hypotheses. The overlapping occurs because if there is no conditional variance or covariance in the data set, the null and alternative hypotheses become identical. Thus there are points in the parameter space where the hypotheses are not testable.

There are several standard approaches to testing non-nested hypotheses. The first developed initially by Cox(1961,1962) is a model selection approach which simply asks which model has the highest value of a (possibly penalized) likelihood. If there are two models being
A second approach introduced by Vuong (1989) is to form a two tailed test of the hypothesis that the models are equally close to the true model. In this case, there are three possible outcomes, model 1, model 2, and that models are equivalent. A third approach is artificial nesting as discussed in Cox (1961, 1962) and advocated by Aitkin (1970). An artificial model is constructed that nests both special cases. When this strategy is used to compare two models, there are generally four possible outcomes – model 1, model 2, both, neither.

A). Overlapping Hypotheses

Consider two data generating processes for the same vector of time series

\[ z_t = (y_t', x_t'), \ t = 1, \ldots, T \]  \hspace{1cm} (15)

In the simplest case, there is only one endogenous variable and k conditioning variables. These can be expressed in terms of the joint distributions conditional on past information as:

\[ z_t | \mathbb{F}_{t-1} \sim f_t (z_t, \theta), \ \theta \in \Theta \]
\[ z_t | \mathbb{F}_{t-1} \sim g_t (z_t, \phi), \ \phi \in \Phi \]  \hspace{1cm} (16)

Without loss of generality we can write each distribution conditional on the past as the product of a conditional and a marginal distribution each conditioned on the past.

\[ f_t (z_t; \theta) = f_{x,t} (y_t, x_t; \theta) f_{x,t} (x_t; \theta) \]
\[ g_t (z_t; \phi) = g_{x,t} (y_t, x_t; \phi) g_{x,t} (x_t; \phi) \]  \hspace{1cm} (17)

The maximized log likelihoods of these models are given by

\[ L_f (\hat{\theta}) = \sum_{t=1}^{T} \log f_{x,t} (y_t, \hat{x}_t; \hat{\theta}) + \log f_{x,t} (x_t; \hat{\theta}) \]
\[ L_g (\hat{\phi}) = \sum_{t=1}^{T} \log g_{x,t} (y_t, \hat{x}_t; \hat{\phi}) + \log g_{x,t} (x_t; \hat{\phi}) \]  \hspace{1cm} (18)
Model $f$ is said to be nested within $g$ if for every $\theta_0 \in \Theta$ there is a $\phi^* \in \Phi$ such that $f_1(z, \theta_0) = g_1(z, \phi^*)$ for all $t$, so that events have the same probability whether $f(z, \theta_0)$ or $g(z, \phi^*)$ is the true DGP. A similar definition applies to the notion that $g$ is nested within $f$.

The models $f$ and $g$ are partially non-nested (Cox(1960), 1961), Pesaran(1999)) or overlapping (Vuong(1989)) if there are some parameters in $\Theta \otimes \Phi$ that lead to identical distributions. Vuong proposes a sequential procedure that tests for such parameters first and upon rejecting these cases, examines the purely non-nested case.

In the DCB case, the parameters $\theta$ and $\phi$ include parameters of the covariance matrix that must be estimated as well as parameters of the regression function such as the betas. The parameters of the covariance matrix could be the parameters of a multivariate GARCH model if that is how the estimation is conducted. If a DCC approach is used, then the parameters would include both GARCH parameters and DCC parameters. From equation (3) and knowledge of a wide range of multivariate GARCH models, it is apparent that the only point in the parameter space where the fixed parameter and varying parameter are equivalent is the point where there is no conditional heteroskedasticity in the observables. Even models like constant conditional correlation, BEKK, VEC and DECO have no parameters that will make the betas in (3) constant unless both $H_{xx,f}$ and $H_{xy,f}$ are themselves constant. Thus it is essential to determine whether there is heteroskedasticity in $y$ and $x$ to ensure that this is a non-nested problem.

As introduced above, several approaches to model testing for non-nested models will be applied.

B) Model Selection
From the maximized likelihood in (18) and a penalty function such as Schwarz or Akaike, the selected model can be directly found. The criterion function is the penalized likelihood ratio given by

\[
LR_T = \left\{ \sum_{t=1}^{T} \log \left( f_{z,t} \left( z_t ; \hat{\theta} \right) \right) \right\} - \left\{ \sum_{t=1}^{T} \log \left( g_{z,t} \left( z_t ; \hat{\phi} \right) \right) \right\} - \left( \text{pen}_f - \text{pen}_g \right)
\]  

(19)

A positive value of LR selects \( f \) and a negative value selects \( g \). For the Akaike Information Criterion (AIC), \( \left( \text{pen}_f - \text{pen}_g \right) = \left( k_f - k_g \right) \) and for the Schwarz or Bayesian Information Criterion, (BIC) it is given by \( \left( k_f - k_g \right) * \log (T) / 2 \) Here \( k \) is the total number of estimated parameters in each model. Clearly, as \( T \) becomes large, the penalty becomes irrelevant in model selection.

