CONDITIONAL COVARIANCE MODELING AND APPLICATIONS IN MUTUAL FUND PERFORMANCE EVALUATION

By

KAINAN WANG

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
College of Business

MAY 2012
To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of KAINAN WANG find it satisfactory and recommend that it be accepted.

Harry J. Turtle, Ph.D., Chair

John R. Nofsinger, Ph.D.

Sung K. Ahn, Ph.D.
ACKNOWLEDGEMENT

I would like to thank my dissertation chair, Dr. Harry Turtle, for his intellectual and moral support for my scientific development. He has encouraged and guided me during the course of my studies.

I would also like to thank my parents, Jinjun Wang and Yunyun Xie, and my wife, Xuan Fei, for their encouragement and support.

My appreciation also goes to all the current and former faculties, staffs, and Ph.D. students at Washington State University for their stimulating discussion and friendship: Dr. Sung K. Ahn, Dr. Gene Lai, Dr. Sheen Liu, Dr. Michael McNamara, Dr. John Nofsinger, Dr. Donna Paul, Dr. Richard Sias, Dr. Harry Turtle, Dr. Nathan Walcott, Dr. David Whidbee, Dr. Lily Xu, Ms. Sandra Boyce, Ms. Nicole Cherry, Dr. Nicole Choi, Chune Y. Chung, Luke DeVault, Dr. Hongyan Fang, Kevin Graham, Dr. Jian Huang, Wei Huang, Chang Liu, Erin Lu, Wen-Ling Lu, JD Morschick, Dr. Bela Reca, Dr. Abhishek Varma, Dr. Weicheng Wang, Dr. Dandan Wu, Gulnara Zaynudinova, and Dr. Chengping Zhang.

Special thanks are given to my dissertation committee Dr. Harry Turtle, Dr. John Nofsinger, and Dr. Sung K. Ahn for the valuable suggestions for this dissertation.
This dissertation contains two essays. The first essay proposes a new model for conditional covariances based on predetermined information instruments. The model is based on a restricted Cholesky-like decomposition and ensures positive definiteness and invariance to variate order by construction. Comparing to existing time-series models, the new model provides a parsimonious and accurate description of second moments.

The second essay develops a new conditional alpha model to investigate the risk-adjusted performance of mutual funds of different size as measured by assets under management. The model incorporates the conditional covariance methodology in the first essay and allows both time-varying risks and expected returns to rationally evolve with economic fundamentals. Using fund flow as an information instrument, we find that small mutual funds significantly outperform large funds and that these performance gains are most apparent following fund inflows. In contrast, larger funds display poorer performance after inflows. Our findings support the Berk and Green (2004) notion that small funds are more nimble in deploying optimal strategies than large funds. Results are consistent in both unconditional and conditional performance models.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>CHAPTER ONE: Modeling conditional covariances with economic information instruments</td>
<td>6</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>6</td>
</tr>
<tr>
<td>2. Model Development</td>
<td>10</td>
</tr>
<tr>
<td>3. Empirical Analysis</td>
<td>19</td>
</tr>
<tr>
<td>4. Concluding Remarks</td>
<td>29</td>
</tr>
<tr>
<td>References</td>
<td>31</td>
</tr>
<tr>
<td>Appendix</td>
<td>33</td>
</tr>
<tr>
<td>Tables</td>
<td>36</td>
</tr>
<tr>
<td>CHAPTER TWO: Does mutual fund size erode performance? Conditional evidence using fund flows</td>
<td>46</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>46</td>
</tr>
<tr>
<td>2. Model</td>
<td>49</td>
</tr>
<tr>
<td>3. Empirical Analysis</td>
<td>52</td>
</tr>
<tr>
<td>4. Concluding Remarks</td>
<td>60</td>
</tr>
<tr>
<td>References</td>
<td>61</td>
</tr>
<tr>
<td>Appendix 1 – Conditional alpha with conditional covariances</td>
<td>64</td>
</tr>
<tr>
<td>Appendix 2 – Inferences with conditional bootstrap</td>
<td>65</td>
</tr>
<tr>
<td>Tables and Figures</td>
<td>66</td>
</tr>
</tbody>
</table>
Modeling conditional covariances

with economic information instruments

1. Introduction

Modeling the second moments of financial asset returns has important implications in many research designs examining economic behavior including risk management, derivative pricing, hedging, and portfolio optimization. Beginning with the seminal work of Engle (1982) and the subsequent extension by Bollerslev (1986), a vast literature on modeling the time varying covariance structure of asset returns has developed. Many economic models examine multivariate relationships between first and second moments – making covariance estimation particularly important in economic research designs.

Recent developments in the estimation of multivariate second moment specifications include, among others, the constant conditional correlation (CCC) model of Bollerslev (1990), the BEKK model of Engle and Kroner (1995), the Factor GARCH model of Engle, Ng and Rothschild (1990), and the Dynamic Conditional Correlation (DCC) model of Engle (2002), and Rangel and Engle (2011). These models have had varying degrees of success in capturing the volatility clustering, leptokurtosis and asymmetry commonly observed in many economic and financial time series. One of the challenges in modeling second moments is to ensure the positive definiteness of the resultant covariance matrix without spuriously imposing a temporal pattern in conditional covariances. To cope with this issue, numerous models have been proposed. For instance, to guarantee positive definiteness, the BEKK model specifies the covariance matrix as the sum of multiple positive definite matrices, where the DCC model imposes restrictions on correlation parameters in a multistep procedure that facilitates large scale estimation. One major drawback with these approaches is that the model structure is ad hoc and may therefore inadvertently impact the underlying economic dynamics describing the evolution of second moments. We propose a simple information instrument model that both ensures the positive definiteness of the resultant covariance matrix and incorporates related economic
Although our approach is very different, the econometric intuition of our specification is similar in spirit to Rangel and Engle (2011), who focus on capturing the importance of economic variables in their multivariate covariance specification.

By design, our conditional covariance model is multivariate, intuitive, and parsimonious. As we demonstrate in our simulations, the approach proves well behaved by standard measures of statistical performance and in comparison with the most familiar approaches. The essence of our model can be captured with two observations. First, any positive definite matrix can be rewritten as a lower triangular matrix multiplied by its transpose. For simplicity, we model each term in this specification as a linear function of economic information instruments. Linearity can be motivated with a Taylor approximation using well-chosen instruments, even for relatively complex nonlinear underlying specifications. Second, we require that a simple reshuffling of the economic variates in the system should not produce a different functional form for the resultant covariance terms. For example, in our system of five book-to-market (BM) portfolios, suppose we hypothesize that covariances are a function of lagged levels of liquidity. Our specification then requires that if the liquidity for the smallest BM portfolio depends on its own liquidity, other portfolios may only include their own liquidity levels. These two modeling choices lead to a parsimonious model specification to capture covariance dynamics.

The essential elements of our approach can be described in a simple bivariate example. Suppose that we are interested in modeling the covariance between the stock returns for two portfolios, a value portfolio, \( r_{1t} \), and a growth portfolio, \( r_{2t} \). For simplicity, let \( \mathbf{r}_t = \begin{bmatrix} r_{1t} \\ r_{2t} \end{bmatrix} \) be a vector with zero mean returns, and conditional covariance matrix \( \mathbf{M}_t = \begin{bmatrix} m_{11t} & m_{12t} \\ m_{21t} & m_{22t} \end{bmatrix} \). Our primary interest is to provide an econometric description of covariances related to underlying variate specific economic information instruments that may include lagged endogenous or known exogenous variables. As a special case of our more general framework suppose all variates are solely dependent on a single own economic information instrument. A researcher may

\[ \text{1 There are a number of related works that include economic variables as independent variables within second moment specifications including, for example, Hodrick (1989), Andersen, Bollerslev, Diebold, and Ebens (2001), Lamoureux and Lastrapes (1990), Engle, Ng and Rothschild (1990), Gray (1996), Hagwara and Herce (1999), and Engle and Patton (2001). Although these works are similar in spirit to our approach, we believe our approach is unique.} \]
hypothesize that returns to value and growth firms are related to their underlying sensitivities to previous research and development expenditures. As an example, a researcher might hypothesize that returns are nonlinearly related to opacity, where opacity increases with previous research development expenditures. Other examples of important own economic information instruments might include accounting information, previous price to earnings ratios, liquidity measures, dividend yields, or any other time varying characteristics that vary in cross-section. Let the known economic information instruments be given by $Z_{t-1} = [Z_{i,t-1}]$ for $i = 1$ and $2$. Recognizing that every positive definite covariance matrix yields a lower triangular decomposition, allows us to write $M_t = L_t L_t'$ where $L_t$ is a lower triangular matrix, and where each element in $L_t = [l_{ijt}]$ is written as a linear function of $Z_{t-1}$. Consider the following functional form for $l_{ijt}$,

$$l_{ijt} \equiv \gamma_{ij0} + \gamma_{ij1} Z_{i,t-1}$$

for $i \geq j, j = 1, \text{ and } 2$; and model parameters $\gamma_{ij0}$ and $\gamma_{ij1}$. Direct multiplication yields the resultant covariance matrix as,

$$M_t = \begin{bmatrix} m_{11t} & m_{12t} \\ m_{21t} & m_{22t} \end{bmatrix} = \begin{bmatrix} \gamma_{110} + \gamma_{111} Z_{1,t-1} & 0 \\ \gamma_{210} + \gamma_{211} Z_{2,t-1} & \gamma_{220} + \gamma_{221} Z_{2,t-1} \end{bmatrix} \begin{bmatrix} \gamma_{110} + \gamma_{111} Z_{1,t-1} & \gamma_{210} + \gamma_{211} Z_{2,t-1} \\ 0 & \gamma_{220} + \gamma_{221} Z_{2,t-1} \end{bmatrix}. \tag{2}$$

Evaluating term by term, we see that $m_{11t}$ is a linear function of $Z_{1,t-1}$ and $Z_{1,t-1}^2$; $m_{22t}$ is a linear function of $Z_{2,t-1}$ and $Z_{2,t-1}^2$; and $m_{12t}$ is a linear function of $Z_{1,t-1}, Z_{2,t-1}$, and $Z_{1,t-1}Z_{2,t-1}$.

This model has two important features. First, the functional form of each element in $M_t$ with respect to the information instruments is intuitive and easy to interpret. For example, the variance of each company’s return depends only on the firm specific information, where the covariance is driven by the economic information for both portfolios. Second, the model has the desirable properties of positive definiteness and invariance to variate order. Positive definiteness of $M_t$ is guaranteed if each diagonal element in $L_t$ is non-zero (see proof in the Appendix). The invariance to variate order property effectively ensures that the linear functional form for each lower triangular matrix element does not change when the order of the variates changes. To see that equation (1) satisfies the invariance to variate order property, notice that if the reordered system $r_t^* = \begin{bmatrix} r_{2t} \\ r_{1t} \end{bmatrix}$
were estimated, the functional form for the two variates would not be changed. We view this property as desirable and consider a model as valid only if it satisfies this property.

Many linear specifications for \( l_{ijt} \) will not satisfy \textit{invariance to variate order}. For example, consider the more exhaustive specification,

\[
l_{ijt} = \begin{cases} 
  y_{i0} + y_{ij1} Z_{it-1}, & \text{for } i = j \\
  y_{i0} + y_{ij1} Z_{it-1} + y_{ij2} Z_{jt-1}, & \text{for } i > j
\end{cases}
\]

for \( j = 1, \) and \( 2; \) and model parameters \( y_{i0}, y_{ij1}, \) and \( y_{ij2}. \) This specification admits terms in both \( Z_{it-1} \) and \( Z_{jt-1} \) for all \( l_{ijt} \) terms. Notice that direct multiplication yields the resultant covariance matrix as,

\[
M_t = \begin{bmatrix}
  m_{11t} & m_{12t} \\
  m_{21t} & m_{22t}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  y_{110} + y_{111} Z_{1t-1} & 0 \\
  y_{210} + y_{211} Z_{2t-1} + y_{212} Z_{1t-1} & y_{220} + y_{221} Z_{2t-1}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  y_{110} + y_{111} Z_{1t-1} & y_{210} + y_{211} Z_{2t-1} + y_{212} Z_{1t-1} & y_{220} + y_{221} Z_{2t-1}
\end{bmatrix}
\]

Evaluating term by term, we observe that \( m_{11t} \) is a linear function of \( Z_{1t-1} \) and \( Z_{1t-1}^2, \) where \( m_{22t} \) is a linear function of \( Z_{1t-1}, Z_{2t-1}, Z_{1t-1}^2, Z_{2t-1}^2, \) and \( Z_{2t-1}^2. \) If the order of two assets were changed, the functional form for the two variates would be changed due \textit{solely} to variate order. Therefore, specification (3) does not satisfy our proposed \textit{invariance to variate order} condition.

