Stochastic Volatility Demand Systems*

Apostolos Serletis† and Maksim Isakin
Department of Economics
University of Calgary
Canada

Forthcoming in: *Econometric Reviews*

September 22, 2014

Abstract:

We address the estimation of stochastic volatility demand systems. In particular, we relax the homoscedasticity assumption and instead assume that the covariance matrix of the errors of demand systems is time-varying. Since most economic and financial time series are nonlinear, we achieve superior modeling using parametric nonlinear demand systems in which the unconditional variance is constant but the conditional variance, like the conditional mean, is also a random variable depending on current and past information. We also prove an important practical result of invariance of the maximum likelihood estimator with respect to the choice of equation eliminated from a singular demand system. An empirical application is provided, using the BEKK specification to model the conditional covariance matrix of the errors of the basic translog demand system.

JEL classification: C32, E52, E44.

Keywords: Flexible functional forms; Volatility.

---

*We would like to thank two anonymous referees for comments that greatly improved the paper.

†Corresponding author. Phone: (403) 220-4092; Fax: (403) 282-5262; E-mail: Serletis@ucalgary.ca; Web: http://econ.ucalgary.ca/serletis.htm.

1 Introduction

The measurement of consumer preferences and the estimation of demand systems has been one of the most interesting and rapidly expanding areas of recent research. Following Diewert's (1971) influential paper, a large part of the empirical demand literature has taken the approach of using a flexible functional form for the underlying aggregator function. Flexible functional forms have revolutionized microeconometrics, by providing access to all neoclassical microeconomic theory in econometric applications. They include the locally flexible generalized Leontief of Diewert (1971), translog of Christensen *et al.* (1975), Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980), Quadratic AIDS (QUAIDS) of Banks *et al.* (1997), Minflex Laurent (ML) model of Barnett (1983), Normalized Quadratic (NQ) of Diewert and Wales (1988), exact affine Stone index (EASI) implicit Marshallian demand system recently introduced by Lewbel and Pendakur (2010), and the globally flexible Fourier functional form of Gallant (1981) and Asymptotically Ideal Model (AIM) of Barnett and Jonas (1983) and Barnett and Yue (1988). See Lewbel (1997) and Barnett and Serletis (2008) for an up-to-date survey of the state-of-the art in consumer demand analysis.

In recent years, economists and finance theorists have also been creating new models in which stochastic variables are assumed to have a time-dependent variance (and are called 'heteroscedastic', as opposed to 'homoscedastic'). In fact, recent leading-edge research in financial econometrics has applied the autoregressive conditional heteroscedasticity (ARCH) model, developed by Engle (1982), to estimate time-varying variances in commodity prices. Other models include Bollerslev's (1986) generalized ARCH (GARCH) model and Nelson's (1991) exponential GARCH (EGARCH) model. The univariate volatility models have also been generalized to the multivariate case. The multivariate models are similar to the univariate ones, except that they also specify equations of how the conditional covariances and correlations move over time. See Bauwens *et al.* (2006) and Silvennoinen and Teräsvirta (2011) for surveys of this literature.

In this paper we merge the empirical demand systems literature with the recent financial econometrics literature. Our primary interest lies in the estimation of stochastic volatility demand systems. In particular, we relax the homoscedasticity assumption and instead assume that the covariance matrix of the errors of demand systems is time-varying. Since most economic and financial time series are nonlinear and time-varying, we expect to achieve superior modeling using parametric nonlinear demand systems in which the unconditional variance is constant but the conditional variance, like the conditional mean, is also a random variable depending on current and past information.

Although most of the applications in the literature on consumer demand systems use cross sections (or a set of pooled cross sections), there is increasing interest in the methods dealing with time series data. For example, Lewbel and Ng (2005) propose a modification of the translog demand system that can be applied in the presence of nonstationary prices with possibly nonstationary errors. The focus of our paper is on the estimation of demand systems with explicit account for intertemporal structure of (unobserved) error term volatility. In particular, we analyze the simultaneous estimation of a demand system and a GARCH specification for the error term using time series data. While empirical models often reveal heteroscedasticity, the sources of the volatility dynamics might be different. For example, the intertemporal structure of volatility might be a result of the aggregation of heterogeneous preferences. Granger (1980) shows how the aggregation of simple dynamic preferences might lead to complicated long-memory processes. At the same time, the current cross section demand literature focuses on dealing with unobserved heterogeneity — see, for example, Hoderlein (2011), Blundell, Kristensen and Matzkin (2013) and Hoderlein and Stoye (2014). In some functional (Gorman) forms aggregable in income, like the AIDS, unobserved heterogeneity might also result in heteroscedastic error terms [see Lewbel (2001)], albeit this type of heteroscedasticity is beyond the scope of our paper.

The paper is organized as follows. Section 2 briefly reviews the neoclassical theory of consumer choice whereas Section 3 provides a discussion of stochastic volatility demand systems. Section 4 considers a model based on the Baba, Engle, Kraft, and Kroner (BEKK) representation [see Engle and Kroner (1995)] for the conditional covariance matrix of the basic translog demand system. Section 5 provides an empirical application of this model using monthly data on monetary asset quantities and their user costs recently produced by Barnett *et al.* (2013) and maintained within the Center of Financial Stability (CFS) program Advances in Monetary and Financial Measurement (AMFM). The final section concludes.

2 Neoclassical Demand Theory

Consider n consumption goods that can be selected by a consuming household. The household's problem is

$$\max_{\boldsymbol{x}} u(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{p}'\boldsymbol{x} = y$$

where $\boldsymbol{x}$ is the $n \times 1$ vector of goods; $\boldsymbol{p}$ is the corresponding vector of prices; and y is the household's total expenditure on goods (often just called 'nominal income' in this literature). The solution of the first-order conditions for utility maximization are the Marshallian ordinary demand functions,

$$\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{p}, y).$$

Demand systems are often expressed in budget share form, $\boldsymbol{s} = (s_1, ..., s_n)'$, where $s_i = p_i x_i(\boldsymbol{p}, y)/y$ is the expenditure share of good i.