In many cases the likelihood will have the same estimate of the likelihood of the conditioning variables, \( x \), and consequently the maximized values will satisfy

\[
\left\{ \sum_{t=1}^{T} \log \left( f_{x,t} \left( x_t ; \hat{\theta} \right) \right) \right\} = \left\{ \sum_{t=1}^{T} \log \left( g_{x,t} \left( x_t ; \hat{\phi} \right) \right) \right\}
\]  

(20)

This will happen if the dynamic covariance matrix of the \( x \) variables is estimated the same way regardless of the specification of the dynamics of the regression coefficients. Therefore model selection will only involve the first term in the log likelihood just as in a conventional regression. Thus model selection can be based on

\[
LR_T = \sum_{t=1}^{T} \log \left( f_{y,t} \left( y_t ; \hat{\theta} | x_t \right) \right) - \sum_{t=1}^{T} \log \left( g_{y,t} \left( y_t ; \hat{\phi} | x_t \right) \right) - \left( \text{pen}_f - \text{pen}_g \right)
\]  

(21)

which can be expressed in terms of the residuals as in equations (10), (11) and (14). Model selection between DCB and constant betas can be based on the penalized log likelihood

\[
LR_T = \sum_{t=1}^{T} \left( \log \left( h_{w,t} \right) - \frac{w_t^2}{h_{w,t}} \right) - \sum_{t=1}^{T} \left( \log \left( h_{w,t} \right) - \frac{u_t^2}{h_{u,t}} \right) - \left( \text{pen}_w - \text{pen}_u \right)
\]  

(22)
The penalties will depend upon the number of parameters estimated in each model. Since the parameters in the x matrix are the same and the parameters in the error process are the same, the only difference is the number of fixed betas that are estimated.

C) Testing Model Equality.

Following Vuong(1989) and Rivers and Vuong(2002), the model selection procedure can be given a testing framework. The criterion in (21) has a distribution under the null hypothesis that the two likelihoods are the same. The hypothesis that the likelihoods are the same is interpreted by Vuong as the hypothesis that they are equally close in a Kulbach Leibler sense, to the true DGP. He shows that under this null, the LR criterion converges to a Gaussian distribution with a variance that can be estimated simply from the maximized likelihood ratio statistic. Letting

\[ m_t = \log \left( \frac{f_{y|x,t}(y_t; \hat{\theta}|x_t)}{g_{y|x,t}(y_t; \hat{\phi}|x_t)} \right) \]  \hspace{1cm} (23)

then

\[ \hat{\omega}^2 = \frac{1}{T} \sum_{t=1}^{T} \left( m_t - \frac{LR_T}{T} \right)^2 \]  \hspace{1cm} (24)

and it is shown in Theorem 5.1 that

\[ T^{-1/2} LR_T / \hat{\omega} \xrightarrow{d} N(0,1) \]  \hspace{1cm} (25)

This can be computed as the t-statistic on the intercept in a regression of \( m_t \) on a constant. Based on this limiting argument, three critical regions can be established – select model f, select model g, they are equivalent.
To apply this approach we substitute the log likelihood of the DCB model and the fixed coefficient GARCH model into (23). From (11) and (14) we get

\[ m_t = L_{f,t} - L_{g,t} = -0.5 \left( \log \left( \frac{h_{f,t}}{h_{g,t}} \right) + \varepsilon_{f,t}^2 - \varepsilon_{g,t}^2 \right) \]  

(26)

where the epsilons are the standardized residuals. If a finite sample penalty for the number of parameters is adapted such as the Schwarz criterion, then the definition of m becomes

\[ m_t = -0.5 \left( \log \left( \frac{h_{f,t}}{h_{g,t}} \right) + \varepsilon_{f,t}^2 - \varepsilon_{g,t}^2 \right) - \left( k_f - k_g \right) \log(T) / T \]  

(27)

This series of penalized log differences is regressed on a constant and the robust t-statistic is reported. If it is significantly positive, then model f is preferred and if significantly negative, then model g is preferred. As m in equation (27) is likely to have autocorrelation, the t-statistic should be computed with a HAC standard error that is robust to autocorrelation and heteroskedasticity. See Rivers and Vuong(2002).

D) The third approach is based on artificial nesting and is simply an augmented regression model which allows for heteroskedasticity. From asymptotic standard errors, tests of both f and g null models can be constructed. The simplicity and power of this approach make it appear to be the most useful way to apply the DCB model. The nested model includes the regressors twice, once with a constant coefficient and once with a time varying coefficient as well as a time invariant coefficient multiplied together element by element as a Hadamard product.

\[ y_i = x_i' \zeta + x_i' \left( \beta_i \circ \xi \right) + \sqrt{h_{e,t}} e_i \]  

(28)

If \( \zeta = 0 \), the model is DCB and if \( \xi = 0 \) it is fixed coefficient. T-tests can determine the outcome.

Although this regression has regressors that depend upon estimated parameters, which leads to the generated regressor problem of Pagan(1984), a theorem from Wooldridge(2002)
shows that parameter estimates are consistent and some inference is supported. The nested DCB model can be expressed as

\[ y_i = x_i' \zeta + x_i' (\beta_i (\theta_0) \circ \xi) + \varepsilon_i = (x_i, (x_i \circ \beta_i (\theta_0)))' \phi + \varepsilon_i, \phi = (\zeta, \xi)' \]  

(29)

when the parameter \( \theta = \theta_0 \). The model that is estimated however uses the estimated value of \( \theta = \hat{\theta} \). Wooldridge(2002) establishes in section 12.4 the conditions for consistency and correct inference in this model. Consistency follows simply from the consistency of \( \hat{\theta} \). To establish that inference is correct at \( \phi = \phi_0 \), requires the condition that

\[ E\left( \frac{\partial^2 L_i (\phi, \theta)}{\partial \phi \partial \theta} \right) \bigg|_{\phi = \phi_0, \theta = \theta_0} = 0 \]  