Our conditional covariance model is distinct from much of the extant literature in a number of important manners. First, our model is built with an intention to provide an economic description of second moments rather than a purely time series motivated specification. This approach has the potential to better nest within asset pricing contexts in describing nonlinear relations between economic variates. Nonetheless, our specification may also be readily adapted to a purely time series specification as demonstrated in our empirical analysis. Second, positive definiteness and uniqueness of the covariance matrix are assured, without resorting to difficult restricted estimation strategies. Finally, among the class of linear functions of \( L_t, \) equation (1) is the only nontrivial specification that satisfies the \textit{invariance to order} property.

In our empirical simulations, we compare our model to a familiar multivariate ARCH specification for the covariance matrix suggested by the BEKK model. By construction, our linear information instrument
model has the potential to capture nonlinear economic relationships between variates. Consistent with this conjecture, we find that our model outperforms the BEKK model for all terms in the resultant conditional covariance matrix by mean absolute errors (MAE), and root mean square errors (RMSE). When using heteroskedasticity-adjusted mean squared errors (HMSE), we find that for all but one covariance term, our model has greater estimation accuracy than the BEKK model. In addition, our linear information instrument design contains fewer parameters than the BEKK design. Our simulation experiments also confirm that the proposed model outperforms the BEKK specification for all multivariate metrics of performance.

The remainder of the paper is organized as follows. Section 2 presents the model, discusses the resultant positive definiteness of the covariance matrix, and examines the invariance to order property of the model. Section 3 presents the empirical analysis and simulation study. Concluding remarks are provided in Section 4.

2. Model Development

2.1 Model

Our goal is to provide an economically driven specification for time varying covariances that is parsimonious, yet rich in describing covariances as a function of the underlying economic state. Consider any \( N \times 1 \) zero mean random vector, \( \mathbf{r}_t \), with associated conditional covariance matrix, \( \mathbf{M}_t \equiv \mathbb{E}_{t-1}[ \mathbf{r}_t \mathbf{r}_t'] \), for \( t = 1, \ldots, T \). We wish to consider potential specifications for \( \mathbf{M}_t \) that are related to underlying economic information instruments that may be either lagged endogenous variables, or any other known observable economic information instruments from the information set. For clarity, we consider two general types of economic information instruments, \( \textit{own} \) economic information instruments, \( \mathbf{Z}_{i,t-1}, i = 1, \ldots, N \) that are specific to each series considered, as well as a \( K \times 1 \) vector of macroeconomic sources of information, \( \mathbf{Z}_{\text{agg}, t-1} \). In equilibrium, we might consider the relation between covariances of abnormal asset returns to both individual asset characteristics as well as observed market characteristics. For example, stock or portfolio covariances may be dependent on \( \textit{own} \) financial characteristics such as liquidity measures, profitability
measures, or valuation measures like price-to-book or cash-flow-to-book. In addition, overall market forces such as general credit market conditions might impact conditional covariances.

Before stating our main proposition, we introduce a helpful conceptual definition.

**Definition:** A model is said to be *(strongly) invariant to variate order* if the specific functional form for all variance and covariance specifications includes the same independent information instruments when the variate order is perturbed. We leave for future research the slightly weaker condition of *weak invariance to variate order* that only requires the same variables, but does not consider functional form.

A model is not *strongly invariant to order* if a simple shuffling of the variates produces a different functional form for the covariance specification for any two of the variates. In our empirical application, we require that each pair of our five portfolio return series has the same functional form for each potential (co)variance specification.

Equation (3) provides an example of a violation of both *weak* and *strong invariance*. To see that *weak invariance* is violated note that the specification for $m_{22t}$ includes both $Z_{1,t-1}$ and $Z_{2,t-1}$ as information instruments. In contrast, the specification for $m_{11t}$ only includes $Z_{1,t-1}$ as an information instrument. Of course, *weak invariance* implies *strong invariance*. Nonetheless, a further example may help to clarify the underlying principle. Notice that the specification for $m_{21t} = m_{12t}$ includes $Z_{1,t-1}^2$, but not $Z_{2,t-1}^2$, as an information instrument. If the system were to be reestimated after changing the order of the variates, the functional form for covariances would be altered solely due to an ad hoc decision about variate order.
Our general framework considers the time varying conditional covariance matrix, $\mathbf{M}_t = [m_{ij,t}]$, for $i, j = 1, 2, ..., N$, and for $t = 1, ..., T$. For each conditional covariance matrix, $\mathbf{M}_t$, we consider the following decomposition,

$$
\mathbf{M}_t = \mathbf{L}_t(\theta|\Omega_{t-1})\mathbf{L}_t(\theta|\Omega_{t-1})',
$$

(5)

where $\theta$ represents a vector of parameters, $\Omega_{t-1}$ represents a vector of information instruments, and $\mathbf{L}_t(\theta|\Omega_{t-1})$ is any lower triangular matrix with nonzero diagonals.\(^2\) Our interest is in a particular form for the nonzero elements within $\mathbf{L}_t(\theta|\Omega_{t-1}) = [l_{ij,t}(\theta|\Omega_{t-1})]$ from the general class of linear functions in aggregate information instruments, and in own information instruments defined as,

$$
l_{ij,t}(\theta|\Omega_{t-1}) = \begin{cases} 
\gamma_{i0} + \gamma_{ij1}'Z_{agg,t-1} + \gamma_{ij2}Z_{lt-1}, & \text{for } i = j \\
\gamma_{ij0} + \gamma_{ij1}'Z_{agg,t-1} + \gamma_{ij2}Z_{lt-1} + \gamma_{ij3}Z_{jt-1}, & \text{for } i > j 
\end{cases}
$$

(6)

where $i, j = 1, 2, ..., N; Z_{agg,t-1}$ is a $K \times 1$ vector of macroeconomic sources of information; $Z_{lt-1}$ and $Z_{jt-1}$ are known own information instruments for the $i^{th}$ and $j^{th}$ variate, respectively; and where $\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2},$ and $\gamma_{ij3}$ are known conformable parameters.\(^3\) For all $i < j = 1, 2, ..., N$ we set $l_{ij,t} = 0$.

To simplify our notation, we may rewrite the class of linear functions as,

$$
l_{ij,t}(\theta|\Omega_{t-1}) = \theta Z
$$

(7)

where $\theta = [\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2}]'$ and $Z = [1, Z_{agg,t-1}, Z_{lt-1}]$ for $i = j$, and $\theta = [\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2}, \gamma_{ij3}]'$ and $Z = [1, Z_{agg,t-1}, Z_{lt-1}, Z_{jt-1}]$ for $i > j$.

**Proposition 1:** Within the general class of functions defined by equation (6), and assuming the information instruments follow any continuous distribution, the functional form given by

$$
l_{ij,t}(\theta|\Omega_{t-1}) = \gamma_{ij0} + \gamma_{ij1}'Z_{agg,t-1} + \gamma_{ij2}Z_{lt-1}
$$

(8)

\(^2\) $\mathbf{L}_t$ is not a Cholesky decomposition of $\mathbf{M}_t$, as we do not require strictly positive diagonal entries in $\mathbf{L}_t$.

\(^3\) Our approach may be readily extended to consider a slightly more general specification,

$$
l_{ij,t}(\theta|\Omega_{t-1}) = \gamma_{ij0} + \gamma_{ij1}'f(Z_{agg,t-1}) + \gamma_{ij2}g(Z_{lt-1}) + \gamma_{ij3}h(Z_{jt-1})
$$

for functions of the information instruments $f, g, h$ with sufficient conditions to ensure that all $l_{ij,t}$ elements are nonzero with probability one.
ensures positive definiteness for the resultant covariance matrix $M_t$ with probability one for any nonzero coefficient $y_{ij0}$ and satisfies the strong invariance to variate order property. Specifications in which either $y_{ij1} = 0$ or $y_{ij2} = 0$ for all $i, j = 1, 2, \ldots, N$ also trivially satisfy positive definiteness and strong invariance to variate order.

Proof: See Appendix.

Our proposed specification is similar in spirit to a simple Cholesky decomposition of the conditional covariance matrix that is often used in constant covariance applications where positive definiteness is required. Because every positive definite matrix yields a unique Cholesky decomposition, we build our specification for second moments within the lower triangular matrix context. Given our concern with uniqueness in the covariance specification and not the lower triangular matrix, $L_t$, we do not impose the familiar Cholesky condition that diagonal elements must be strictly positive. Our approach requires only that diagonal elements are nonzero to ensure a positive definite covariance matrix.\(^4\) The invariance property provides the structure to develop our model and produces a rich and parsimonious functional form not available in the extant literature.

The number of parameters required to estimate the proposed system is improved over existing models. In general each of the $\frac{N(N+1)}{2}$ unique $I_{ij}$ terms requires a specification in our context. In the case of equation (8) with a scalar $Z_{agg,t-1}$ this results in $\frac{3N(N+1)}{2}$ total covariance parameters.

Standard ML approaches may be used to obtain parameter estimates. Under standard regularity conditions, parameter estimates converge to population values. According to the functional invariance property of ML, the resultant conditional covariance matrices are ML and therefore are consistent estimators of the true conditional covariances.

### 2.2 Comparisons with existing models

Although our interest is primarily in economically motivated changes in second moment matrices, our approach can be compared to the extant time series literature by specifying our information instruments as

\(^4\) Unreported estimation results using the additional condition that diagonal elements of the lower triangular matrix, $L_t$, are positive for all times and all variates produce very similar results.
time series variables. In particular, we compare the multivariate ARCH representation of the BEKK model to a restricted version of equation (8) with no macroeconomic information instruments,

\[ l_{ijt}(\theta|\Omega_{t-1}) = \gamma_{ij0} + \gamma_{ij1}Z_{it,t-1}. \]  

(9)

A similar specification can be compared to the CCC or DCC models. To limit the parameterization of the linear information instrument and facilitate comparison with the factor ARCH model of Engle, Ng and Rothschild (1990), we consider the following restricted version of equation (8) with only macroeconomic information instruments,

\[ l_{ijt}(\theta|\Omega_{t-1}) = \gamma_{ij0} + \gamma_{ij1}^\prime Z_{agg,t-1} \]  

(10)

where \( Z_{it,t-1} \) and \( Z_{agg,t-1} \) are own and macroeconomic lagged model disturbances, respectively.

2.2.1 Comparison with the BEKK model

To compare our specification with the BEKK specification, we construct the conditional covariance matrices as

\[ M_t = L_t L_t' = [m_{ijt}], \text{ and} \]

\[ M_{BEKK} = C_0 C_0' + A_1 A_t r_{t-1} r_{t-1}' A_1', \]

(11) (12)

respectively, where \( C_0 \) and \( A_1 \) are \( N \times N \) parameter matrices and where \( C_0 \) is lower triangular. To estimate the BEKK covariance matrix, \( M_{BEKK} \), we require \( \frac{N(N+1)}{2} + N^2 \) parameters.\(^5\)

In the simplest bivariate case, \( M_t \) is given by

\[ M_t = \begin{bmatrix} m_{11t} & m_{12t} \\ m_{12t} & m_{22t} \end{bmatrix} \]

\(^5\) The most general representation of the multivariate ARCH model is the \texttt{vech} form (cf., Bollerslev, Engle and Wooldridge (1988)). Assuming the \texttt{vech} representation, the multivariate ARCH(1) model could be written as \( m_t = \omega + A r_{t-1}, \) where \( m_t = \texttt{vech}(M_t), \ r_t = \texttt{vech}(r_{1t}, r_{2t})' \) and \texttt{vech} is the operator that stacks the lower triangular part of a symmetric \( N \times N \) matrix \( M_t \) into a \( \frac{N(N+1)}{2} \) dimensional vector, where \( \omega \) and \( A_1 \) are parameter matrices with dimensions \( \frac{N(N+1)}{2} \times 1 \) and \( \frac{N(N+1)}{2} \times \frac{N(N+1)}{2} \), respectively. Because the \texttt{vech} representation does not explicitly deal with the inherent restrictions required in covariance estimation, we focus on the most familiar and widely used models.