The maximum level of utility at given prices and income, $h(\boldsymbol{p}, y) = u\left[\boldsymbol{x}(\boldsymbol{p}, y)\right]$, is the indirect utility function. The direct utility function and the indirect utility function are equivalent representations of the underlying preference preordering. Using $h(\boldsymbol{p}, y)$, we can derive the demand system by straightforward differentiation, without having to solve a system of simultaneous equations, as would be the case with the direct utility function first order conditions. In particular, Diewert's (1974, p. 126) modified version of Roy's identity,

$$s_i(\boldsymbol{v}) = \frac{v_j \nabla h(\boldsymbol{v})}{\boldsymbol{v}' \nabla h(\boldsymbol{v})}, \qquad i = 1, ..., n \tag{1}$$

can be used to derive the budget share equations, where $\boldsymbol{v} = [v_1, ..., v_n]'$ is a vector of expenditure normalized prices, with the jth element being $v_j = p_j/y$, and $\nabla h(\boldsymbol{v}) = \partial h(\boldsymbol{v})/\partial \boldsymbol{v}$. See Barnett and Serletis (2008) for more details.

Suppose that the indirect utility function h and, therefore, the share equation system, is defined up to the set of parameters $\boldsymbol{\theta}$. In order to estimate share equation systems using observations of prices and shares $(\boldsymbol{v}_t, \boldsymbol{s}_t)_{t=1}^T$, a stochastic version is specified. Also, since only exogenous variables appear on the right-hand side of such systems, it seems reasonable to assume that at time t the observed share in the ith equation deviates from the true share by an additive disturbance term ϵ_{it}. Thus, the share equation system at time t is written in matrix form as

$$\boldsymbol{s}_t = \boldsymbol{s}(\boldsymbol{v}_t, \boldsymbol{\theta}) + \boldsymbol{\epsilon}_t \tag{2}$$

where $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, ..., \epsilon_{nt})'$ and $\boldsymbol{\theta}$ is the parameter vector to be estimated. It has also been typically assumed that

$$\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Omega}) \tag{3}$$

where $\mathbf{0}$ is n-dimensional null vector and $\boldsymbol{\Omega}$ is the $n \times n$ symmetric positive definite error covariance matrix.

Finally, since the demand system (1) satisfies the adding-up property, i.e., the budget shares sum to 1, the error covariance matrix $\boldsymbol{\Omega}$ is singular. Barten (1969) shows that maximum likelihood estimates can be obtained by arbitrarily dropping any equation in the system. McLaren (1990) also establishes invariance by virtue of observational equivalence of the subsystems with different deleted equations.

3 Stochastic Volatility Demand Systems

In this paper, we relax the homoscedasticity assumption in (3) and instead assume that

$$\boldsymbol{\epsilon}_t \sim F(\mathbf{0}, \boldsymbol{\Omega}_t) \tag{4}$$

where F is a distribution from the family of elliptical distributions and $\boldsymbol{\Omega}_t$ is the time-varying covariance matrix of the errors.

As before, the error terms of the demand system sum to zero, $(\boldsymbol{\Omega}_t)_{t=1}^T$ are singular and we can drop one equation to avoid singularity. The natural question is whether it matters which equation is deleted. The following theorem establishes the result.

Theorem 1 *Let the rank of the covariance matrices* $\boldsymbol{\Omega}$ *be* $n-1$. *Then any two subsystems obtained from (2) by deleting different conditional mean equations, under assumption (4), are observationally equivalent, in the sense that they have the same likelihood for any possible sample.*

The proof is given in Appendix A. Theorem 1 is the counterpart of the invariance claim in Barten (1969) and McLaren (1990) for the case when the covariance matrix varies over time. The theorem posits that it does not matter which equation is eliminated from the demand system (2) in order to avoid singularity. Moreover, under assumption (4), the covariance matrices of the different subsystems are related by equation (17) in Appendix A.

Although Theorem 1 is a nice theoretical result, it is not very attractive empirically, because it assumes no intertemporal dependence between the covariance matrices $(\boldsymbol{H}_t)_{t=1}^T$. If one introduces a particular GARCH representation for the covariance matrix, the observational equivalence does not hold in general, although it does hold for some particular cases. In other words, whether invariance holds depends on the intertemporal structure imposed on the covariance matrix. In the next section we show that the invariance result is correct for the GARCH BEKK specification. In particular, we consider a case, with the conditional mean equation given by the basic translog demand system and the conditional variance equation parameterized by a GARCH BEKK.

4 A Specific Case

Consider the basic translog (BTL) reciprocal indirect utility function of Christensen *et al.* (1975)

$$\ln h(\boldsymbol{v}) = \alpha_0 + \sum_{k=1}^{n} \alpha_k \ln v_k + \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} \gamma_{jk} \ln v_k \ln v_j \tag{5}$$

where $\boldsymbol{\Gamma} = [\gamma_{ij}]$ is an $n \times n$ symmetric matrix of parameters and $\boldsymbol{\alpha} = (\alpha_0, ..., \alpha_n)$ is a vector of other parameters, for a total of $(n^2 + 3n + 2)/2$ parameters. The BTL share equations, derived using the logarithmic form of Roy's identity, are

$$s_i = \frac{\alpha_i + \displaystyle\sum_{k=1}^{n} \gamma_{ik} \log v_k}{\displaystyle\sum_{k=1}^{n} \alpha_k + \sum_{k=1}^{n} \sum_{j=1}^{n} \gamma_{jk} \log v_k}, \qquad i = 1, \ldots, n. \tag{6}$$

Estimation of (6) requires some parameter normalization, as the share equations are homogeneous of degree zero in the α's. Usually the normalization $\sum_{i=1}^{n} \alpha_k = 1$ is used.