(30)

Under the standard likelihood assumptions and under the condition (7) this term in the Hessian can be expressed as

\[ \frac{\partial^2 L_i (\phi, \theta)}{\partial \phi \partial \theta} \bigg|_{\theta = \theta_0, \phi = \phi_0} = \frac{\partial}{\partial \theta} \left[ \varepsilon_i (x_i, x_i \beta_i (\theta)) \right] = \varepsilon_i (0, x_i G_i (\theta_0)) - (x_i, x_i \beta_i (\theta_0))' x_i G_i (\theta_0) \gamma_0 \]  

(31)

Because G consists only of lagged variables, the first term has expected value zero. The second term will generally be non-zero except when \( \gamma_0 = 0 \). Thus inference on the nested DCB model can be done without correcting the standard errors when \( \gamma_0 = 0 \) which is the fixed beta null hypothesis. Thus the artificially nested model is perfectly appropriate for testing the null hypothesis of fixed betas and is consistent under the alternative, but inference under the alternative should be adjusted. An approach to this inference problem can be developed following Wooldridge but this has not yet been worked out.

IV. Asset Pricing and the DCB
Multifactor asset pricing theories begin with a pricing kernel and then derive cross-sectional and time series implications that can be tested. This derivation is completely consistent with the Dynamic Conditional Beta as described above.

Let \((r, f_t)\) be vectors of nx1 asset returns and kx1 pricing or risk factors respectively. The pricing factors are taken to be tradable with returns given by \(f_t\). A pricing kernel can be specified as

\[
m_t = a_t - b_t \cdot f_t
\]

where the parameters a and b potentially change over time but the changes are not priced. From the pricing kernel it is straightforward to derive the following expression:

\[
1 = E_{t-1} \left( m_t \left( r_t + 1 \right) \right), \quad 1 + r_t' = 1 / E_{t-1} \left( m_t \right)
\]

\[
E_{t-1} \left( r_t - r_t' \right) = (1 + r_t') \text{Cov}_{t-1} \left( r_t, f_t \right) b_t
\]

\[
= \beta_{r,f,t} \left( 1 + r_t' \right) \text{Var}_{t-1} \left( f_t \right) b_t
\]

where \(r_t'\) is the risk free rate and the betas follow the DCB prescription

\[
\beta_{r,f,t} = \text{Cov}_{t-1} \left( r_t, f_t \right) \left[ \text{Var}_{t-1} \left( f_t \right) \right]^{-1}
\]

Assuming the factors are also priced by this pricing kernel

\[
E_{t-1} \left( f_t \right) = \left( 1 + r_t' \right) \text{Var}_{t-1} \left( f_t \right) b_t
\]

so that the factor expected returns is a potentially time varying linear combination of the conditional variances and covariances of the factors. Substituting this into (33) we get the familiar asset pricing result now expressed in terms of DCB

\[
E_{t-1} \left( r_t - r_t' \right) = \beta_{r,f,t} E_{t-1} \left( f_t \right)
\]

This nxk matrix has the betas from regressions of asset returns on factor returns and these betas are conditional on past information. The themes of asset pricing come through completely in that
the expected return on an asset is linear in beta and depends upon the risk premium embedded in the factor and there should be no alpha.

In Bali, Engle and Tang (2012) this empirical model is challenged to explain cross sectional returns in a one factor dynamic model. It performs very well providing new evidence of the power of beta when it is estimated dynamically. Furthermore, they find strong evidence of hedging demand in the conditional ICAPM. In Bali and Engle (2010) a similar model is examined from a pooled time series cross section point of view and again evidence is found for hedging demands. It is found that the most useful second factor is a volatility factor.

V. Asset Pricing for Industry Portfolios

To illustrate the usefulness of the DCB in an asset pricing context, the 12 industry portfolios will be examined in the context of the Fama French 3 factor model. Data are from Ken French’s web site and cover the period 1963-2011 which is over 12,000 daily observations per return series. For each industry the following model is estimated:

\[
r_t^j - r_t^f = \alpha_j + \beta_{j,m} (r_t^m - r_t^f) + \beta_{j,hml} r_t^{hml} + \beta_{j,smb} r_t^{smb} + \sqrt{h_t} \varepsilon_t^j
\]  

(37)

Here \( (r_t^m, r_t^f, r_t^{hml}, r_t^{smb}) \) are the returns on the market, on the risk free rate, on the high minus the low book to market value portfolio, and on the small minus big firm portfolio, respectively. The DCC model is used to estimate the correlations between all variables. For each equation there are 4 variables, the three factors and an endogenous variable. Because the factors are the same for each equation, it is natural to use the same covariance matrix for them in each equation. This is accomplished by estimating a 3x3 DCC once and using it for all equations. Assuming the remaining correlations in a 4x4 specification follow a DCC, then the DCC dynamic parameters from the factors can be applied to the endogenous variables directly. This is in the spirit of the
Composite Likelihood approach to estimating vast covariance matrices of Engle et al (2014). Alternatively, when the correlations with the endogenous variables were estimated individually without constraints, several had persistence exceeding one. The Schwarz criterion uniformly preferred the model with restricted covariance to the model with unrestricted covariances.