\(^6\) There are \( \frac{N(N+1)}{2} \) and \( N^2 \) parameters in \( C_0 \) and \( A_1 \), respectively.
with

\[ m_{11t} = \gamma_{110}^2 + 2\gamma_{110}\gamma_{111}Z_{t-1} + \gamma_{111}^2 Z_{t-1}^2, \]

\[ m_{21t} = m_{12t} = \gamma_{210}\gamma_{110} + \gamma_{210}\gamma_{111}Z_{t-1} + \gamma_{211}\gamma_{111}Z_{t-1} + \gamma_{211}^2 Z_{t-1}^2, \]

and

\[ m_{22t} = (\gamma_{210}^2 + \gamma_{220}^2) + 2(\gamma_{210}\gamma_{211} + \gamma_{220}\gamma_{221})Z_{t-1} + (\gamma_{211}^2 + \gamma_{221}^2)Z_{t-1}^2. \quad (13) \]

The related BEKK representation is given by

\[
M_{BEKK} = \begin{bmatrix}
  m_{BEKK11t} & m_{BEKK12t} \\
  m_{BEKK12t} & m_{BEKK22t}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
  c_{11} & 0 \\
  c_{21} & c_{22}
\end{bmatrix}
\begin{bmatrix}
  c_{11} & c_{21} \\
  0 & c_{22}
\end{bmatrix}
+ \begin{bmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
  Z_{1,t-1}^2 & Z_{1,t-1}Z_{2,t-1} \\
  Z_{1,t-1}Z_{2,t-1} & Z_{2,t-1}^2
\end{bmatrix}
\begin{bmatrix}
  a_{11} & a_{21} \\
  a_{12} & a_{22}
\end{bmatrix}
\]

\[ m_{BEKK11t} = c_{11}^2 + a_{11}^2 Z_{1,t-1}^2 + 2a_{11}a_{12}Z_{1,t-1}Z_{2,t-1} + a_{12}^2 Z_{2,t-1}^2, \]

\[ m_{BEKK21t} = m_{BEKK12t} = c_{11}c_{21} + a_{11}a_{21}Z_{1,t-1}^2 \\
+ (a_{12}a_{21} + a_{11}a_{22})Z_{1,t-1}Z_{2,t-1} + a_{12}a_{22}Z_{2,t-1}^2, \]

and

\[ m_{BEKK22t} = c_{21}^2 + c_{22}^2 + a_{21}^2 Z_{1,t-1}^2 + 2a_{21}a_{22}Z_{1,t-1}Z_{2,t-1} + a_{22}^2 Z_{2,t-1}^2. \quad (14) \]

A direct comparison of equations (13) and (14) reveals the following similarities and differences between our linear information instrument model and the BEKK model. First, both models satisfy the \textit{strong invariance to variate order} property. Second, our model given by equation (9) has fewer parameters than the BEKK model.

In particular, for any \( N \times N \) covariance matrix, \( M_t \) has \( N(N+1) \) parameters using equation (9), where \( M_{BEKK} \) requires \( \frac{N(N+1)}{2} + N^2 \) parameters. Thus, it is straightforward to show that for any \( N \geq 2, \frac{N(N+1)}{2} + N^2 > N(N+1) \).

Third, our covariance matrix, \( M_t \), is the unique covariance matrix representation given the general class of models described in equation (6). The BEKK specification may have good empirical properties, especially if our linear instrument specification does not provide a good empirical description of second moment dynamics. Fourth, our conditional covariance matrix has a different functional form relative

---

7 In many empirical applications a restricted form of \( M_{BEKK} \) is considered in which the parameter matrix \( A_1 \) is a diagonal matrix (diagonal BEKK) or \( A_1 \) is equal to a scalar parameter times the identity matrix (scalar BEKK). In these restricted versions of \( M_{BEKK} \), the number of parameters are reduced. For example, the diagonal BEKK model requires \( \frac{N(N+1)}{2} + N \) parameters and the scalar BEKK model has \( \frac{N(N+1)}{2} + 1 \) parameters. For brevity, we avoid a comparison of restricted versions of the BEKK model and similarly restricted versions of our linear information instrument model.
to BEKK. Each $i,j^{th}$ element of the BEKK covariance matrix is a function of $Z_{i,t-1}^2, Z_{j,t-1}^2$, and $Z_{i,t-1}Z_{j,t-1}$.

In contrast, our approach yields a covariance matrix whose variances are solely functions of own instruments and whose covariances between disturbance $i$ and $j$ are functions of $Z_{i,t-1}, Z_{j,t-1}$, and $Z_{i,t-1}Z_{j,t-1}$. Our approach should therefore have the ability to capture interesting own effects involving the lower moments of the instruments that may be obscured within the BEKK functional form. Our particular empirical application may not reveal the full benefits to these nonlinear covariance terms; however, the rapid growth in highly nonlinear hedge fund payoffs may provide an interesting application for our model in further research. We address this question further in our empirical analysis of these models in describing realized covariance matrices.

2.2.2 Comparison with the factor ARCH model

Factor models provide another popular form of covariance specification motivated by economic theory (cf., Engle, Ng and Rothschild (1990), van der Weide (2002), Vrontos et al. (2003), Lanne and Saikkonen (2007)). For example, Engle, Ng and Rothschild (1990) propose a factor structure for the conditional covariance matrix as,

$$ M_{Ft} = \Omega + \sum_{k=1}^{K} \beta_k \beta_k' f_{k,t} $$

(15)

where $\Omega$ is an $N \times N$ positive semi-definite matrix; $\beta_k, k = 1, ..., K$, is a linearly independent $N \times 1$ weight vector for factor $k$; and $f_{k,t}$ is the $k^{th}$ factor that is derived from the stochastic process of $r_t$.

Assuming $f_{k,t}$ follows a first-order ARCH structure, we have,

$$ f_{k,t} = \lambda_k + \phi_k (r_{t-1}' \omega_k)^2 $$

(16)

where $\lambda_k$, and $\phi_k$ are scalar parameters and $\omega_k$ is an $N \times 1$ vector of weights.

The model can be estimated using a two-step ML method as described in Engle, Ng and Rothschild (1990). In particular, consistent estimates of $f_{k,t}$ are typically first obtained by ML. The estimates of $\beta_k$ are then obtained in a second step using first stage consistent estimates of $f_{k,t}$.
To compare our specification with the factor ARCH model, we consider a simple bivariate case ignoring the typical multistep estimation procedure. For simplicity and comparison with the single factor ARCH model, we consider a scalar macroeconomic information instrument \( Z_{agg,t-1} = r_{t-1}' \omega \) in equation (10).

The conditional covariance matrix of our linear information instrument model is given by,

\[
M_t = \begin{bmatrix}
m_{11t} & m_{12t} \\
m_{21t} & m_{22t}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\gamma_{110} + \gamma_{111}(r_{t-1}' \omega) & 0 \\
\gamma_{210} + \gamma_{221}(r_{t-1}' \omega) & \gamma_{220} + \gamma_{221}(r_{t-1}' \omega)
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\gamma_{110} + \gamma_{111}(r_{t-1}' \omega) & \gamma_{210} + \gamma_{221}(r_{t-1}' \omega) \\
0 & \gamma_{220} + \gamma_{221}(r_{t-1}' \omega)
\end{bmatrix}
\]

(17)

where \( r_{t-1} \) is a 2 \times 1 vector of lagged model disturbances and \( \omega \) is a 2 \times 1 vector of weights. Evaluating term by term, we have,

\[
m_{11t} = \gamma_{110}^2 + 2\gamma_{110}\gamma_{111}(r_{t-1}' \omega) + \gamma_{111}^2(r_{t-1}' \omega)^2,
\]

\[
m_{21t} = m_{12t} = \gamma_{110}\gamma_{210} + (\gamma_{110}\gamma_{221} + \gamma_{210}\gamma_{111})(r_{t-1}' \omega) + \gamma_{111}\gamma_{221}(r_{t-1}' \omega)^2,
\]

\[
m_{22t} = \gamma_{210}^2 + \gamma_{220}^2 + 2(\gamma_{210}\gamma_{221} + \gamma_{220}\gamma_{221})(r_{t-1}' \omega)
\]

\[+ 2\gamma_{221}^2(r_{t-1}' \omega)^2\]

(18)

For the factor ARCH model with a single factor, the conditional covariance matrix is given by,

\[
M_{Ft} = \begin{bmatrix}
m_{F11t} & m_{F12t} \\
m_{F21t} & m_{F22t}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\gamma_{11} & \gamma_{21} \\
\gamma_{21} & \gamma_{22}
\end{bmatrix} + \begin{bmatrix}
\beta_1 \\
\beta_2
\end{bmatrix}\begin{bmatrix}
\lambda + \phi(r_{t-1}' \omega)
\end{bmatrix}
\]

(19)

Evaluating term by term, we get,

\[
m_{F11t} = \gamma_{11} + \beta_1^2[\lambda + \phi(r_{t-1}' \omega)^2],
\]

\[
m_{F21t} = m_{F12t} = \gamma_{21} + \beta_1\beta_2[\lambda + \phi(r_{t-1}' \omega)^2],\text{ and}
\]

\[
m_{F22t} = \gamma_{22} + \beta_2^2[\lambda + \phi(r_{t-1}' \omega)^2].
\]

(20)

---

8 In a general case, to match with the \( N \) variate, \( K \) factor ARCH model, we define \( Z_{agg,t-1} = r_{t-1}' \omega \), where \( r_{t-1} \) is an \( N \times 1 \) vector of lagged model disturbances and \( \omega \) is an \( N \times K \) matrix of weights. We do not consider potential restrictions to ensure weights sum to one, or other potential interesting restrictions on the weight vector that are commonly used in practice. Instead, we focus only on differences in these specifications for a given weight vector.
Comparing equations (18) and (20), we observe that our linear information instrument model differs from the factor ARCH model in a number of manners. First, our linear information instrument model has a richer specification than the factor ARCH model. In particular, our linear information instrument model involves terms in both \( r_{it-1} \) and \( r_{it-1} r_{jt-1} \). In contrast, the factor ARCH model only contains terms in \( r_{it-1} r_{jt-1} \). Second, our linear information instrument model has fewer parameters than the factor ARCH model. In particular, for any \( N \times N \) covariance matrix with \( K \) factors, our linear information instrument model requires \( \frac{N(N+1)}{2} + 2NK \) parameters, while the factor ARCH model has \( \frac{N(N+1)}{2} + 2NK + 2K \) parameters. Third, our linear information instrument model requires fewer restrictive assumptions than the factor ARCH model to determine the number of factors.

2.2.3 Comparison with the DCC model

Correlation models decompose the conditional covariance matrix into conditional standard deviations and conditional correlations. This class of models includes, among others, the CCC model by Bollerslev (1990) and the DCC model by Engle (2002), or Tse and Tsui (2002). The DCC model is an extension of the CCC model to allow time-varying conditional correlations. A typical DCC model is defined as,

\[
D_t = \rho_{ij} D_t (21)
\]

where \( D_t = \text{diag}(\sqrt{d_{1t}}, \ldots, \sqrt{d_{nt}}) \) and \( \rho_t = [\rho_{ij}] \) is a positive correlation matrix with \( \rho_{ii} = 1 \) for \( i = 1, \ldots, N \).

If conditional variances follow a first-order ARCH structure, they can be written as,

\[
d_t = \omega + A_1 r_{t-1} \odot r_{t-1} (22)
\]

where \( \omega \) is an \( N \times 1 \) vector; \( A_1 \) is a diagonal \( N \times N \) matrix; and \( \odot \) denotes the Hadamard matrix product operator (i.e., elementwise multiplication).