As before, consider a stochastic version of the demand system (in matrix form)

$$\boldsymbol{s}_t = \boldsymbol{s}(\boldsymbol{v}_t, \boldsymbol{\theta}) + \boldsymbol{\epsilon}_t \tag{7}$$

where $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, ..., \epsilon_{nt})'$ is an additive disturbance term and $\boldsymbol{\theta}$ is the parameter vector to be estimated. Assume that the n-dimensional error vector is normally distributed with zero mean and time-varying covariance matrix

$$\boldsymbol{\epsilon}_t | I_{t-1} \sim N(\mathbf{0}, \boldsymbol{\Omega}_t) \tag{8}$$

where $\boldsymbol{\Omega}_t$ is measurable with respect to information set I_{t-1}. To avoid singularity, delete (any) one equation from (7) and consider the corresponding $(n-1) \times (n-1)$ covariance matrix $\boldsymbol{\Phi}_t$ of the error vector.

We assume the Baba, Engle, Kraft, and Kroner (BEKK) GARCH(p, q) representation for the $(n-1) \times (n-1)$ covariance matrix of the error vector with generality parameter K [see Engle and Kroner (1995)]

$$\boldsymbol{\Phi}_t = \boldsymbol{C}'\boldsymbol{C} + \sum_{i=1}^{K} \sum_{i=1}^{p} \boldsymbol{B}'_{ik} \boldsymbol{\Phi}_{t-i} \boldsymbol{B}_{ik} + \sum_{k=1}^{K} \sum_{i=1}^{q} \boldsymbol{A}'_{ik} \boldsymbol{u}_{t-i} \boldsymbol{u}'_{t-i} \boldsymbol{A}_{ik}. \tag{9}$$

The BEKK model has the attractive property of having the conditional covariance matrix, $\boldsymbol{H}_t$, positive definite by construction. This model has $[n(n+3) - 2]/2$ free parameters in the conditional mean equations (6) and $(n-1)n/2 + K^2pqn^2$ free parameters in the conditional variance equations (9), for a total of $(K^2pq + 1)n^2 + n - 1$ free parameters. The following theorem claims the invariance of the maximum likelihood (ML) estimator with respect to the deleted equation for this model.

Theorem 2 *Let the covariance matrices* $(\boldsymbol{\Omega}_t)_{t=0}^{T}$ *have rank* $n-1$ *and the initial covariance matrix* $\boldsymbol{\Omega}_0 = \boldsymbol{\Lambda}$. *Then any two subsystems of (7) consisting of* $n-1$ *equations with the corresponding conditional variance equation (9) are observationally equivalent. Also, the ML estimates of the parameters of one subsystem can be recovered from the ML estimates of the parameters of another subsystem.*

The proof is given in Appendix B. Theorem 2 states that the result of the maximum likelihood estimation of the system does not depend on the choice of the $n-1$ equations to be estimated from the n equations of the demand system in the following sense. The estimators of the conditional mean equations are the same for any set of $n-1$ equations; the estimators of the conditional variance equations for any set of $n-1$ equations can be obtained from the estimators of any other set of $n-1$ equations, using the linear transformation (25)-(27) in Appendix B.

5 Empirical Application

Consider the model defined in the previous section with $n = 3$. Since (6) is a singular system we delete one equation (say the third equation) and consider the following two conditional mean equations

$$s_1 = \frac{\alpha_1 + \gamma_{11}\log v_1 + \gamma_{12}\log v_2 + \gamma_{13}\log v_3}{\sum_{k=1}^{n}\alpha_k + \sum_{k=1}^{n}\sum_{j=1}^{n}\gamma_{jk}\log v_k} + e_1; \tag{10}$$

$$s_2 = \frac{\alpha_2 + \gamma_{12}\log v_1 + \gamma_{22}\log v_2 + \gamma_{23}\log v_3}{\sum_{k=1}^{n}\alpha_k + \sum_{k=1}^{n}\sum_{j=1}^{n}\gamma_{jk}\log v_k} + e_2. \tag{11}$$

We assume a BEKK GARCH(1,1) with $K = 1$ representation for the covariance matrix of e_1 and e_2 in (10) and (11). In particular, the 2×2 covariance matrix of the errors can be written as

$$\boldsymbol{H}_t = \boldsymbol{C}'\boldsymbol{C} + \boldsymbol{B}'\boldsymbol{H}_{t-1}\boldsymbol{B} + \boldsymbol{A}'\boldsymbol{e}_{t-1}\boldsymbol{e}'_{t-1}\boldsymbol{A}. \tag{12}$$

Thus, the BTL demand system with a BEKK specification for the covariance matrix $\boldsymbol{H}_t$, consists of the conditional mean equations (10) and (11) and the following conditional variance and covariance equations

$$\begin{aligned} h_{11,t} &= c_{11}^2 + b_{11}^2 h_{11,t-1} + 2b_{11}b_{21}h_{12,t-1} + b_{21}^2 h_{22,t-1} \\ &\quad + a_{11}^2 e_{1,t-1}^2 + 2a_{11}a_{21}e_{1,t-1}e_{2,t-1} + a_{21}^2 e_{2,t-1}^2; \end{aligned} \tag{13}$$

$$\begin{aligned} h_{12,t} &= c_{11}c_{12} + b_{11}b_{12}h_{11,t-1} + (b_{11}b_{22} + b_{12}b_{21})\, h_{12,t-1} + b_{21}b_{22}h_{22,t-1} \\ &\quad + a_{11}a_{12}e_{1,t-1}^2 + (a_{11}a_{22} + a_{12}a_{21})\, e_{1,t-1}e_{2,t-1} + a_{21}a_{22}e_{2,t-1}^2; \end{aligned} \tag{14}$$

$$\begin{aligned} h_{22,t} &= c_{12}^2 + c_{22}^2 + b_{12}^2 h_{11,t-1} + 2b_{12}b_{22}h_{12,t-1} + b_{22}^2 h_{22,t-1} \\ &\quad + a_{12}^2 e_{1,t-1}^2 + 2a_{12}a_{22}e_{1,t-1}e_{2,t-1} + a_{22}^2 e_{2,t-1}^2. \end{aligned} \tag{15}$$

This model has a total of 19 free parameters to be estimated.