The model (37) is therefore estimated under four assumptions:

a) OLS with constant coefficients and robust standard errors
b) GJR-GARCH with constant coefficients
c) DCB with regression coefficients from (3). Correlation matrices estimated with trivariate DCC for the factors. Bivariate DCC is used for the correlations between the dependent variable and each of the factors using the dynamic parameters estimated from the trivariate factor correlation matrix. The GJR-GARCH model is used for each series.
d) NESTED DCB includes each factor both with a constant coefficient and the time varying coefficient.

Table 1 shows the test criteria for choosing between constant betas and time varying betas for each of the 12 industries. The model selection criterion simply chooses the model with the highest value of the penalized likelihood. In this case, all twelve industries are better fit with time varying betas. The Vuong test examines whether these differences are significant. These are t-ratios and consequently at conventional levels, all twelve industries show significant improvement using purely time varying betas.

Table 1 about here

When the nested model is applied there are t-statistics for each beta corresponding to pure time variation and pure fixed beta. The t-statistics in Table 2 indicate that almost all the coefficients are significantly different from zero. This means that the null hypothesis of zero fixed beta as well as the null hypothesis of no time variation are both rejected almost every time.
There is clear evidence that time varying betas add to the explanatory power of this regression. Interestingly, the coefficients on the time varying betas are typically less than one in this application indicating the advantage of some shrinkage.

Table 2 about here

From an asset pricing perspective, it is interesting to learn if the Fama French three factor model is sufficient to price industry portfolios when the DCB model estimates are used. The multifactor pricing model predicts that the alpha of these regressions should be zero if these three factors price these assets. Table 3 presents t-statistics for each of these models from the full time series.

Table 3 about here

It is clear from this table that many of the alphas are significant for particular industries regardless of how the model is estimated. From the NESTED DCB model, 4 industries have significantly negative alphas and one, health, has a significantly positive alpha. This is slightly better than for the fixed parameter GARCH where there are 4 significantly negative and 4 significantly positive alphas.

A few pictures reveal differences across sectors in the time series patterns of their betas. Because there are 12 thousand daily betas, quarterly averages are plotted. From these figures, it is clear that typically the beta on HML is the most variable. Business Equipment has a negative beta on HML in the 90s reaching a beta of -2 and then returning to about -1. This sector therefore trades like a growth stock with time varying intensity. Energy appears to be a value stock from 1982-90 and 2003-2007 but was a growth stock in 1980-82. Money becomes a growth stock in 2004-2007 and then with a dramatic reversal becomes a value stock as its equity values decline during the financial crisis. The market beta also rises for the Money sector
during the financial crisis but not nearly as much as it does in a one factor model. The beta on SMB becomes negative after 1990 indicating that the sector moves like a large cap stock. All the nested betas for all sectors are shown in the figures in the appendix.

Figures 1,2,3 about here

These estimates are useful for risk management and style analysis of a portfolio. ‘The style of a portfolio will change over time as the betas change. Furthermore, the portfolio risk of any factor movement will change as the betas change.

VI. Global Systemic Risk with DCB

A central question in the analysis of systemic risk is the following: “how much capital would a firm need to raise in order to carry on its normal business, if we have another financial crisis?” This is a question of systemic risk because the only source of capital in a financial crisis is likely to be the taxpayer. The taxpayer will consider the consequences of bailing out this institution or letting it fail. In a financial crisis, this institution will not be alone in requesting capital. The bigger the total request the more severe the situation as the real economy cannot function without a viable financial sector. Prudent regulators will foresee this event and require a sufficient cushion that the firms do not need to raise capital, or at least not as much, in a crisis. Financial firms will also foresee this event; however, their costs are limited to their salaries and ownership positions and therefore do not include the costs imposed on the rest of society. This systemic risk externality makes it clear that risks rationally taken by financial firms may be greater than are socially optimal. This argument is developed in more detail in Acharya, Pedersen, Phillipon and Richardson(2010).
To measure the capital shortfall, Brownlees and Engle (2014) propose a time series approach which essentially estimates the beta of a firm’s equity on a broad market index. They estimate a bivariate volatility model between the return on the broad market and the equity return on the firm being analyzed. This model allows volatilities and correlations to evolve and the process is used to simulate the probability of severely negative outcomes over an extended period. Assuming that firms can only operate if capital is a non-trivial proportion of their total liabilities, we define SRISK as the capital that would be needed to achieve a market cap that is 8% of the book value of assets in the event of another crisis. The inputs to SRISK are size, leverage and risk. Each is important and a firm that wants to reduce its SRISK, can operate on any of these characteristics. A crisis is taken to be a 40% drop in global equity values over six months. SRISK is computed weekly for firms from about 70 countries and published on the internet at http://vlab.stern.nyu.edu.

In extending this analysis to international markets, the model is naturally generalized to a multivariate volatility model. The data are considered to be the equity return on one firm, the equity return on a global index of equity returns, and perhaps some lags. The use of daily data for assets that are priced in different time zones means that closing prices are not measured at the same time and consequently there may be some important effect from lagged factor returns.

To adapt the DCB model to this setting requires first adjusting for the timing of returns. In figure 4, closing prices in New York and in a foreign market are plotted on a time line. It is clear that the foreign market will close before New York on day t-1 and again on day t. Consequently the company return will correlate with NY returns on the same day and also on the day before. Thus it appears that there is a lagged effect of NY returns on foreign returns, but this is merely a consequence of non-synchronous trading.
By using an exchange traded fund (ETF) traded in NY which is indexed to a global equity portfolio as the market return $R_{m,t}$, the model can be expressed simply as

$$R_{i,t} - r_t^f = \alpha + \beta R_{m,t} + \gamma R_{m,t-1} + \epsilon_t$$

To estimate this model we wish to allow the betas to vary over time and the DCB is a natural model. Because of the lag however, it is important to reformulate equation (1) by allowing an additional lag. Otherwise the lagged dependent variable will be in the information set.