Engle (2002) defines the dynamics of the conditional correlation matrix, \( R_t \), in a multi-step procedure. First, consider a dynamic matrix process,

\[
Q_t = (1 - a - b)S + ar_{t-1} r_{t-1}' + bQ_{t-1} (23)
\]
where \( a \) and \( b \) are parameters, and \( S \) is the unconditional covariance matrix of \( r_t \). Under the condition that \( a + b < 1 \) and \( Q_0 \) is positive definite, \( Q_t \) is positive definite. The conditional correlation matrix \( R_t \) is then obtained as,

\[
R_t = (I \otimes Q_t)^{-1/2} Q_t (I \otimes Q_t)^{-1/2}
\]

(24)

where \( \otimes \) denotes the Kronecker matrix product operator. Positive definiteness of \( M_{DCC} \) is ensured if \( R_t \) is positive definite, and if the elements of \( \omega \) and the diagonal elements of \( A_1 \) are positive. The estimates of the DCC model can be obtained using ML.\(^9\)

Previous literature has shown that the DCC and BEKK models perform similarly in forecasting conditional covariances. For example, Caporin and McAleer (2008) find that the scalar versions of the two models are similar in forecasting conditional covariances and value-at-risk thresholds. Moreover, Massimiliano and Michael (2010) suggest that the DCC and BEKK models produce highly comparable conditional covariances and correlations in both univariate and large scale contexts. Given these findings, and our primary interest in own effects, we focus on a detailed comparison of our linear information instrument model and the BEKK model.

### 3. Empirical Analysis

We present empirical results for the covariance matrix estimation using our model and a comparable specification of the BEKK model. Our empirical analysis approach may be described as follows. First, we construct a set of realized conditional covariance matrices that we treat as population values. We then construct return samples by taking consecutive draws from these matrices for various intertemporal multivariate subsamples. Because our interest is in the temporal properties of the conditional covariance matrix, we always retain the temporal links in the generated covariance matrices. We begin our analysis with

\(^9\) Tse and Tsui (2002) propose a similar structure for the correlation matrix, \( R_t = (1 - a - b)S + aS_{t-1} + bR_{t-1} \), where \( S \) is a positive definite parameter matrix with ones on the diagonal, \( a \) and \( b \) are scalar valued, and \( S_{t-1} \) is the sample correlation matrix of the past \( M \) model disturbances. In comparison with Engle (2002), the Tse and Tsui (2002) specification uses considerably more parameters that arise in their specification of \( S \) within \( R_t \).
an example showing that the constant covariance matrix is readily estimated using a simplified version of the linear information instrument model. Next, we compare estimates of the conditional covariance matrix using both the linear information instrument model and the BEKK model. We use an average of previous daily absolute returns over a four week window (beginning one week prior) as our information instrument to allow a comparison between models. In our final table we present Monte Carlo simulation results for both covariance specifications using various return vector dimensions ($N = 3$, or $5$) and sample sizes ($T = 260$, $1,040$, or $2,419$) for 1,000 replications.

3.1 Conditional covariance matrix construction

We construct weekly conditional covariance matrices from the daily returns of the value-weighted BM quintile portfolios from July 5, 1963 to December 30, 2009. To avoid spurious weekend and holiday effects, we define a week from Thursday-open to Wednesday-close. We then construct the weekly realized conditional covariance matrices as follows. First, for each Wednesday ending on day $t$, we compute the weekly realized covariance matrices from the past eight-weeks of daily returns as,

$$\text{cov}_t = 5 \times \frac{1}{N_{\text{days}}} \sum_{j=0}^{N_{\text{days}}} r_{t-j} r_{t-j}'$$

(25)

where $N_{\text{days}}$ is the number of days in the past eight-week window, and $r_{t-j}$ is a vector of percentage daily returns on day $t - j$. Similar to Anderson et al. (2001) and the empirical Mixed Data Sampling (MIDAS) results from Ghysels et al. (2006), we create weekly realized conditional covariance matrices by multiplying each daily realized covariance matrix by five for a standard trading week. We treat this time series of realized

10 The portfolio return data are downloaded from French’s website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
11 Our sample begins on Friday, July 5, 1963 as Thursday, July 4, 1963 represents a market closure for the Independence Day holiday.
12 By choosing Thursday-open to Wednesday-close weekly periods, we mitigate the impact of many market closures and provide the greatest number of possible five day weeks. The scaling factor $N_{\text{days}}$ reflects the number of daily returns employed for any given eight-week window of daily data. Because our analysis is focused at the portfolio level, missing observations reflect US market closures. Our daily sample contains 2,419 Thursday to Wednesday calendar weeks of which 2,413 have either four (413) or five (2000) trading days per week. The worst week for available data was the period from Thursday, July 4 to July 10, 1968. The market was closed on July 4 (Independence Day), July 5 (the day after Independence Day), July 6 and 7 (the weekend), and July 10 (due to the 1968 Paperwork crisis). This produced a two-day
covariance matrices as the population covariance matrices to be subsequently estimated using either the linear information instrument model or the BEKK model in the empirical analysis.\textsuperscript{13}

### 3.2 Conditional covariance matrix estimation

At each point of time \( t \), we draw a weekly return vector \( r_t \) from a multivariate normal distribution with mean of zero and covariance matrix of \( \mathbf{Cov}_t \) as developed in equation (25). We treat this multivariate time series as our base case sample for later empirical simulations. Using our sample of daily returns from July 5, 1963 to December 30, 2009, we create 2,419 weekly realized covariances for the base case period from Thursday, August 22, 1963 to Wednesday, December 30, 2009. The ML estimator of the covariance matrix, \( \frac{1}{T} \sum_{t=1}^{T} (r_t - \bar{r})(r_t - \bar{r})' \), for the \( 5 \times 1 \) set of book-to-market simulated portfolio returns, \( r_t \), with sample mean \( \bar{r} \), is reported for the weekly sample of 2,419 simulated observations in Table 1. The simulated series is based on a normal draw with a \( 5 \times 1 \) zero mean vector and a time-varying covariance matrix given by equation (25).

*** Insert Table 1 about here ***

\textsuperscript{13} In an unreported table, we show that our empirical results are robust to different realized conditional covariance construction methods including modifications to consider a) the previous four weeks of daily returns, b) the previous twelve weeks of daily returns, and c) different decay weights of 0.9, 0.7, and 0.5 for lags of daily squared returns. Although the general magnitudes of the model performance metrics change with model assumptions, the superior relative performance for the linear instrument model relative to the BEKK model persists in all cases.
As a point of departure for our empirical analysis, we first report the estimated constant covariance matrix using our ML routine for the simulated series. For the case of a constant covariance matrix, we parameterize the lower triangular matrix \( L = [l_{ij}] \) as,

\[ l_{ij} = y_{ij} \]

for \( i \geq j, \ i, j = 1, 2, \ldots, 5 \); and where \( l_{ij} \) is the \( i, j \)th element of \( L \). The covariance matrix \( M \) is then constructed as,

\[ M = LL'. \]

Table 2 reports the constant covariance matrix in lower triangular form, \( L = [y_{ij}] \), along with associated standard errors in parentheses.15

*** Insert Table 2 about here ***

We note that all \( y_{ij} \) terms are highly significant. In addition, we find that the resultant estimated unconditional covariance matrix, \( M = LL' \), is virtually identical to the ML estimate of the covariance matrix in Table 1.

For our conditional covariance estimation, we consider both the linear information instrument model given by equation (9) versus the BEKK model after computation of \( M_t \) using equation (11). The information instrument for both models is \( Z_{t-1} = [Z_{t-1}, Z_{t-1}] \), where \( Z_{t-1} \) is the instrument variable realized at week \( t-1 \).

For our instrument, we average the daily absolute returns over the calendar period beginning one week prior to week \( t \), to five weeks prior, for a four week window in total. To prevent potential scale problems, we scale this variable by multiplying by 500.16 Our instrument choice provides a natural comparison to the BEKK

---

14 The log likelihood function is given by 
\[ \log(\theta | \mathbf{r}_t) = -\frac{1}{2} \sum_{t=1}^{T} \left[ N \log(2\pi) + \log|M_t| + \mathbf{r}_t \cdot M_t^{-1} \mathbf{r}_t \right], \]

where \( M_t \) is the conditional covariance matrix at \( t \).

15 Starting values for the mean vector are given by zeros and the initial value of \( y_{ij} \) is the \( i, j \)th element of the Cholesky decomposition of unconditional sample covariance matrix. We use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method within Matlab.

16 We find the scaling has little impact on the resultant covariance matrix estimates, and in retrospect there is little need for any scaling in the estimation routines. We adopt this modified instrument for both the linear instrument and BEKK models as it results in a greater than 300 percent improvement in the MAPE for both the linear instrument and BEKK models in comparison with a single lag of previous returns.
model after computation of \( M_t \).\(^{17}\) To specify the BEKK model, we adopt equation (12) for \( M_{BEKK} \), where

\[
\mathbf{r}_{t-1} = \mathbf{Z}_{t-1}.\(^{18, 19}\)
\]

Parameter estimates and standard errors are reported in Table 3.

*** Insert Table 3 about here ***

Table 3 reports results for the linear information instrument model in Panel A and for the BEKK Model in Panel B. In Panel A we observe that virtually all coefficients are highly significant. The sole exception appears to be the estimate of \( \gamma_{420} \). Panel B shows that the full BEKK model has a considerably larger number of parameters, and further many of these coefficients are not significant. The constant component of the BEKK specification is highly significant for all portfolios. Interestingly, all above diagonal terms in the \( A_1 \) coefficient matrix are not significant. The lower triangular portion of the \( A_1 \) coefficient matrix, and especially the diagonal elements, appear much more significant.

To provide an alternative method to gauge the parameter estimates in Panel B of Table 3, Table 4 presents the coefficient estimates and standard errors for the parameters of the conditional covariance matrix after expanding the \( M_t = L_t L_t' \) linear information instrument specification. The \( i,j \)th element in \( M_t, m_{ij,t} \), may be rewritten as

\[
m_{ij,t} = \phi_{ij0} + \phi_{ij1}Z_{t-1} + \phi_{ij2}Z_{t} + \phi_{ij3}Z_{t-1}Z_{t}, \quad \text{where} \quad \phi_{ij0} = \sum_{k=1}^{I} \gamma_{ik0}\gamma_{jk0},
\]

\[
\phi_{ij1} = 2 \sum_{k=1}^{I} \gamma_{ik1}\gamma_{jk0}, \quad \phi_{ij2} = 0, \quad \text{and} \quad \phi_{ij3} = \sum_{k=1}^{I} \gamma_{ik1}\gamma_{jk1} \quad \text{when} \quad i = j; \quad \text{and} \quad \phi_{ij0} = \sum_{k=1}^{I} \gamma_{ik0}\gamma_{jk0},
\]

\[
\phi_{ij1} = \sum_{k=1}^{I} \gamma_{ik1}\gamma_{jk0}, \quad \phi_{ij2} = \sum_{k=1}^{I} \gamma_{ik0}\gamma_{jk1}, \quad \text{and} \quad \phi_{ij3} = \sum_{k=1}^{I} \gamma_{ik1}\gamma_{jk1} \quad \text{when} \quad i \neq j. \quad \text{By the functional invariance property of ML, the coefficients estimates of all} \ \phi_{ij} \ \text{terms are also ML. Further, under standard}
\]

---

\(^{17}\) Initial values of the parameters for the linear information instrument model are obtained as follows. At each point of time \( t \), we construct a square matrix in which the \( i,j \)th element is \( e_{i,t}e_{j,t} \), where \( e_{i,t} = \frac{1}{2D} \sum_{k=0}^{2D} (r_{i,t-k} - \bar{r}_i) \) and \( \bar{r}_i \) is the time-series average of \( r_{i,t} \) over the entire sample period. We take the Cholesky decomposition of the square matrix \( [e_{i,t}e_{j,t}] \) at each time \( t \). For each element in the resultant lower triangular matrix after the decomposition, we run an ordinary least squares (OLS) regression on \( Z_{t-1} \) to obtain the starting parameter values.

\(^{18}\) Initial values for the BEKK model of the parameter matrix \( C_0 \) are obtained as the lower triangular matrix of the unconditional sample covariance matrix after Cholesky decomposition. The initial values of \( A_1 \) are set to 0.05 for each element. Robustness checks with alternative starting values suggest the conditional covariance matrix estimates are not particularly sensitive to the choice of the initial value of \( A_1 \). We also examined less precise starting values for the linear information instrument model to confirm that the superior empirical performance was not driven by starting parameter values.

\(^{19}\) For both models, we follow Engle and Kroner (1995) and estimate parameters using the simplex algorithm for the initial 20 iterations, and then the BFGS algorithm to obtain final parameter estimates.
regularity conditions, these ML estimates have an asymptotically normal distribution. Thus, we use the delta
method to derive the standard errors for the coefficient estimates.