Applied demand analysis uses two types of data, time series data and cross sectional data. Time series data offer substantial variation in relative prices and less variation in income whereas cross sectional data offer limited variation in relative prices and substantial variation in income levels. In this application, we use the monthly time series data on monetary asset quantities and their user costs recently produced by Barnett *et al.* (2013) and maintained within the Center of Financial Stability (CFS) program Advances in Monetary and Financial Measurement (AMFM). The sample period is from 1967:2 to 2011:12 (a total of 539 observations). For a detailed discussion of the data and the methodology for the calculation of user costs, see Barnett *et al.* (2013) and http://www.centerforfinancialstability.org.

In particular, we model the demand for three monetary assets: demand deposits (x_1), small time deposits at commercial banks (x_2), and large time deposits (x_3). As we require real per capita asset quantities for the empirical work, we divided each quantity series by the CPI (all items) and total population. The estimation is performed in Estima RATS. We first estimate equations (10) and (11) under the homoscedasticity assumption in (3) and report the results in Table 1. To verify the presence of ARCH effects in the residuals of (10) and (11), estimated under the homoscedasticity assumption (3), we plot the estimated squared residuals $\hat{e}_1^2$ and $\hat{e}_2^2$ in Figures 1 and 2, respectively. Moreover, Lagrange multiplier tests for ARCH in each of $\hat{e}_1$ and $\hat{e}_2$ indicate significant evidence of ARCH effects; the null hypotheses of no ARCH (of different orders) in each of $\hat{e}_1$ and $\hat{e}_2$ are rejected with p-values less than 0.000001.

Next, we estimate the model under the heteroscedasticity assumption in (4), assuming the BEKK specification (12) for the error covariance matrix, $\boldsymbol{H}_t$. That is, we estimate the conditional mean equations, (10) and (11), and the conditional variance equations, (13)-(15). The estimation results are reported in Table 2. We also estimated the model using the subsystems with the second and first equations deleted (see Table 3 and 4, respectively). Consistent with Theorem 2 the parameters of the mean equations ($\alpha_1, \alpha_2, \alpha_3, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{21}, \beta_{22}, \beta_{23}, \beta_{31}, \beta_{32}, \beta_{33}$) are exactly the same in all three estimations (see panels A in Tables 2-4). The parameters of the variance equations are related by equations (25)-(27) in Appendix B.

It is worth mentioning that the log likelihood function of the described model exhibits manifold local maxima. To ensure reliability of the result it is necessary to implement the optimization with different initial parameters (e.g. randomly assigned from the predetermined area) and/or different eliminated equations. Furthermore, in most cases it is helpful to perform derivative-free search (such as Simplex algorithm) as a preliminary step for the derivative-based optimization.

6 Conclusion

Uncertainty is a very important concept in economics and finance, if not the most important. Motivated by the fact that the current demand systems literature ignores the role of uncertainty, in this paper we introduce recent advances in financial econometrics to model the covariance matrix of the errors of flexible demand systems, thereby improving the flexibility of these systems to capture certain important features of the data. We prove an important practical result of invariance of the ML estimator with respect to the deleted equation for the BTL demand system with conditional variance in the BEKK form. We also provide an empirical application based on the use of this model.

Although we study the BEKK specification of the error covariance matrix of the basic translog, our approach could be applied to any other known demand system (including those mentioned in the Introduction). Moreover, other variance specifications could be used such as, for example, the VECH model of Bollerslev *et al.* (1988), the constant conditional correlation (CCC) model of Bollerslev (1990), as well as the dynamic conditional correlation (DCC) models of Engle (2002) and Tse and Tsui (2002). Extension of the methodology to demand systems that focus on cross-sectional data or pooled cross-sectional data is an area for potential productive future research.

Appendix A
Proof of Theorem 1

Since the rank of $\boldsymbol{\Omega}_t$ is $n-1$, the covariance matrix for any subsystem of $n-1$ equations is not singular. Moreover, the error term of the deleted equation can be recovered from the error terms of the other $n-1$ equations using a linear injection (one-to-one mapping). Therefore, the vector of $n-1$ error terms, with the nth error term being deleted, denoted by $\boldsymbol{u}_t$, can be transformed to any other vector of $n-1$ error terms by eliminating the ith error term ($i = 1, ..., n-1$), denoted by $\boldsymbol{e}_t$, as follows

$$\boldsymbol{e}_t = \boldsymbol{T}\boldsymbol{u}_t \tag{16}$$

where $\boldsymbol{T}$ is a unity $(n-1) \times (n-1)$ transformation matrix with the ith row replaced by a vector of -1. The transformation matrix $\boldsymbol{T}$ has the property that $\boldsymbol{T}\boldsymbol{T} = \boldsymbol{I}$ and, therefore, $\boldsymbol{T} = \boldsymbol{T}^{-1}$. Hence, the time-varying covariance matrices $\boldsymbol{\Phi}_t$ and $\boldsymbol{H}_t$ of the error vectors $\boldsymbol{u}_t$ and $\boldsymbol{e}_t$, respectively, are linked by the following

$$\boldsymbol{\Phi}_t = \boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1} = \boldsymbol{T}\boldsymbol{H}_t\boldsymbol{T}'. \tag{17}$$