Assuming the means are zero for ease of notation, the covariance of returns can be expressed as

$$(R_{i,t}, R_{m,t}, R_{m,t-1}) \mid \mathbb{F}_{t-2} \sim N(0, H_t)$$

Thus the DCB equation becomes

$$R_{i,t} = \beta_t R_{m,t} + \gamma_t R_{m,t-1} + u_{i,t}$$

From this expression it is clear that $E_{t-2}(u_t) = 0$ but this means that $u$ potentially has a first order moving average term. In fact, this must be the case if the returns of asset $i$ measured in its own country time is to be a Martingale Difference. From (40) $E(R_{i,t} R_{i,t-1})$ will be non-zero unless $u$ has an MA(1) representation with a negative coefficient just sufficient to offset the autocorrelation induced by the non-synchronous data. As in almost all asset return equations, there is heteroskedasticity so $u$ must have a GARCH–MA(1) representation.

The Nested DCB model is a natural approach to specifying this equation. It has proven useful in the previous example and will be used here.
For each lag, there is both a fixed and a time varying component. The residual is an MA(1) GJR-GARCH.

If global equity returns are serially independent, then an even easier expression is available since the covariance between these two factors is zero. The expression for these coefficients can be written as

$$\begin{bmatrix} \beta_{i,t} \\ \gamma_{i,t} \end{bmatrix} = \begin{bmatrix} H_{R_{m,j}, R_{m,j}} & H_{R_{m,j}, R_{m,j-1}} \\ H_{R_{m,j}, R_{m,j-1}} & H_{R_{m,j-1}, R_{m,j-1}} \end{bmatrix}^{-1} \begin{bmatrix} H_{R_{m,j}, R_{i,t}} \end{bmatrix}$$

which simplifies to

$$\beta_{i,t} = \frac{E\left( R_{i,t} R_{m,j} \big| F_{t-2} \right)}{E\left( R_{m,j}^2 \big| F_{t-2} \right)}, \quad \gamma_{i,t} = \frac{E\left( R_{i,t} R_{m,j-1} \big| F_{t-2} \right)}{E\left( R_{m,j-1}^2 \big| F_{t-2} \right)}$$

There are two non-standard elements in (43). These are the two period lag and the correlation between variables at different times. Because these change little from one day to the next, the following approximations are quite good.

$$\beta_{i,t} \approx \frac{E\left( R_{i,t-1} R_{m,j-1} \big| F_{t-2} \right)}{E\left( R_{m,j-1}^2 \big| F_{t-2} \right)} = \rho_{i,m,t-1} \sqrt{h_{t-1}} \sqrt{h_{m,t-1}}$$

$$\gamma_{i,t} \approx \rho_{i,m-,t-1} \sqrt{h_{t-1}} \sqrt{h_{m,t-1}}$$

where the latter correlation is computed by simply applying DCC to the pair \( R_{i,t}, R_{m,j-1} \) and then lagging the result by one day. If DCC does not converge, then the parameters (.01,.98) are used to essentially exponentially smooth this correlation. With these definitions of \( (\beta_i, \gamma_i) \), the
Nested DCB model in equation (41) is estimated. The beta that is reported is therefore the sum of four numbers. This is natural because the shock is a six month decline in market returns and therefore a two day beta, called $\beta_i^*$, is the starting point.

$$\beta_{i,t}^* = \left( \phi_1 \beta_{i,t} + \phi_2 \right) + \left( \phi_3 \gamma_{i,t} + \phi_4 \right)$$

(46)

This statistical model is now operating in V-LAB where it computes SRISK for almost 1200 global financial institutions weekly and ranks them. The average results from Feb 23, 2015 for 1068 global financial firms are shown in Table 4.

Table 4 around here

The GJR GARCH models look fairly standard. On average $\alpha = .08$ with standard errors one tenth of the coefficient. The asymmetry is estimated as $\gamma = .056$ which is typical for individual stocks. Beta is $.87$ which is may be slightly low but gives substantial persistence. The average DCC parameters are $\alpha = .01$, $\beta = .97$ revealing the high persistence that correlations typically show. The DCC estimated on lags shows similar persistence but much bigger standard errors. In the nested DCB, 77% of the constant betas and 72% of the DCB betas are significant. For the lagged market returns, the results find gamma is significant 59% as a constant and 52% for DCB.

The importance of the lagged beta should depend upon the time zone of the firm, at least if the market has no arbitrage opportunities. So in Table 5 we examine the Nested DCB tests for different continents. In the Americas, the current beta is generally significant both in constant and time varying form while there is much less evidence for a lagged effect through gamma. In Asia, the lagged DCB is more likely to be significant than the current one but the constant part is about the same for current and lagged market returns. In Europe, the current DCB is almost always significant while the lag is only significant about half the time. In Africa the four probabilities are pretty similar.
In summary, the America financials are less sensitive to the previous day return than the other three regions. Over all the DCB is generally significant but so is the constant portion. Only in Europe is there a much stronger effect from the time varying beta than the constant one.