For our application, let \( \gamma \) be a vector of parameters in \( L_t \) and \( h(\gamma) \) be a vector of coefficients in \( M_t \). According to the delta method, the coefficient estimate \( h(\hat{\gamma}) \) follows an asymptotically normal distribution

\[
\sqrt{T}(h(\hat{\gamma}) - h(\gamma)) \sim N(0, \nabla h(\gamma)' \cdot \Sigma \cdot \nabla h(\gamma))
\]

where \( T \) is the number of observations; \( \Sigma \) is the covariance matrix of \( \gamma \); and \( \nabla h(\gamma) \) is the gradient function of \( h(\gamma) \). We compute the standard errors of \( h(\hat{\gamma}) \) using \( \nabla h(\hat{\gamma})' \cdot \hat{\Sigma} \cdot \nabla h(\hat{\gamma}) \), where \( \hat{\Sigma} \) is the estimated covariance matrix of \( \gamma \) and is obtained as the inverse of the Fisher information matrix. We report coefficient estimates and standard errors in Table 4.

*** Insert Table 4 about here ***

Interpretation of the coefficients in the expanded model is simplified in that estimates relate to the
covariance matrix directly. For example, the coefficient estimate of \( \phi_{11} \) is 0.3664 and suggests that a unit
increase in the instrument value, \( Z_{1,t-1} \), will give rise to a 0.3664 unit increase in the smallest BM portfolio variance. In addition to this impact, we find a nonlinear impact related to \( Z_{1,t-1}^2 \). The coefficient estimate of \( \phi_{13} \) of 0.2246 is highly significant and suggests that, after controlling for absolute previous errors, a one unit increase in the squared instrument gives rise to a 0.2246 unit increase in the smallest BM portfolio variance. In general, we observe that the reported coefficients are highly significant and suggest strong sensitivities to all instruments at conventional levels.

3.3 Conditional covariance matrix estimation by information instrument percentile

Table 5 presents an alternative manner to gauge the importance of changes in the covariance matrix with underlying information instrument realizations. For each instrument considered we evaluate the 2.5, 25, 50, 75, and 97.5 percentile values of the information instruments within the sample period. We then report the
resultant conditional covariance matrix estimate for that vector of instrument realizations using the parameter estimates from Table 4.

Table 5 shows that the variability in the estimated covariance matrix elements is economically important. In particular, all of the (co)variance estimates based on the 97.5 percentile instrument realization is more than 15 times larger than the corresponding estimate based on the 2.5 percentile realization. Interestingly, we find more variability and a greater range in variances for the most extreme book to market portfolios, relative to the second, third, and fourth BM quartiles.

3.4 Loss evaluation

Following Lopez (2001), we use several loss functions to evaluate the relative accuracy of different covariance matrix estimates. In particular, we compute the MAE, the RMSE, and the HMSE for each element in the conditional covariance matrices estimated by the linear information instrument model and the BEKK model. The three loss functions are defined as follows,

\[
MAE_{ij} = \frac{1}{T} \sum_{t=1}^{T} \left| h_{i,j,t} - \sigma_{i,j,t} \right| \tag{27}
\]

\[
RMSE_{ij} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (h_{i,j,t} - \sigma_{i,j,t})^2} \tag{28}
\]

\[
HMSE_{ij} = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{h_{i,j,t}}{\sigma_{i,j,t}} - 1 \right)^2 \tag{29}
\]

where \( h_{i,j,t} \) is the forecast for the \( i, j^{th} \) element in the conditional covariance matrix at \( t \) and \( \sigma_{i,j,t} \) is the population parameter value.
To examine the statistical differences between the linear information instrument model covariance matrix and the BEKK model covariance matrix we use the Diebold and Mariano (1995) statistic for each of the three loss functions. The asymptotic Diebold-Mariano (DM) statistic is given by,

\[ DM_{ij} = \frac{\hat{d}}{\sqrt{\hat{\sigma}_d^2/T}} \sim N(0,1) \]  

(30)

where \( DM_{ij} \) is the DM statistic corresponding to the \( i,j \)th element in the conditional covariance matrix, \( \hat{d} \) is the sample mean of the difference series of loss function values between the two competing models, and \( \hat{\sigma}_d^2 \) is a consistent estimator of its variance. We report the DM statistics in Table 6 with significance levels denoted for the larger of the estimated values.

*** Insert Table 6 about here ***

Table 6 provides the element-wise calculations for the MAE, RMSE and HMSE for both models in Panels A, B and C, respectively. In addition, we also report significance levels for the DM test that the two losses are equal. At a glance, we observe that for all but one occurrence, the loss for the linear information instrument loss is smaller than the comparable BEKK loss. In this single exception fourth BM portfolio using the HMSE loss function, the difference is not significant. In sum, we have strong element-wise evidence that the linear information instrument model offers superior performance. On average, our model outperforms the BEKK model by 17 percent and 12 percent using the MAE and RMSE, respectively. When using the HMSE, the average outperformance is more than 60 percent. Using the Diebold-Mariano statistics, we find that the linear information instrument model significantly outperforms the BEKK model for all element-wise comparisons at the one percent level using both the MAE and RMSE, and for most element-wise comparisons using HMSE.
3.5 Finite sample simulations

Our previous results suggest that the linear information instrument model performs well in our five portfolio example for a lengthy time series. In this section we consider the generalizability of these findings to a subset of our assets, and for various finite samples. In all cases, we consider 1,000 replications of the procedure described in Section 3.2 with $T = 260, 1040$, and $2419$ and $N = 3$ and 5.\(^{20}\) For each replication for $T = 260$ or $1040$, we draw a uniform random variable between $t = 1$ and $t = 2419 - T$, so that all admissible dates for the series to start are equally plausible for each simulation replication. We then use the next $T = 260$ or $T = 1040$ covariance matrices for our estimation for that replication. The $N = 3$ subset is created by removing the second and fourth BM portfolios. For each of the 1,000 replications, we then calculate the mean absolute percentage error ($MAPE$) and the root mean square percentage error ($RMSPE$) for both the linear information instrument model and the BEKK model. The MAPE and RMSPE loss functions are defined as,

$$MAPE = \frac{1}{T} \sum_{t=1}^{T} \frac{|M_t - \Sigma_t e'|}{e' \Sigma_t e'}, \quad \text{and}$$

$$RMSPE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \frac{(M_t - \Sigma_t) \odot (M_t - \Sigma_t)e'}{e' \Sigma_t e'}}, \quad (32)$$

where $e$ is a conformable vector of ones; and $M_t$ and $\Sigma_t$ are the estimate and the population value for the conditional covariance matrix at time $t$, respectively; and $T$ is the number of estimated conditional covariance matrices. Summary results are reported in Table 7.

*** Insert Table 7 about here ***

Panels A and B report the MAPE and RMSPE for both models in the case of the smallest intertemporal sample of $T = 260$ for $N = 3$ and 5, respectively. Focusing on the average and median MAPE and RMSPE,

\(^{20}\) Assuming 52 weeks in a year, $T = 260$ and $T = 1040$ provide results for five and ten year estimation intervals, respectively.
we find that the linear information instrument model produces consistently smaller loss outcomes relative to the BEKK model. For example, when $N = 3$ the MAPE suggests the average error in our model is approximately 13 percent smaller (=47.36 - 34.60 percent) than for the BEKK model. Further, the linear instrument model appears to provide less variability in modeling errors across replications. For example in the $N = 3$ sampling experiment the standard deviation in the computed MAPE values is dramatically smaller for the linear information instrument model versus the BEKK Model (0.06 versus 0.14).

In our largest samples, when $T = 2419$ and $N = 3$, the linear information instrument model improves the mean MAPE by five percent (=0.3296-0.2764) and the mean RMSPE by nine percent (=0.4419-0.3503) relative to the BEKK model. The comparable results for $N = 5$ are an improvement in the mean MAPE of ten percent (=0.3945-0.2901) and an improvement of the RMSPE of 18 percent (=0.5516-0.3743). Similar results are found for the fifth percentile, median, and 95th percentile of the two accuracy measures.

As the sample size grows, we observe the expected reduction in the variability of MAPE across replications. For example, comparing $N = 3$ results again, the standard deviation in the MAPE values across replications shrinks from 0.06 to 0.03 to 0.02 as $T$ increases from 260 to 1040 to 2419, respectively, for the linear information instrument model. The comparable standard deviations for the BEKK model shrink from 0.14 to 0.14, to 0.03. Interestingly, the benefits of the linear information instrument model are likely to be most prevalent in finite samples suggesting our earlier results may be somewhat understated due to the large samples considered.

The benefits to increasing the sample size are more apparent in the larger case of $N = 5$ variates. For example, with a sample size of $T = 260$ we observe a range of RMSPE values across replications of 0.48 (=0.72-0.24), and 1.63 (=1.88-0.24) for the linear information instrument and BEKK models, respectively. As the sample size increases to 1,040 and 2,419, the range in RMSPE values fall further to 0.22 and then to 0.15
for the linear information instrument model. The comparable ranges for the BEKK model are dramatically larger at 1.0166 and 0.69,\textsuperscript{21}

3.6 Robustness to alternative realized conditional covariances

To ensure that our results are not driven in any meaningful manner by the method we use to construct realized conditional covariances, we consider a range of alternative approaches including lags of 12 weeks of daily returns in place of four weeks, as well as a decay function of previous squared returns with a decay weight of $\gamma^j$ where $\gamma$ is one of 0.9, 0.7, or 0.5. In general, the results are qualitatively unchanged. The proposed model produces smaller MAPE and RMSPE values across 1,000 replications that are highly comparable to the values reported in Table 7.

4. Concluding Remarks

The modeling of the second moments of financial asset returns plays an important role in risk management, derivative pricing, hedging, and portfolio optimization. We propose a new model for conditional covariance estimation based on available information instruments. Our model is based on a restricted Cholesky-like decomposition and ensures positive definiteness and invariance to variate order by construction. Comparing to existing time-series models, our instrument model provides a parsimonious and accurate description of second moments. The proposed approach is expected to be especially valuable in the context of important own effects, where variances are likely to be driven by variate specific information. Simulation results suggest that our linear information instrument model has higher estimation accuracy than the BEKK model.

The linear information instrument model should provide a natural application in an asset pricing context where return moments evolve with predetermined economic information instruments. Our model can be used to estimate both time varying risk premiums and asset pricing risks that arise as a function of asset

\textsuperscript{21} We observe an interesting anomaly as $T$ increases from 1,040 to 2,419 where the loss measures increase with sample size. After further investigation, we hypothesize that this is due to a few very large realized covariance matrices (in the latter portion of our sample period) that are more likely to be included in the experiment as $T$ increases. To examine the cause of this finding, we consider the linear information instrument model at $T = 60, 130$ or 260 and $N = 3$ for 1,000 replications. These results show a monotonic decrease in measured losses as $T$ increases.
moments and economic fundamentals. In sum, we add to the burgeoning conditional covariance literature with a simple, yet parsimonious manner to incorporate economic information. Future work may be helpful to consider different functional forms and instrument choices within the decomposition suggested.
References


Appendix

**Proposition 1:** Within the general class of functions defined by equation (6), and assuming the information instruments follow any continuous distribution, only the functional form given by

\[
l_{ijt}(\theta|\Omega_{t-1}) = \gamma_{ij0} + \gamma_{ij1}Z_{agg,t-1} + \gamma_{ij2}Z_{lt-1}
\]

ensures positive definiteness for the resultant covariance matrix \(M_t\) with probability one for any nonzero coefficients \(\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2}\), and satisfies the \textit{strong invariance to variate order} property.

**Proof:**

To see that equation (A1) results in a positive definite covariance matrix, \(M_t\), note that because \(L_t\) is a lower triangular matrix, the determinant of \(L_t\) is given as \(\text{Det}(L_t) = \prod_{i=1}^{N} l_{iit}\), where \(l_{iit}\) represents the \(i\)th diagonal element in \(L_t\). For any nonzero coefficients \(\gamma_{ij0}, \gamma_{ij1}, \gamma_{ij2}\), we have that \(l_{iit} \neq 0\) with probability one. Therefore, \(\text{Det}(L_t) \neq 0\), hence \(L_t\) is invertible. According to the equivalent conditions of matrix positive definiteness, for any lower triangular invertible matrix \(L_t\), we have that \(L_tL_t'\) is positive definite. Therefore, \(M_t\) is positive definite.