Moreover, the Jacobian of the transformation is

$$\mathbb{J}\left(\boldsymbol{T}\right) = -1. \tag{18}$$

We assume that the error term (4) is elliptically distributed. Therefore, its marginal distributions are also elliptical with characteristic functions of the form

$$\phi\left(\boldsymbol{t}\right) = E[e^{i\boldsymbol{t}'\boldsymbol{e}}] = \Psi\left(\boldsymbol{t}'\boldsymbol{H}_t\boldsymbol{t}\right)$$

for some function Ψ which does not depend on n, and its density function can be expressed as

$$f\left(\boldsymbol{e}_t\right) = \left|\boldsymbol{H}_t\right|^{-1/2} g\left(\boldsymbol{e}_t'\boldsymbol{H}_t\boldsymbol{e}_t; n-1\right) \tag{19}$$

where $g(\cdot; n-1)$ is a univariate function with parameter $(n-1)$.

The log likelihood function associated with equation (19) is given by

$$L\left(\boldsymbol{\vartheta}\right) = \sum_{t=1}^{T} l_t\left(\boldsymbol{\vartheta}\right)$$

where $\boldsymbol{\vartheta}$ is the set of parameters $\boldsymbol{\theta}$ of the mean equation (2) and covariance matrices $\left(\boldsymbol{H}_t\right)_{t=1}^{T}$ and

$$l_t\left(\boldsymbol{\vartheta}\right) = -\frac{1}{2}\ln\left|\boldsymbol{H}_t\right| + \ln g\left(\boldsymbol{u}_t'\boldsymbol{H}_t\boldsymbol{u}_t; n-1\right)$$

with error vectors $(\boldsymbol{u}_t)_{t=1}^{T}$ determined by the subsystem of (2) without the nth equation.

Using (16), (17), (18) and the fact that $\left|\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right| = |\boldsymbol{H}_t|$, we can write (19) as

$$\begin{aligned} f\left(\boldsymbol{e}_t\right) &= \left|\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right|^{-1/2} g\left(\boldsymbol{e}_t'\left(\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right)^{-1}\boldsymbol{e}_t; n-1\right) \\ &= |\boldsymbol{\Phi}_t|^{-1/2} g\left(\boldsymbol{u}_t'\boldsymbol{\Phi}_t^{-1}\boldsymbol{u}_t; n-1\right) \\ &= f\left(\boldsymbol{u}_t\right). \end{aligned}$$

Therefore, the log likelihood functions based on observations of the error vectors $\boldsymbol{u}_t$ and $\boldsymbol{e}_t$ are the same. This proves the theorem. Q.E.D.

Appendix B
Proof of Theorem 2

Without loss of generality, we provide the proof for the case of a GARCH(1,1) BEKK with $K = 1$ representation for the conditional variance equation. Consider the subsystem of (7) with the last equation deleted, i.e. the error vector is $\boldsymbol{u}_t = (\epsilon_{1t}, \epsilon_{2t}, ..., \epsilon_{n-1,t})$. The Gaussian log likelihood function based on the sample $(\boldsymbol{v}_t, \boldsymbol{s}_t)_{t=1}^{T}$ can be written as

$$L\left(\boldsymbol{\Theta} | \left(\boldsymbol{u}_t\right)_{t=1}^{T}\right) = -\frac{T(N-1)}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log|\boldsymbol{\Phi}_t| + \boldsymbol{u}_t'\boldsymbol{\Phi}_t^{-1}\boldsymbol{u}_t\right) \tag{20}$$

where

$$\boldsymbol{\Phi}_t = \boldsymbol{C}'\boldsymbol{C} + \boldsymbol{B}'\boldsymbol{\Phi}_{t-1}\boldsymbol{B} + \boldsymbol{A}'\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}'\boldsymbol{A} \tag{21}$$

and $\boldsymbol{\Theta} = (\boldsymbol{\alpha}, vech(\boldsymbol{\Gamma}), vech(\boldsymbol{C}), vec(\boldsymbol{B}), vec(\boldsymbol{A}))$ is the vector of all parameters to be estimated and $(\boldsymbol{u}_t)_{t=1}^{T}$ is determined by the subsystem (7) without the nth equation.

As before, consider the subsystem of (7) without the ith equation and corresponding error vector $\boldsymbol{e}_t = (\epsilon_{1t}, ..., \epsilon_{i-1,t}, \epsilon_{i+1,t}, ..., \epsilon_{nt})$. It has been shown that the error vectors $\boldsymbol{u}_t$ and $\boldsymbol{e}_t$ are linked by the non-singular linear transformation (16). The Gaussian log likelihood function for this subsystem is

$$L\left(\widetilde{\boldsymbol{\Theta}} \middle| \left(\boldsymbol{e}_t\right)_{t=1}^{T}\right) = -\frac{T(N-1)}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log|\boldsymbol{H}_t| + \boldsymbol{e}_t'\left(\boldsymbol{H}_t\right)^{-1}\boldsymbol{e}_t\right) \tag{22}$$

where

$$\boldsymbol{H}_t = \tilde{\boldsymbol{C}}'\tilde{\boldsymbol{C}} + \left(\tilde{\boldsymbol{B}}\right)'\boldsymbol{H}_{t-1}\tilde{\boldsymbol{B}} + \left(\tilde{\boldsymbol{A}}\right)'\boldsymbol{e}_{t-i}\left(\boldsymbol{e}_{t-i}\right)'\tilde{\boldsymbol{A}} \tag{23}$$

and $\widetilde{\boldsymbol{\Theta}} = (\widetilde{\boldsymbol{\alpha}}, vech(\widetilde{\boldsymbol{\Theta}}), vech(\tilde{\boldsymbol{C}}), vec(\tilde{\boldsymbol{B}}), vec(\tilde{\boldsymbol{A}}))$ is the vector of all the parameters.