Plots of the betas of the four highest SRISK European banks will serve to illustrate the effectiveness of the DCB model. In each case the beta is the sum of the current and the lagged beta from the nested model as in (46).

The beta of BNP Paribas averaged about 1.5 since 2000, however it rose dramatically to almost 3 during the European financial crisis in 2011-2013. This makes the bank more vulnerable to financial downturns. However BNP beta showed little movement during the 2007-2009 financial crisis. Deutsche Bank beta averaged about 1.50 but it rose to more than 2.50 as the sovereign financial crisis became the dominant news in global equity markets. The rise in both 2008 and 2003 was only to approximately 2. Credit Agricole reached betas of 3.5 during the sovereign debt crisis but these have now returned to around 1.5. Similar observations apply to Barclays. Interestingly, only Barclays showed a comparable rise in beta during the financial crisis of 2007-2009. The SRISK of all European firms is calculated from these betas and from the leverage and market cap of each firm. Adding these together indicates the systemic risk facing Europe over time. Clearly it has declined substantially as shown in Figure 9, both from the financial crisis of 2007-9 and from the sovereign financial crisis of 2011-13. However it remains stubbornly substantially higher than in 2007 so there is more to be done to stabilize the European financial sector.

Figure 5, Figure 6, Figure 7, Figure 8, Figure 9 about here
IX Conclusion

This paper has introduced a new method to estimate time series regressions that allow for time variation in the regression coefficients called Dynamic Conditional Beta or DCB. The approach makes clear that the regression coefficient should be based on the predicted covariances of endogenous and exogenous variables as well as potentially the predicted means. Estimation of the model is discussed in a likelihood context with some betas varying and others constant. Testing the constancy of beta is a non-nested hypothesis and several approaches are suggested and implemented. The most attractive appears to be artificial nesting which motivates the NESTED DCB model where both constant and time varying coefficients are introduced.

The model is applied in two contexts, multi-factor asset pricing and systemic risk assessment. The Fama French three factor model is applied to industry portfolios and the fixed beta assumption is tested. There is strong evidence that the betas are time varying. From plots it is apparent that the HML beta is the most volatile. In the global systemic risk context a dynamic two factor model is estimated which allows foreign equities to respond to current as well as lagged global prices which is expected because of non-synchronous markets. As a result, betas for global financial institutions can be followed over time. The big European banks have had rising betas over the fall of 2011 as the sovereign debt crisis grows in strength. This is an input to the NYU Stern Systemic Risk Ranking displayed on V-LAB which documents the serious nature of systemic risk in 2012 and its moderate improvement in 2014.
REFERENCES


Figure A.1
Nested Betas for Three Factors by Sector
Figure A.2
Nested Betas for Three Factors by Sector
Table 1

Model Selection Test for Constant Betas

<table>
<thead>
<tr>
<th>NAME</th>
<th>SCHWARZ_DCB</th>
<th>SCHWARZ_GARCH</th>
<th>VUONGTEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buseq</td>
<td>1.597089</td>
<td>1.706647</td>
<td>9.726441</td>
</tr>
<tr>
<td>Chems</td>
<td>1.097934</td>
<td>1.131116</td>
<td>3.333880</td>
</tr>
<tr>
<td>Durbl</td>
<td>1.835669</td>
<td>1.893769</td>
<td>6.559011</td>
</tr>
<tr>
<td>Enrgy</td>
<td>1.973194</td>
<td>2.076260</td>
<td>9.118627</td>
</tr>
<tr>
<td>Hlth</td>
<td>1.365149</td>
<td>1.433991</td>
<td>6.551288</td>
</tr>
<tr>
<td>Manuf</td>
<td>0.291902</td>
<td>0.348881</td>
<td>4.329228</td>
</tr>
<tr>
<td>Money</td>
<td>0.716225</td>
<td>0.875159</td>
<td>11.95807</td>
</tr>
<tr>
<td>Nodur</td>
<td>0.669520</td>
<td>0.873242</td>
<td>15.94370</td>
</tr>
<tr>
<td>Other</td>
<td>0.559703</td>
<td>0.620909</td>
<td>3.983843</td>
</tr>
<tr>
<td>Shops</td>
<td>1.113731</td>
<td>1.172615</td>
<td>5.679295</td>
</tr>
<tr>
<td>Telcm</td>
<td>1.738182</td>
<td>1.787497</td>
<td>5.564302</td>
</tr>
<tr>
<td>Utils</td>
<td>0.894953</td>
<td>1.011304</td>
<td>8.349354</td>
</tr>
</tbody>
</table>

Schwarz information criteria are computed for DCB and for constant beta. These are the negative of the log likelihood so smaller is better. The Vuong test from equation (28) is a t-test where positive values favor DCB. Tests are computed from EViews computed regressions with DCB on all coefficients in column 2 and with constant coefficients in column 3. Data are daily from 1963 to 2011 from Ken French’s website.
In the nested model of equation (28) the t-statistics of the factor with constant beta are given first and the t-statistics of the time varying betas are given second. Factor 1 is mkt_rf, Factor 2 is HML and Factor 3 is SMB. Tests are computed from EViews computed regressions of nested DCB with GJR-GARCH residuals. Data are daily from 1963 to 2011 from Ken French’s website.