To show that equation (A1) satisfies the \textit{strong invariance to variate order} property, we first show sufficiency followed by necessity within the class of linear functions defined by equation (6).

1) \textit{Sufficiency}

Consider the subgroup of linear functions given by equation (A1).

Given \(M_t = L_tL_t'\), the \(j\)th diagonal element in \(M_t\), \(m_{iit}\), is

\[
m_{iit} = \sum_{k=1}^{i} l_{ikt}^2 = \varphi_{iit0} + \varphi_{iit1}Z_{agg,t-1} + \varphi_{iit2}Z_{lt-1} + \varphi_{iit3}Z_{agg,t-1} \otimes Z_{agg,t-1} + \varphi_{iit4}Z_{agg,t-1}Z_{lt-1} + \varphi_{iit5}Z_{lt-1}^2
\]

where \(\varphi_{iit0} = \sum_{k=1}^{i} \gamma_{ik0}^2\), \(\varphi_{iit1} = \sum_{k=1}^{i} 2\gamma_{ik0}\gamma_{ik1}\), \(\varphi_{iit2} = \sum_{k=1}^{i} 2\gamma_{ik0}\gamma_{ik2}\), \(\varphi_{iit3} = \sum_{k=1}^{i} \gamma_{ik1} \otimes \gamma_{ik1}\), \(\varphi_{iit4} = \sum_{k=1}^{i} 2\gamma_{ik2}\gamma_{ik1}\), \(\varphi_{iit5} = \sum_{k=1}^{i} \gamma_{ik2}^2\), and \(\otimes\) denotes the Kronecker product operator.

The \(m_{ijt}\) element is

\[
m_{ijt} = \sum_{k=1}^{i} l_{ikt}l_{jkt} = \phi_{ij0} + \phi_{ij1}Z_{agg,t-1} + \phi_{ij2}Z_{lt-1} + \phi_{ij3}Z_{jt-1} + \phi_{ij4}Z_{agg,t-1} \otimes Z_{agg,t-1} + \phi_{ij5}Z_{agg,t-1}Z_{lt-1} + \phi_{ij6}Z_{agg,t-1}Z_{jt-1} + \phi_{ij7}Z_{lt-1}Z_{jt-1},
\]
for any \( i > j \), and where \( \phi_{ij0} = \sum_{k=1}^{j} y_{ik0} y_{jk0} \), \( \phi_{ij1} = \sum_{k=1}^{j} (y_{ik0} y_{jk1} + y_{ik1} y_{jk0}) \), \( \phi_{ij2} = \sum_{k=1}^{j} y_{ik2} y_{jk0} \), \( \phi_{ij3} = \sum_{k=1}^{j} y_{ik0} y_{jk2} \), \( \phi_{ij4} = \sum_{k=1}^{j} y_{ik1} y_{jk1} \), \( \phi_{ij5} = \sum_{k=1}^{j} y_{ik2} y_{jk1} \), \( \phi_{ij6} = \sum_{k=1}^{j} y_{ik1} y_{jk2} \), and \( \phi_{ij7} = \sum_{k=1}^{j} y_{ik2} y_{jk2} \).

Note that \( m_{ijt} \) is a linear function of \( Z_{agg,t-1}, Z_{i,t-1}, Z_{agg,t-1} \otimes Z_{agg,t-1}, Z_{i,t-1}^2 \); where \( m_{ijt} \) is a function of \( Z_{agg,t-1}, Z_{i,t-1}, Z_{agg,t-1} \otimes Z_{agg,t-1}, Z_{agg,t-1} \otimes Z_{i,t-1}, Z_{agg,t-1} Z_{i,t-1} \); \( Z_{agg,t-1} Z_{i,t-1} \), and \( Z_{i,t-1} Z_{i,t-1} \). Therefore, \( M_t \) is invariant to variate order.

2) Necessity

By way of contradiction, we will show that any subgroup of the defined class of linear functions other than equation (A1) violates the strong invariance to order property.

In the case of the general specification given by equation (6), we need only show that strong invariance is violated in a simple \( 2 \times 2 \) case when \( Z_{i,t-1} \) are admitted into the empirical specification. Consider a lower triangular matrix \( L_t = [l_{ijt}] \) with

\[
 l_{ijt} = y_{ij0} + y_{ij1}' Z_{agg,t-1} + y_{ij2} Z_{i,t-1} + y_{ij3}' Z_{i,t-1} \tag{A2} 
\]

Given \( M_t = L_t L_t' \), it is straightforward to confirm that

\[
 m_{11t} = y_{110}^2 + (2y_{110} y_{111})' Z_{agg,t-1} + (2y_{110} y_{112} + 2y_{110} y_{113}) Z_{i,t-1} \\
 + (y_{111} y_{111})' (Z_{agg,t-1} \otimes Z_{agg,t-1}) + (2y_{111} y_{112} + 2y_{111} y_{113})' Z_{agg,t-1} Z_{i,t-1} \\
 + (y_{112}^2 + 2y_{112} y_{113} + y_{113}^2) Z_{i,t-1}^2 
\]

and

\[
 m_{22t} = (y_{210}^2 + y_{220}^2) + (2y_{210} y_{211} + 2y_{220} y_{221})' Z_{agg,t-1} + 2y_{210} y_{213} Z_{i,t-1} \\
 + (2y_{210} y_{212} + 2y_{220} y_{222} + 2y_{220} y_{223}) Z_{i,t-1} \\
 + (y_{211} y_{211} + y_{221} y_{221})' (Z_{agg,t-1} \otimes Z_{agg,t-1}) + (2y_{211} y_{213})' Z_{agg,t-1} Z_{i,t-1} \\
 + (2y_{211} y_{212} + 2y_{221} y_{222} + 2y_{221} y_{223})' Z_{agg,t-1} Z_{i,t-1} \\
 + 2y_{212} y_{213} Z_{i,t-1} Z_{i,t-1} + y_{213}^2 Z_{i,t-1}^2 \\
 + (y_{212}^2 + y_{222}^2 + y_{223}^2 + 2y_{222} y_{223}) Z_{i,t-1}^2 
\]

22 Because \( M_t \) is symmetric, it follows that \( m_{ijt} = m_{jit} \), for any \( i > j \).
Observe that $m_{11t}$ is a function of $Z_{agg,t-1}, Z_{1,t-1}, Z_{agg,t-1} \otimes Z_{agg,t-1}, Z_{agg,t-1} Z_{1,t-1}$, and $Z_{1,t-1}^2$.

In contrast, $m_{22t}$ is a function of $Z_{agg,t-1}, Z_{1,t-1}, Z_{2,t-1}, Z_{agg,t-1} \otimes Z_{agg,t-1}, Z_{agg,t-1} Z_{1,t-1}, Z_{agg,t-1} Z_{2,t-1}, Z_{1,t-1}^2$, and $Z_{2,t-1}^2$. Thus, equation (A2) violates both the weak and strong invariance to order properties.

The only other alternative linear specification from equation (6) is given by $L_t = [l_{ijt}]$ with

$$l_{ijt} = y_{i0} + y_{ijt}' Z_{agg,t-1} + y_{ij3} Z_{jt-1}$$  \hfill (A3)

We can now construct $M_t = L_t L_t'$, to demonstrate that

$$m_{11t} = y_{110}^2 + (2y_{110} y_{111})' Z_{agg,t-1} + 2y_{110} y_{113} Z_{1,t-1}$$

$$+ (y_{111} y_{111})' (Z_{agg,t-1} \otimes Z_{agg,t-1}) + (2y_{111} y_{113})' Z_{agg,t-1} Z_{1,t-1} + y_{113}^2 Z_{1,t-1}^2$$

and

$$m_{22t} = (y_{210}^2 + y_{220}^2) + (2y_{210} y_{211} + 2y_{220} y_{221})' Z_{agg,t-1} + 2y_{210} y_{213} Z_{1,t-1}$$

$$+ 2y_{220} y_{223} Z_{2,t-1} + (y_{211} y_{211} + y_{221} y_{221})' (Z_{agg,t-1} \otimes Z_{agg,t-1})$$

$$+ (2y_{211} y_{213})' Z_{agg,t-1} Z_{1,t-1} + (2y_{221} y_{223})' Z_{agg,t-1} Z_{2,t-1}$$

$$+ y_{213}^2 Z_{1,t-1}^2 + y_{223}^2 Z_{2,t-1}^2$$

We observe that $m_{11t}$ is a function of only $Z_{agg,t-1}$ and $Z_{1,t-1}$, where $m_{22t}$ is a function of $Z_{agg,t-1}$, $Z_{1,t-1}$, and $Z_{2,t-1}$. Thus, equation (A3) violates both the weak and strong invariance to order property.

Therefore, all subgroups of the defined class of linear functions except equation (A1) violate the strong invariant to order property. Thus, within the class of linear functions having both aggregate and own information instruments, any model that satisfies the strong invariant to order property must take the form of equation (A1).
Table 1
Descriptive statistics for the base case sample.

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallest</td>
<td>5.3747</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>4.3709</td>
<td>4.2613</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>4.1067</td>
<td>3.9191</td>
<td>4.1303</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>3.8184</td>
<td>3.6828</td>
<td>3.7135</td>
<td>3.9825</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>4.0617</td>
<td>3.8932</td>
<td>3.8856</td>
<td>3.9678</td>
<td>4.7037</td>
</tr>
</tbody>
</table>

We report the maximum likelihood estimator of the covariance matrix for the simulated weekly return sample of 2419 observations for book-to-market quintile portfolios. The simulated series is based on a normal draw with zero mean and time-varying covariance given by equation (25).
Table 2
Unconditional sample mean and covariance estimation.

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3183</td>
<td>(0.0334)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1.8854</td>
<td>0.8407</td>
<td>(0.0321)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1.7714</td>
<td>0.6892</td>
<td>0.7194</td>
<td>(0.0326)</td>
<td>(0.0177)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>1.6471</td>
<td>0.6869</td>
<td>0.4483</td>
<td>0.7725</td>
<td>(0.0330)</td>
<td>(0.0207)</td>
</tr>
<tr>
<td>1.7520</td>
<td>0.7019</td>
<td>0.4149</td>
<td>0.5360</td>
<td>0.8260</td>
<td>(0.0363)</td>
</tr>
</tbody>
</table>

We report the ML estimates of the unconditional covariance matrix for the simulated weekly return sample of 2419 observations. The unconditional covariance matrix $\mathbf{M}$ is modeled as $\mathbf{M} = \mathbf{L} \mathbf{L}'$, where $\mathbf{L} = [l_{ij}]$ is a lower triangular matrix with the $i,j^{th}$ element $l_{ij} = \gamma_{ij}$. We report the ML parameter estimates of $\gamma_{ij}$ along with standard errors in parentheses.
Table 3
Conditional covariance matrix estimation.