Using (16), (18), and the fact that $\left|\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right| = |\boldsymbol{H}_t|$ we can write the log likelihood function (**??**)-(**??**) as

$$L\left(\widetilde{\boldsymbol{\Theta}} \middle| \left(\boldsymbol{e}_t\right)_{t=1}^{T}\right) = -\frac{T(N-1)}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log\left|\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right| + \boldsymbol{u}_t'\left(\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1}\right)^{-1}\boldsymbol{u}_t\right)$$

with

$$\begin{aligned}\boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1} = {} & \boldsymbol{T}^{-1}\tilde{\boldsymbol{C}}'\tilde{\boldsymbol{C}}\left(\boldsymbol{T}'\right)^{-1} + \boldsymbol{T}^{-1}\tilde{\boldsymbol{B}}'\boldsymbol{T}\left(\boldsymbol{T}^{-1}\boldsymbol{H}_{t-1}\left(\boldsymbol{T}'\right)^{-1}\right)\boldsymbol{T}'\tilde{\boldsymbol{B}}\left(\boldsymbol{T}'\right)^{-1} \\ & + \boldsymbol{T}^{-1}\tilde{\boldsymbol{A}}'\boldsymbol{T}\boldsymbol{e}_{t-i}\boldsymbol{e}_{t-i}'\boldsymbol{T}'\tilde{\boldsymbol{A}}\left(\boldsymbol{T}'\right)^{-1}.\end{aligned}$$

Now we observe that systems (20)-(21) and (22)-(23) are identical if the following conditions hold

$$\boldsymbol{\Phi}_t = \boldsymbol{T}^{-1}\boldsymbol{H}_t\left(\boldsymbol{T}'\right)^{-1} \tag{24}$$

$$\boldsymbol{C} = Cholesky\left(\boldsymbol{T}^{-1}\tilde{\boldsymbol{C}}'\tilde{\boldsymbol{C}}\left(\boldsymbol{T}'\right)^{-1}\right) \tag{25}$$

$$\boldsymbol{B} = \boldsymbol{T}'\tilde{\boldsymbol{B}}\left(\boldsymbol{T}'\right)^{-1} \tag{26}$$

$$\boldsymbol{A} = \boldsymbol{T}'\tilde{\boldsymbol{A}}\left(\boldsymbol{T}'\right)^{-1} \tag{27}$$

where $Cholesky(\cdot)$ stands for the Cholesky decomposition.

Clearly, the ML estimators of $\boldsymbol{\Theta}$ and $\widetilde{\boldsymbol{\Theta}}$ are related as follows. The parameters of the conditional mean equations are the same: $\boldsymbol{\alpha} = \widetilde{\boldsymbol{\alpha}}$ and $\boldsymbol{\Gamma} = \widetilde{\boldsymbol{\Gamma}}$ whereas the parameters of the conditional variance equations are related through (25)-(27). Note that condition (24) is satisfied for the initial covariance matrices which are predetermined by $\boldsymbol{\Lambda}$. In addition, since both $\boldsymbol{C}$ and $\tilde{\boldsymbol{C}}$ are triangular matrices in the BEKK representation, in order to preserve the triangular form, equation (25) involves transformation using Cholesky decomposition.

Hence, the likelihood functions of different subsystems of (7) consisting of $n-1$ equations with corresponding conditional variance equation (9) are the same up to the parameters transformation which relates the maximum likelihood estimators for these subsystems. Q.E.D.

References

[1] Banks, J., R. Blundell, and A. Lewbel. "Quadratic Engel Curves and Consumer Demand." *Review of Economics and Statistics* 79 (1997), 527-539.

[2] Barnett, W.A. "Definitions of Second-Order Approximation and of Flexible Functional Form." *Economics Letters* 12 (1983), 31-35.

[3] Barnett, W.A. and A. Jonas. "The Müntz-Szatz Demand System: An Application of a Globally Well Behaved Series Expansion." *Economics Letters* 11 (1983), 337-342.

[4] Barnett, W.A. and A. Serletis. "Consumer Preferences and Demand Systems." *Journal of Econometrics* 147 (2008), 210-224.

[5] Barnett, W.A. and P. Yue. "Seminonparametric Estimation of the Asymptotically Ideal Model: The AIM demand system." In: G.F. Rhodes and T.B. Fomby, (Eds.), Nonparametric and Robust Inference: Advances in Econometrics, Vol. 7. JAI Press, Greenwich CT (1988).

[6] Barnett, W.A., J. Liu, R.S. Mattson, and J. van den Noort. "The New CFS Divisia Monetary Aggregates: Design, Construction, and Data Sources." *Open Economies Review* 24 (2013), 101-124.

[7] Barten, A.P. "Maximum Likelihood Estimation of a Complete System of Demand Equations." *European Economic Review* 1 (1969), 7-73.

[8] Bauwens, L., S. Laurent, and J.V.K. Rombouts. "Multivariate GARCH Models: A Survey." *Journal of Applied Econometrics* 21 (2006), 79-109.

[9] Blundell, R., D. Kristensen, and R. Matzkin. "Stochastic Demand and Revealed Preference." *Journal of Econometrics* (2014, forthcoming).

[10] Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31 (1986), 307-27.