<table>
<thead>
<tr>
<th>NAME</th>
<th>FIX1</th>
<th>FIX2</th>
<th>FIX3</th>
<th>DCB1</th>
<th>DCB2</th>
<th>DCB3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buseq</td>
<td>22.2</td>
<td>-10.0</td>
<td>-3.25</td>
<td>36.0</td>
<td>38.0</td>
<td>30.8</td>
</tr>
<tr>
<td>Chems</td>
<td>16.8</td>
<td>-2.43</td>
<td>-9.45</td>
<td>25.7</td>
<td>26.9</td>
<td>9.16</td>
</tr>
<tr>
<td>Durbl</td>
<td>19.0</td>
<td>9.61</td>
<td>1.05</td>
<td>22.6</td>
<td>28.8</td>
<td>25.5</td>
</tr>
<tr>
<td>Enrgy</td>
<td>18.7</td>
<td>0.55</td>
<td>-7.28</td>
<td>29.0</td>
<td>46.8</td>
<td>16.2</td>
</tr>
<tr>
<td>Hlth</td>
<td>14.7</td>
<td>-10.1</td>
<td>-11.0</td>
<td>33.6</td>
<td>34.9</td>
<td>15.7</td>
</tr>
<tr>
<td>Manuf</td>
<td>18.5</td>
<td>8.33</td>
<td>10.4</td>
<td>21.9</td>
<td>33.0</td>
<td>19.1</td>
</tr>
<tr>
<td>Money</td>
<td>16.0</td>
<td>6.76</td>
<td>10.8</td>
<td>52.6</td>
<td>42.4</td>
<td>32.9</td>
</tr>
<tr>
<td>Nodur</td>
<td>13.6</td>
<td>-2.10</td>
<td>-3.09</td>
<td>40.0</td>
<td>35.2</td>
<td>45.4</td>
</tr>
<tr>
<td>Other</td>
<td>24.7</td>
<td>3.28</td>
<td>14.5</td>
<td>25.3</td>
<td>38.9</td>
<td>37.1</td>
</tr>
<tr>
<td>Shops</td>
<td>18.6</td>
<td>-6.08</td>
<td>2.37</td>
<td>32.3</td>
<td>27.2</td>
<td>17.5</td>
</tr>
<tr>
<td>Telcm</td>
<td>15.5</td>
<td>5.15</td>
<td>-10.5</td>
<td>26.2</td>
<td>27.4</td>
<td>12.4</td>
</tr>
<tr>
<td>Utils</td>
<td>17.7</td>
<td>7.25</td>
<td>3.00</td>
<td>43.2</td>
<td>40.5</td>
<td>34.8</td>
</tr>
</tbody>
</table>
Table 3

Testing that Alpha is Zero in the Three Factor FF Model

<table>
<thead>
<tr>
<th>NAMEN</th>
<th>TSTAT_DC</th>
<th>TSTAT_NEST</th>
<th>TSTAT_GARCH</th>
<th>TSTAT_OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buseq</td>
<td>0.54</td>
<td>0.98</td>
<td>1.63</td>
<td>2.46</td>
</tr>
<tr>
<td>Chems</td>
<td>-0.67</td>
<td>-0.12</td>
<td>0.76</td>
<td>0.87</td>
</tr>
<tr>
<td>Durbl</td>
<td>-2.60</td>
<td>-2.81</td>
<td>-2.70</td>
<td>-2.52</td>
</tr>
<tr>
<td>Energy</td>
<td>0.06</td>
<td>0.54</td>
<td>2.16</td>
<td>0.85</td>
</tr>
<tr>
<td>Hlth</td>
<td>2.81</td>
<td>3.83</td>
<td>5.04</td>
<td>3.42</td>
</tr>
<tr>
<td>Manuf</td>
<td>-3.27</td>
<td>-3.06</td>
<td>-1.82</td>
<td>-0.35</td>
</tr>
<tr>
<td>Money</td>
<td>-2.69</td>
<td>-2.76</td>
<td>-2.01</td>
<td>-3.10</td>
</tr>
<tr>
<td>Nodur</td>
<td>1.19</td>
<td>1.71</td>
<td>4.15</td>
<td>3.74</td>
</tr>
<tr>
<td>Other</td>
<td>-1.47</td>
<td>-1.91</td>
<td>-2.60</td>
<td>-2.75</td>
</tr>
<tr>
<td>Shops</td>
<td>0.65</td>
<td>1.19</td>
<td>2.08</td>
<td>1.76</td>
</tr>
<tr>
<td>Telcm</td>
<td>-0.88</td>
<td>-0.38</td>
<td>0.69</td>
<td>-0.14</td>
</tr>
<tr>
<td>Utils</td>
<td>-3.17</td>
<td>-3.23</td>
<td>-2.55</td>
<td>-0.39</td>
</tr>
</tbody>
</table>

T-statistics are presented on the intercept of a factor model with four different estimates. First is DCB, then Nested DCB, then constant coefficient regression with GARCH residuals, and finally, constant coefficients with robust standard errors. Tests are computed from EViews computed regressions. Data are daily from 1963 to 2011 from Ken French's website.
### Table 4
Mean Results for Nested DCB model for Global Financials