Panel A: ML estimates for linear information instrument model

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Coefficient</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\gamma_{10}$</td>
<td>0.3866 (0.0476)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{11}$</td>
<td>0.4739 (0.0163)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>$\gamma_{20}$</td>
<td>0.2201 (0.0412)</td>
<td>0.3986 (0.0214)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{21}$</td>
<td>0.4580 (0.0157)</td>
<td>0.1094 (0.0072)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd</td>
<td>$\gamma_{30}$</td>
<td>0.1616 (0.0402)</td>
<td>0.1015 (0.0318)</td>
<td>0.3574 (0.0211)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{31}$</td>
<td>0.4550 (0.0163)</td>
<td>0.1460 (0.0116)</td>
<td>0.0913 (0.0074)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th</td>
<td>$\gamma_{40}$</td>
<td>0.2248 (0.0430)</td>
<td>0.0459 (0.0360)</td>
<td>0.0857 (0.0328)</td>
<td>0.3385 (0.0215)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{41}$</td>
<td>0.4196 (0.0176)</td>
<td>0.1563 (0.0137)</td>
<td>0.0793 (0.0121)</td>
<td>0.1049 (0.0078)</td>
<td></td>
</tr>
<tr>
<td>Largest</td>
<td>$\gamma_{50}$</td>
<td>0.1785 (0.0495)</td>
<td>0.1105 (0.0454)</td>
<td>0.1014 (0.0415)</td>
<td>0.1241 (0.0381)</td>
<td>0.3680 (0.0244)</td>
</tr>
<tr>
<td></td>
<td>$\gamma_{51}$</td>
<td>0.4317 (0.0185)</td>
<td>0.1327 (0.0158)</td>
<td>0.0621 (0.0143)</td>
<td>0.0790 (0.0129)</td>
<td>0.1135 (0.0082)</td>
</tr>
</tbody>
</table>

Panel B: ML estimates for BEKK model

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Coefficient</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$C_0$</td>
<td>1.2504 (0.0541)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>0.4439 (0.0549)</td>
<td>0.1425 (0.0955)</td>
<td>0.0293 (0.0855)</td>
<td>-0.0872 (0.0794)</td>
<td>-0.0725 (0.0602)</td>
</tr>
<tr>
<td>2nd</td>
<td>$C_0$</td>
<td>0.9440 (0.0591)</td>
<td>0.7639 (0.0117)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>-0.0825 (0.0506)</td>
<td>0.5820 (0.0902)</td>
<td>0.0905 (0.0819)</td>
<td>-0.0998 (0.0772)</td>
<td>-0.0471 (0.0571)</td>
</tr>
<tr>
<td>3rd</td>
<td>$C_0$</td>
<td>0.7786 (0.0602)</td>
<td>0.5318 (0.0161)</td>
<td>0.6181 (0.0101)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>-0.1304 (0.0499)</td>
<td>0.1394 (0.0839)</td>
<td>0.5778 (0.0804)</td>
<td>-0.0552 (0.0751)</td>
<td>-0.0711 (0.0545)</td>
</tr>
<tr>
<td>4th</td>
<td>$C_0$</td>
<td>0.8186 (0.0595)</td>
<td>0.4996 (0.0183)</td>
<td>0.3018 (0.0167)</td>
<td>0.6694 (0.0110)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>-0.2081 (0.0494)</td>
<td>0.1609 (0.0869)</td>
<td>0.1043 (0.0786)</td>
<td>0.3275 (0.0800)</td>
<td>0.0511 (0.0563)</td>
</tr>
<tr>
<td>Largest</td>
<td>$C_0$</td>
<td>0.7647 (0.0646)</td>
<td>0.4809 (0.0208)</td>
<td>0.2492 (0.0196)</td>
<td>0.3414 (0.0176)</td>
<td>0.7326 (0.0132)</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>-0.1720 (0.0513)</td>
<td>0.1926 (0.0916)</td>
<td>-0.1062 (0.0813)</td>
<td>0.0733 (0.0838)</td>
<td>0.4671 (0.0626)</td>
</tr>
</tbody>
</table>
We present ML estimates for the linear information instrument model and the BEKK model for the simulated weekly return sample of 2,419 observations. For the instrument model, the conditional covariance matrix $M_t$ is computed as $M_t = L_t L_t'$, where $L_t = [l_{ijt}]$ is the lower triangular matrix at $t$, and $l_{ijt} = \gamma_{ij0} + \gamma_{ij1} Z_{lt-1}$, where $\gamma_{ij0}$ and $\gamma_{ij1}$ are parameters and $Z_{lt-1}$ is the instrument variable realized at $t-1$.

For the BEKK model, the conditional covariance matrix $M_{BEKK}$ is modeled as $M_{BEKK} = C_0 C_0' + A_1 Z_{t-1} Z_{t-1}' A_1'$, where $C_0$ and $A_1$ are parameter matrices with $C_0$ being lower triangular and $Z_{t-1}$ is the instrument vector realized at $t-1$. In Panel A and B, we report the ML estimates for the linear information instrument model and the BEKK model, respectively. In both models, the information instrument $Z_{lt-1}$ is computed as the previous four-week mean of absolute returns beginning one week prior. The standard errors of the ML estimates are obtained as the square root of the diagonal elements of the inverse of the Fisher information matrix generated by BFGS algorithm and are reported in parentheses.
We present the elementwise estimates of the conditional covariance matrix as a function of coefficients and information instruments for the simulated weekly return sample of 2,419 observations. The conditional covariance matrix is estimated by the linear information instrument model defined as

$$ \Phi_{ij} = \Phi_{i0} + \Phi_{i1} Z_{t-1}, $$

where $\Phi_{ij}$ is the $i,j$th element in the lower triangular matrix $\Phi$. The covariance matrix is then constructed as $\Phi = \Phi \Phi'$. The coefficients with respect to information instruments in $\Phi_{ij}$ are functions of parameters $\Phi_{i0}$ and $\Phi_{i1}$, and are estimated using ML. We report the ML estimates of the coefficients for each element in the conditional covariance matrix with delta method standard errors given in parentheses.
### Table 5
Conditional covariance matrix estimates for percentiles of instruments.

<table>
<thead>
<tr>
<th>BM Portfolios</th>
<th>Instrument Percentile</th>
<th>Smallest</th>
<th>2\text{nd}</th>
<th>3\text{rd}</th>
<th>4\text{th}</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5%</td>
<td>1.1419</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>2.4894</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Smallest</td>
<td>50%</td>
<td>3.6453</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75%</td>
<td>6.0198</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>97.5%</td>
<td>20.7994</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2\text{nd}</td>
<td>2.5%</td>
<td>0.8031</td>
<td>0.8410</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>1.9298</td>
<td>1.9033</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>2.8984</td>
<td>2.8067</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75%</td>
<td>4.8816</td>
<td>4.6331</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>97.5%</td>
<td>17.1545</td>
<td>15.6970</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3\text{rd}</td>
<td>2.5%</td>
<td>0.7778</td>
<td>0.6958</td>
<td>0.8318</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>1.7364</td>
<td>1.6031</td>
<td>1.6711</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>2.6367</td>
<td>2.4456</td>
<td>2.5122</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75%</td>
<td>4.4697</td>
<td>4.1454</td>
<td>4.1975</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>97.5%</td>
<td>17.5169</td>
<td>16.0424</td>
<td>17.5955</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4\text{th}</td>
<td>2.5%</td>
<td>0.8703</td>
<td>0.7516</td>
<td>0.7609</td>
<td>1.0089</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>1.6970</td>
<td>1.5470</td>
<td>1.4643</td>
<td>1.6531</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>2.4943</td>
<td>2.3013</td>
<td>2.1999</td>
<td>2.3632</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>75%</td>
<td>4.0868</td>
<td>3.7926</td>
<td>3.6511</td>
<td>3.7305</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>97.5%</td>
<td>16.2800</td>
<td>15.0347</td>
<td>16.1531</td>
<td>16.3071</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>2.5%</td>
<td>0.8873</td>
<td>0.7875</td>
<td>0.7843</td>
<td>0.9156</td>
<td>1.1745</td>
</tr>
<tr>
<td></td>
<td>25%</td>
<td>1.7790</td>
<td>1.6358</td>
<td>1.5326</td>
<td>1.5814</td>
<td>1.9596</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>2.6030</td>
<td>2.4061</td>
<td>2.2802</td>
<td>2.2807</td>
<td>2.7360</td>
</tr>
<tr>
<td></td>
<td>75%</td>
<td>4.3354</td>
<td>4.0075</td>
<td>3.8264</td>
<td>3.7002</td>
<td>4.3791</td>
</tr>
<tr>
<td></td>
<td>97.5%</td>
<td>18.6431</td>
<td>17.0092</td>
<td>18.1097</td>
<td>17.7502</td>
<td>21.5263</td>
</tr>
</tbody>
</table>

We present the conditional covariance matrix estimates for the 2.5, 25, 50, 75, and 97.5 percentiles of the information instruments for the simulated weekly return sample of 2419 observations. The percentile values are computed for each information instrument across the sample period. The conditional covariance matrix estimates are computed using the coefficients described in Table 5.
Table 6
Comparing the linear information instrument and BEKK covariance models.

Panel A: MAE

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Inst. Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>1.3668</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>1.1879</td>
<td>1.1609</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>1.1484</td>
<td>1.0912</td>
<td>1.1363</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>1.1414</td>
<td>1.0761</td>
<td>1.0823</td>
<td>1.1594</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>1.1581</td>
<td>1.0987</td>
<td>1.0944</td>
<td>1.1354</td>
<td>1.2387</td>
</tr>
<tr>
<td>BEKK Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>1.6043***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>1.4244***</td>
<td>1.4897***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>1.3348***</td>
<td>1.3596***</td>
<td>1.3893***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>1.3286***</td>
<td>1.3522***</td>
<td>1.3294***</td>
<td>1.5001***</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>1.2870***</td>
<td>1.2954***</td>
<td>1.2602***</td>
<td>1.3396***</td>
<td>1.4767***</td>
</tr>
</tbody>
</table>

Panel B: RMSE

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Inst. Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>3.6467</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>3.5979</td>
<td>3.6596</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>3.5320</td>
<td>3.5200</td>
<td>3.5711</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>3.6375</td>
<td>3.6580</td>
<td>3.6731</td>
<td>4.0809</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>3.4627</td>
<td>3.4807</td>
<td>3.5114</td>
<td>3.7803</td>
<td>3.6410</td>
</tr>
<tr>
<td>BEKK Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>4.2026***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>4.1305***</td>
<td>4.3418***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>3.9549***</td>
<td>4.0848***</td>
<td>4.1370***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>4.0189***</td>
<td>4.2099***</td>
<td>4.1916***</td>
<td>4.7774***</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>3.7557***</td>
<td>3.8890***</td>
<td>3.8481***</td>
<td>4.1823***</td>
<td>4.0586***</td>
</tr>
</tbody>
</table>

Panel C: HMSE

<table>
<thead>
<tr>
<th>Quintile BM Portfolios</th>
<th>Smallest</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Largest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Inst. Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>0.1268</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>0.1921</td>
<td>0.1423</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>0.2405</td>
<td>0.2564</td>
<td>0.1157</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>2.0406</td>
<td>0.2282</td>
<td>0.1463</td>
<td>0.1176</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>0.2454</td>
<td>0.2475</td>
<td>0.1554</td>
<td>0.1398</td>
<td>0.1173</td>
</tr>
<tr>
<td>BEKK Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smallest</td>
<td>0.3495***</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2nd</td>
<td>0.6852***</td>
<td>0.6749***</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3rd</td>
<td>0.5825**</td>
<td>1.3193</td>
<td>0.3928***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4th</td>
<td>1.6527</td>
<td>0.7970***</td>
<td>0.4243***</td>
<td>0.3664***</td>
<td>-</td>
</tr>
<tr>
<td>Largest</td>
<td>0.3900***</td>
<td>0.7845*</td>
<td>0.3654**</td>
<td>0.3249***</td>
<td>0.2788***</td>
</tr>
</tbody>
</table>
We compare the estimation accuracy of the conditional covariance matrix between the linear information instrument model and the BEKK model for the simulated weekly return sample of 2,419 observations. In Panels A, B, and C, we report the mean absolute error (MAE), the root mean square error (RMSE), and the heteroskedasticity-adjusted MSE (HMSE), respectively, for each element in the covariance matrix, where

\[
MAE_{ij} = \frac{1}{T} \sum_{t=1}^{T} |h_{ij,t} - \sigma_{ij,t}|
\]

\[
RMSE_{ij} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (h_{ij,t} - \sigma_{ij,t})^2}
\]

\[
HMSE_{ij} = \frac{1}{T} \sum_{t=1}^{T} \left( \frac{h_{ij,t}}{\sigma_{ij,t}} - 1 \right)^2
\]

for \(h_{ij,t}\) = the \(ij^{th}\) element of the estimated covariance matrix with related population value, \(\sigma_{ij,t}\).

Significance levels for the Diebold-Mariano statistic examine each estimate relative to the most accurate estimate at the one, five, and ten percent levels with significance denoted by ***, **, and *, respectively. We report significance levels for the larger estimate in each panel.
Table 7
Conditional covariance matrix simulation results.