[11] Bollerslev, T. "Modeling the Coherence in Short-Term Nominal Exchange Rates: A Multivariate Generalized ARCH Approach." *Review of Economics and Statistics* 72 (1990), 498-505.

[12] Bollerslev, T., R.F. Engle, and J. Wooldridge. "A Capital Asset Pricing Model with Time-Varying Covariances." *Journal of Political Economy* 96 (1988), 116-131.

[13] Christensen, L.R., D.W. Jorgenson, and L.J. Lau. "Transcendental Logarithmic Utility Functions." *American Economic Review* 65 (1975), 367-383.

[14] Deaton, A. and J.N. Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70 (1980), 312-326.

[15] Diewert, W.E. "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function." *Journal of Political Economy* 79 (1971), 481-507.

[16] Diewert, W.E. "Applications of Duality Theory." In *Frontiers of Quantitative Economics* (Vol. 2), eds. M.D. Intriligator and D.A. Kendrick, Amsterdam: North-Holland (1974), pp. 106-171.

[17] Diewert, W.E. and T.J. Wales. "Normalized Quadratic Systems of Consumer Demand Functions." *Journal of Business and Economic Statistics 6* (1988), 303-312.

[18] Engle, R.F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation." *Econometrica* 50 (1982), 987–1008.

[19] Engle, R.F. "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models." *Journal of Business and Economic Statistics* 20 (2002), 339-350.

[20] Engle, R.F. and K.F. Kroner. "Multivariate Simultaneous Generalized ARCH." *Econometric Theory* 11 (1995), 122-150.

[21] Gallant, A.R. "On the Bias in Flexible Functcional Forms and an Essentially Unbiased Form: The Fourier Flexible Form." *Journal of Econometrics* 15 (1981), 211-245.

[22] Granger, C.W.J. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14 (1980), 227-238.

[23] Hoderlein, S. "How Many Consumers are Rational?" *Journal of Econometrics* 164 (2011), 294-309.

[24] Hoderlein, S. and J. Stoye. "Revealed Preferences in a Heterogeneous Population." *Review of Economics and Statistics* (2014, forthcoming).

[25] Lewbel, A. "Consumer Demand Systems and Household Equivalent Scales." In Pesaran M.H. and P. Schmidt (Eds.), *Handbook of Applied Econometrics*, Volume II Microeconomics. Oxford: Blackwell Publishers Ltd. (1997).

[26] Lewbel, A. "Demand Systems with and without Errors." *American Economic Review* 91 (2001), 611-618.

[27] Lewbel, A. and S. Ng. "Demand Systems with Nonstationary Prices." *Review of Economics and Statistics* 87 (2005), 479-494.

[28] Lewbel, A. and K. Pendakur. "Tricks with Hicks: The EASI Demand System." *American Economic Review* 99 (2009), 827-863.

[29] McLaren, K.R. "A Variant on the Arguments for the Invariance of Estimators in a Singular System of Equations." *Econometric Reviews* 9 (1990), 91-102.

[30] Nelson, D.B. "Conditional Heteroskedasticity in Asset Returns." *Econometrica* 59 (1991), 347-370.

[31] Silvennoinen, A. and T. Teräsvirta. "Multivariate GARCH Models." In T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch (Eds.), *Handbook of Financial Time Series*. New York: Springer (2011).

[32] Tse, Y.K. and A.K.C. Tsui. "A Multivariate GARCH Model with Time-Varying Correlations." *Journal of Business and Economic Statistics* 20 (2002), 351-362.