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VOLATILITY</strong></td>
<td></td>
</tr>
<tr>
<td>$\text{Vol } \omega$</td>
<td>2.5E-05</td>
</tr>
<tr>
<td>$\text{Vol } \omega \text{ se}$</td>
<td>2.28E-06</td>
</tr>
<tr>
<td>$\text{Vol } \alpha$</td>
<td>0.082251</td>
</tr>
<tr>
<td>$\text{Vol } \alpha \text{ se}$</td>
<td>0.007339</td>
</tr>
<tr>
<td>$\text{Vol } \gamma$</td>
<td>0.055865</td>
</tr>
<tr>
<td>$\text{Vol } \gamma \text{ se}$</td>
<td>0.010552</td>
</tr>
<tr>
<td>$\text{Vol } \beta$</td>
<td>0.867158</td>
</tr>
<tr>
<td>$\text{Vol } \beta \text{ se}$</td>
<td>0.007767</td>
</tr>
<tr>
<td><strong>DCC</strong></td>
<td></td>
</tr>
<tr>
<td>$\text{DCC } \alpha$</td>
<td>0.013313</td>
</tr>
<tr>
<td>$\text{DCC } \alpha \text{ se}$</td>
<td>0.022629</td>
</tr>
<tr>
<td>$\text{DCC } \beta$</td>
<td>0.972083</td>
</tr>
<tr>
<td>$\text{DCC } \beta \text{ se}$</td>
<td>0.058646</td>
</tr>
<tr>
<td><strong>DCC LAGGED</strong></td>
<td></td>
</tr>
<tr>
<td>$\text{DCC } \alpha(-1)$</td>
<td>0.009681</td>
</tr>
<tr>
<td>$\text{DCC } \alpha(-1) \text{ se}$</td>
<td>0.014452</td>
</tr>
<tr>
<td>$\text{DCC } \beta(-1)$</td>
<td>0.970297</td>
</tr>
<tr>
<td>$\text{DCC } \beta(-1) \text{ se}$</td>
<td>0.057834</td>
</tr>
<tr>
<td><strong>NESTED DCB</strong></td>
<td></td>
</tr>
<tr>
<td>$\text{DCB Const}$</td>
<td>0.240926</td>
</tr>
<tr>
<td>$\text{DCB Const \text{ se}}$</td>
<td>0.063868</td>
</tr>
<tr>
<td>$\text{DCB Const \text{ sig}}$</td>
<td>0.770599</td>
</tr>
<tr>
<td>$\text{DCB coef}$</td>
<td>0.457231</td>
</tr>
<tr>
<td>$\text{DCB coef \text{ se}}$</td>
<td>0.157594</td>
</tr>
<tr>
<td>$\text{DCB coef \text{ sig}}$</td>
<td>0.72191</td>
</tr>
<tr>
<td>$\text{DCB(-1) Const}$</td>
<td>0.136385</td>
</tr>
<tr>
<td>$\text{DCB(-1) Const \text{ se}}$</td>
<td>0.058848</td>
</tr>
<tr>
<td>$\text{DCB(-1) Const \text{ sig}}$</td>
<td>0.595506</td>
</tr>
<tr>
<td>$\text{DCB(-1) coeff}$</td>
<td>0.374818</td>
</tr>
<tr>
<td>$\text{DCB(-1) coeff \text{ se}}$</td>
<td>0.241127</td>
</tr>
<tr>
<td>$\text{DCB(-1) coeff \text{ sig}}$</td>
<td>0.520599</td>
</tr>
</tbody>
</table>

Results for parameters and standard errors for Global Financials estimated on 2/23/15 with data beginning in 2000 when available. Averages of 1068 individual models are reported. The Volatility model for each equity series is GJR GARCH. DCC is estimated pairwise with MSCI World Index, ACWI, an exchange traded fund. Lagged DCC estimates the correlation between equity return and previous day world index return. Rows labeled sig record the frequency that coefficient/standard error exceeds 1.64 and is therefore significant at a 5% level with a one tailed test.
Table 5

Rejection Frequencies by Region for Nested DCB

<table>
<thead>
<tr>
<th>REGION</th>
<th>DCB const sig</th>
<th>DCB coef sig</th>
<th>DCB(-1) const sig</th>
<th>DCB(-1) coef sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASIA</td>
<td>0.68</td>
<td>0.433</td>
<td>0.673</td>
<td>0.578</td>
</tr>
<tr>
<td>EUROPE</td>
<td>0.763</td>
<td>0.95</td>
<td>0.661</td>
<td>0.537</td>
</tr>
<tr>
<td>AMERICAS</td>
<td>0.95</td>
<td>0.904</td>
<td>0.352</td>
<td>0.377</td>
</tr>
<tr>
<td>AFRICA</td>
<td>0.739</td>
<td>0.739</td>
<td>0.608</td>
<td>0.652</td>
</tr>
</tbody>
</table>

The fraction of firms in each region for which the 5% one tailed t-test rejects that an effect is zero in the Nested DCB regression. DCB const is the constant beta model, DCB coef is the DCB beta model, DCB(-1) is the constant gamma model and DCB(-1) coef is the DCB gamma model as specified in equation (41).
Business Equipment Dynamic Conditional Betas
Computed from Nested DCB model
Figure 1
Dynamic Conditional Betas for Energy
Computed from Nested DCB
Figure 2
Dynamic Conditional Betas for Money
Computed from Nested DCB
Figure 3
Time Line for Non-Synchronous Global Markets

Figure 4.
Beta of BNP Paribas SA
Source – V-LAB

Figure 5
Beta of Deutsche Bank AG
Source – V-LAB

Figure 6
Beta of Credit Agricole
Source – V-LAB

Figure 7
Beta of Barclays PLC
Source – V-LAB

Figure 8
Figure 9

Total SRISK in US $ of European Financials
Source – V-LAB