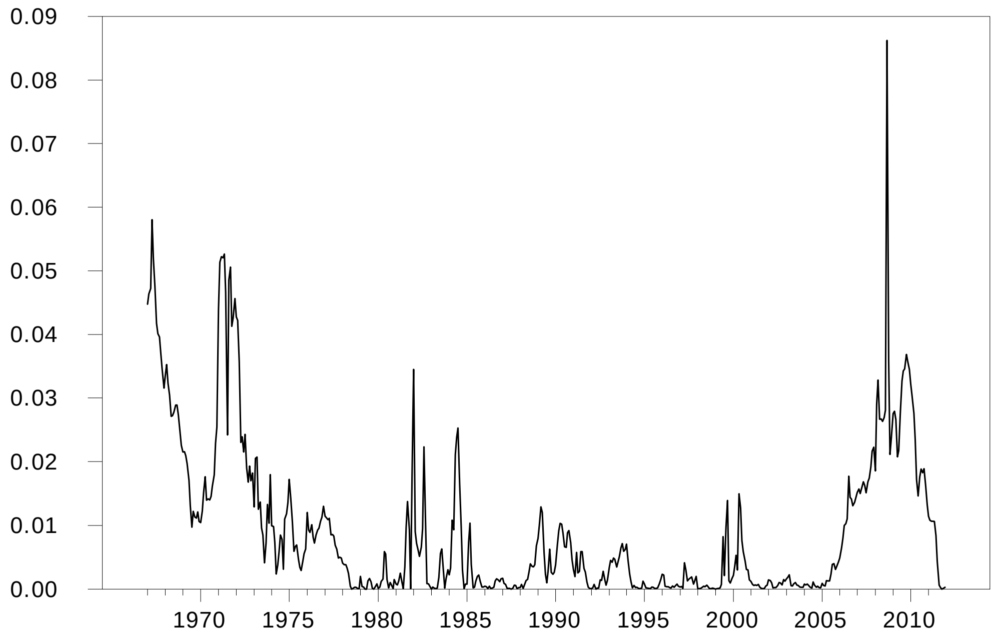

Figure 1: Squared residuals of equation (22), $\hat{e}_1^2$

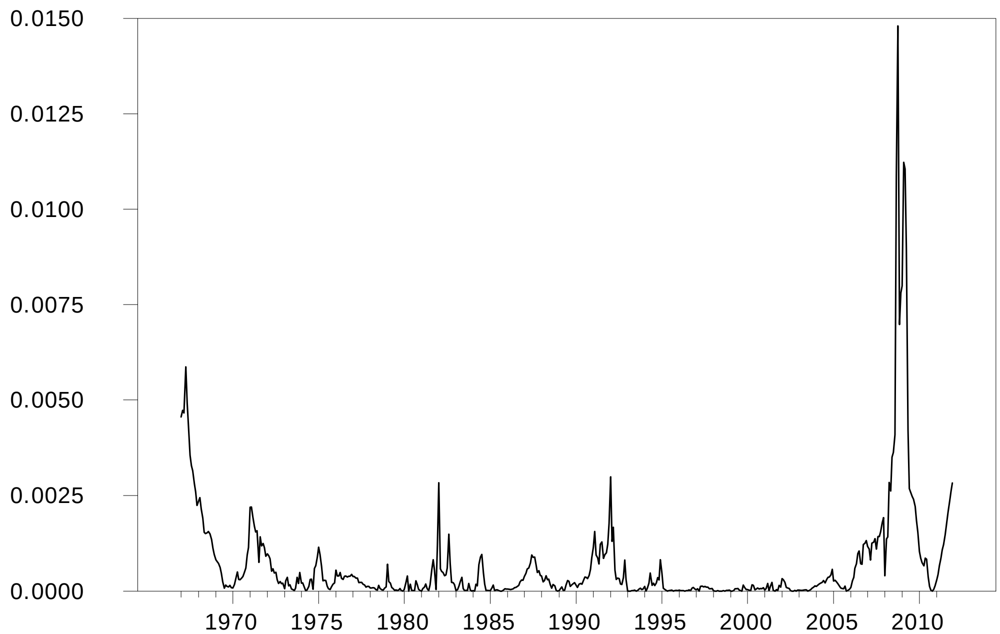

Figure 2: Squared residuals of equation (23), $\hat{e}_2^2$

Table 1

BTL Parameter Estimates

Parameter	Estimate	Standard error
a_1	0.6578	0.0030
a_2	0.1853	0.0015
a_3	0.1569	0.0024
γ_{11}	−0.2436	0.0634
γ_{12}	−0.0635	0.0148
γ_{13}	−0.2796	0.0196
γ_{22}	0.1700	0.0053
γ_{23}	−0.0233	0.0070
γ_{33}	0.1919	0.0115
Log L	1947.1214	

Notes: Sample period, monthly data 1967:2-2011:12.

Table 2

Estimates Of BTL With BEKK Errors

Parameter	Estimate	Standard error
Mean equations:		
a_1	0.6199	0.0005
a_2	0.1843	0.0006
a_3	0.1958	0.0008
γ_{11}	0.0229	0.0141
γ_{12}	-0.0402	0.0028
γ_{13}	-0.3469	0.0042
γ_{22}	0.1618	0.0019
γ_{23}	-0.0922	0.0019
γ_{33}	0.0437	0.0029
Variance equations (with the 3rd mean equation deleted):		
c_{11}	0.0038	0.0003
c_{12}	-0.0010	0.0002
c_{13}	0.0017	0.0001
a_{11}	0.5225	0.0244
a_{12}	-0.0258	0.0046
a_{21}	0.1455	0.0297
a_{22}	0.2867	0.0493
b_{11}	0.9302	0.0316
b_{12}	0.0113	0.0021
b_{21}	-0.0474	0.0220
b_{22}	1.0249	0.0237
Log L	3080.6445	

Notes: Sample period, monthly data 1967:2-2011:12.

Table 3

Estimates Of BTL With BEKK Errors

Parameter	Estimate	Standard error
Mean equations:		
a_1	0.6199	0.0003
a_2	0.1843	0.0001
a_3	0.1958	0.0001
γ_{11}	0.0229	0.0000
γ_{12}	−0.0402	0.0007
γ_{13}	−0.3469	0.0012
γ_{22}	0.1618	0.0005
γ_{23}	−0.0922	0.0005
γ_{33}	0.0437	0.0006
Variance equations (with the 2rd mean equation deleted):		
c_{11}	0.0038	0.0001
c_{12}	−0.0027	0.0001
c_{13}	0.0017	0.0000
a_{11}	0.3769	0.0018
a_{12}	−0.0645	0.0006
a_{21}	−0.1455	0.0006
a_{22}	0.4320	0.0014
b_{11}	0.9777	0.0064
b_{12}	0.0360	0.0006
b_{21}	0.0473	0.0005
b_{22}	0.9777	0.0062
Log L	3080.6444	

Notes: Sample period, monthly data 1967:2-2011:12.

Table 4

Estimates Of BTL With BEKK Errors

Parameter	Estimate	Standard error
Mean equations:		
a_1	0.6199	0.0003
a_2	0.1843	0.0001
a_3	0.1958	0.0000
γ_{11}	0.0229	0.0040
γ_{12}	−0.0402	0.0008
γ_{13}	−0.3469	0.0013
γ_{22}	0.1618	0.0006
γ_{23}	−0.0922	0.0006
γ_{33}	0.0437	0.0007
Variance equations (with the 1rd mean equation deleted):		
c_{11}	0.0032	0.0003
c_{12}	−0.0001	0.0002
c_{13}	0.0020	0.0001
a_{11}	0.4967	0.0052
a_{12}	0.0258	0.0015
a_{21}	0.0645	0.0076
a_{22}	0.3126	0.0071
b_{11}	0.9415	0.0104
b_{12}	−0.0113	0.0018
b_{21}	−0.0360	0.0153
b_{22}	1.0136	0.0097
Log L	3080.6445	

Notes: Sample period, monthly data 1967:2-2011:12.