<table>
<thead>
<tr>
<th>Panel A: T=260, N=3</th>
<th>Model</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>Linear Inst.</td>
<td>0.2782</td>
<td>0.0586</td>
<td>0.1898</td>
<td>0.6091</td>
<td>0.2044</td>
<td>0.2678</td>
<td>0.3905</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.3713</td>
<td>0.1400</td>
<td>0.1977</td>
<td>1.4011</td>
<td>0.2327</td>
<td>0.3323</td>
<td>0.6428</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Linear Inst.</td>
<td>0.3460</td>
<td>0.0717</td>
<td>0.2356</td>
<td>0.7190</td>
<td>0.2554</td>
<td>0.3329</td>
<td>0.4811</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.4736</td>
<td>0.1923</td>
<td>0.2419</td>
<td>1.8756</td>
<td>0.2901</td>
<td>0.4178</td>
<td>0.8569</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: T=260, N=5</th>
<th>Model</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>Linear Inst.</td>
<td>1.0589</td>
<td>0.8682</td>
<td>0.3743</td>
<td>6.7154</td>
<td>0.4187</td>
<td>0.7333</td>
<td>2.8035</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>1.1311</td>
<td>0.9129</td>
<td>0.3783</td>
<td>7.9732</td>
<td>0.4566</td>
<td>0.7880</td>
<td>2.9408</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Linear Inst.</td>
<td>1.3477</td>
<td>1.0663</td>
<td>0.4555</td>
<td>7.8145</td>
<td>0.5039</td>
<td>0.9693</td>
<td>3.4936</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>1.7120</td>
<td>1.6164</td>
<td>0.4665</td>
<td>15.7998</td>
<td>0.5642</td>
<td>1.1670</td>
<td>4.3742</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: T=1,040, N=3</th>
<th>Model</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>Linear Inst.</td>
<td>0.2678</td>
<td>0.0287</td>
<td>0.2060</td>
<td>0.3846</td>
<td>0.2267</td>
<td>0.2646</td>
<td>0.3185</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.3484</td>
<td>0.1370</td>
<td>0.2210</td>
<td>1.0037</td>
<td>0.2413</td>
<td>0.2993</td>
<td>0.7051</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Linear Inst.</td>
<td>0.3384</td>
<td>0.0355</td>
<td>0.2599</td>
<td>0.4736</td>
<td>0.2859</td>
<td>0.3357</td>
<td>0.4002</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.4576</td>
<td>0.2059</td>
<td>0.2775</td>
<td>1.3794</td>
<td>0.3041</td>
<td>0.3812</td>
<td>1.0014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: T=1,040, N=5</th>
<th>Model</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>Linear Inst.</td>
<td>0.2760</td>
<td>0.0293</td>
<td>0.2180</td>
<td>0.3956</td>
<td>0.2357</td>
<td>0.2720</td>
<td>0.3303</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.3626</td>
<td>0.1163</td>
<td>0.2254</td>
<td>0.8868</td>
<td>0.2559</td>
<td>0.3230</td>
<td>0.6165</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Linear Inst.</td>
<td>0.3531</td>
<td>0.0375</td>
<td>0.2770</td>
<td>0.4999</td>
<td>0.3007</td>
<td>0.3477</td>
<td>0.4219</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.4802</td>
<td>0.1745</td>
<td>0.2839</td>
<td>1.3005</td>
<td>0.3258</td>
<td>0.4161</td>
<td>0.8796</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel E: T=2,419, N=3</th>
<th>Model</th>
<th>Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>Linear Inst.</td>
<td>0.2764</td>
<td>0.0151</td>
<td>0.2417</td>
<td>0.3320</td>
<td>0.2541</td>
<td>0.2753</td>
<td>0.3037</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.3296</td>
<td>0.0285</td>
<td>0.2662</td>
<td>0.4325</td>
<td>0.2880</td>
<td>0.3277</td>
<td>0.3816</td>
</tr>
<tr>
<td>RMSPE</td>
<td>Linear Inst.</td>
<td>0.3503</td>
<td>0.0192</td>
<td>0.3034</td>
<td>0.4202</td>
<td>0.3209</td>
<td>0.3490</td>
<td>0.3847</td>
</tr>
<tr>
<td></td>
<td>BEKK</td>
<td>0.4419</td>
<td>0.0461</td>
<td>0.3372</td>
<td>0.6173</td>
<td>0.3762</td>
<td>0.4379</td>
<td>0.5221</td>
</tr>
</tbody>
</table>
Panel F: T=2,419, N=5

<table>
<thead>
<tr>
<th>Model</th>
<th>MAPE Mean</th>
<th>Std</th>
<th>Min</th>
<th>Max</th>
<th>5th Pctl.</th>
<th>Median</th>
<th>95th Pctl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Inst.</td>
<td>0.2901</td>
<td>0.0178</td>
<td>0.2495</td>
<td>0.3616</td>
<td>0.2650</td>
<td>0.2886</td>
<td>0.3235</td>
</tr>
<tr>
<td>BEKK</td>
<td>0.3945</td>
<td>0.0442</td>
<td>0.2812</td>
<td>0.6773</td>
<td>0.3255</td>
<td>0.3925</td>
<td>0.4707</td>
</tr>
<tr>
<td>RMSPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Inst.</td>
<td>0.3743</td>
<td>0.0239</td>
<td>0.3183</td>
<td>0.4658</td>
<td>0.3400</td>
<td>0.3729</td>
<td>0.4177</td>
</tr>
<tr>
<td>BEKK</td>
<td>0.5516</td>
<td>0.0744</td>
<td>0.3565</td>
<td>1.0505</td>
<td>0.4350</td>
<td>0.5477</td>
<td>0.6765</td>
</tr>
</tbody>
</table>

We report the mean absolute percentage error (MAPE) and the root mean square percentage error (RMSPE) for 1,000 simulation replications for both the linear information instrument model and the BEKK model. For each replication, we compute the MAPE and RMSPE as,

\[
MAPE = \frac{1}{T} \sum_{t=1}^{T} \frac{e| |M_t - \Sigma_t| |e'}{e| |\Sigma_t| |e'}
\]

\[
RMSPE = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{e(M_t - \Sigma_t) (M_t - \Sigma_t)e'}{e\Sigma_t (M_t - \Sigma_t)e'}
\]

where \( e \) is a vector of ones, \( M_t \) and \( \Sigma_t \) are the estimate and the population value of the conditional covariance matrix at \( t \), respectively, and \( T \) is the number of estimated conditional covariance matrices with dimension \( N \). We consider sample sizes for \( T = 260, 1040, \) and 2,419 and \( N = 3 \) and 5. In panels A through F, we report the descriptive statistics of MAPE and RMSPE for different \( T \) and \( N \), and different models across 1,000 simulation replications.
Does mutual fund size erode performance?
Conditional evidence using fund flows

1. Introduction

The mutual fund industry has witnessed tremendous growth. According to the Investment Company Institute, the combined assets in US mutual funds has doubled from approximately $11.9 trillion in 2000 to approximately $25 trillion in 2010, corresponding to an annual growth rate of assets under management (AUM) of approximately 10 percent. Within the US, demographic changes have resulted in a great number of investors that rely on mutual funds as investment and retirement vehicles. For instance, more than half of all households invest their financial assets in mutual funds (Investment Company Institute, 2011).

The literature examining mutual fund performance is extensive. Seminal works examining performance include Sharpe (1966), Treynor (1965), Jensen (1968), Merton (1981), and Henriksson and Merton (1981). More recent work examines performance in the context of conditioning information. Representative work in this area includes Ferson and Schadt (1996), Christopherson et al. (1998), Ferson and Siegel (2001), Jha et al. (2009), and Tang and Whitelaw (2011). Our paper builds upon this literature. We model asset return moments as functions of underlying economic information instruments, and then specify performance metrics with these conditional moments. The analysis is similar in spirit to Ferson and Schadt (1996) and Jha et al. (2009) with the additional assumption that both the mean vector and covariance matrix both evolve with underlying economic information instruments.

Although there is debate regarding the predictability of first moments of asset returns (for two good competing articles discussing return predictability see Welch and Goyal (2008) and Campbell and Thompson (2008)), there is little argument that variances and covariances display strong persistence over time. For example, the extensive literature ARCH (Engle 1982), GARCH (Bollerslev 1986) and related recent literature using techniques such as the Dynamic Conditional Correlation (Engle 2002) approach suggest procedures for modeling first and second moments of asset returns have the potential to add value to investor’s portfolios.
Specific examples modeling multivariate asset return moments with links to the conditional investment opportunity set include Korkie and Turtle (1998), Avramov and Wermers (2006), and Tang and Whitelaw (2011), amongst others. We adopt a new specification in which both means and covariances evolve with a known information instrument – mutual fund flows. Within our model, we examine the risk-adjusted performance of mutual funds for a range of funds with different levels of AUM.

The correlation between fund flow and performance has been well documented in literature. For instance, Gruber (1996) and Zheng (1999) both that funds with higher past flows have superior performance relative to funds with smaller past flows. Hendricks, Patel, and Zeckhauser (1993) document that funds tend to flow towards successful prior performers. Chevalier and Ellison (1997), Sirri and Tufano (1998), and Del Guercio and Tkac (2000) show that the relationship between fund flow and prior performance is asymmetric in that poor performers are not as heavily penalized with fund outflows.

A related literature examining the link between mutual fund size and performance has also received substantial attention in literature. Berk and Green (2004) develop a rational model of active portfolio management. In their model, fund managers with more skills manage more assets. However, as the asset base increases, with fund inflows, managerial talent is diluted so the more-skilled managers earn the same expected excess return as their less-skilled peers. As a result, according to Berk and Green (2004), it is rational to expect that fund performance will decrease with fund size. The empirical evidence on fund size and performance is mixed. Some research suggests that a larger asset base erodes fund performance. For example, Perold and Salomon (1991) show that as a fund grows larger, fund managers have to trade larger volume of stock, resulting in higher transaction costs. Chen et al. (2004) and Yan (2008) find evidence that fund size and performance are negatively correlated, and this relationship is more pronounced among funds that lack liquidity. Aggarwal and Jorion (2010) report that managers in small and emerging hedge funds often show strong investment performance. Grinblatt and Titman (1989) provide empirical evidence that fund returns decrease with fund size, but this effect disappears for fund returns net of expenses. Other researchers, however, argue that growth in fund size provides resources for research and lowers expense ratios, thereby increasing fund performance. For example, Tufano and Sevick (1997) find an inverse relationship between
fund size and fees, and Christoffersen and Sarkissaian (2009) show that larger funds based in financial centers perform significantly better than other funds. In sum, there is little consensus among academics regarding the relationship between fund size and performance.

To better understand the time-varying relationship between fund size and performance, we implement a new conditional alpha model in which fund flow is an information instrument for both conditional means and covariances. We then relate changes in means and covariance to the underlying conditional regression alpha in a manner that is consistent with the underlying conditional regression theory.

As an important by-product of our analysis, we develop a bootstrapping approach to determine the significance of our conditional alphas within the underlying conditional regression framework. To our knowledge, this is a novel and important innovation that will facilitate conditional inferences regarding manager performance.

Our conditional covariance specification follows the linear information instrument model of Turtle and Wang (2011). The model is based on a restricted Cholesky-like decomposition and ensures positive definiteness and invariance to variate order. This approach is beneficial in our fund performance setting in that covariances are driven by an underlying economic information instrument. This allows us to link both means and (co)variances to changes in underlying fund flows. Turtle and Wang (2011) demonstrate that the model has fewer parameters and smaller mean absolute percentage error (MAPE) and root mean square percentage error (RMSPE) than a comparably specified BEKK model. Wald and likelihood ratio tests provide strong evidence that conditional covariances are not constant.

Our primary interest in this paper is to empirically examine the risk-adjusted performance of mutual funds of different sizes, as measured by AUM. We find that small funds significantly outperform large funds after management and administrative fees, and this outperformance persists over the sample period from 1991 to 2011. In addition, we find a positive response of conditional alphas to fund flow for small and mid funds, and a negative response for large funds. Finally, we find that comparing to other homoscedastic

---

23 The invariance to variate order property ensures that the functional form for each lower triangular matrix element does not change when the order of the variates changes.
conditional models, our new conditional alpha model produces similar magnitude of conditional alphas but with different levels of significance.

The reminder of this paper is organized as follows. Section 2 presents the model. Section 3 presents the empirical analysis and findings. Concluding remarks are given in Section 4.

2. Model

Our goal is to provide a conditional alpha performance measure that evolves with economic information instruments. Economic information instruments may be either lagged endogenous variables, or any other known observable economic information instruments from the information set. For simplicity, we define the economic information instruments as own information instruments, $Z_{t-1}, i = 1, 2, ... , N$ that are specific to each asset considered. Examples of own information instruments may include key financial characteristics such as liquidity measures, profitability measures, and valuation measures; or any other variable that is specific to each asset considered. We assume that the assets’ expected returns and covariances are dependent on these information instruments.

To construct conditional alphas, we consider an efficient mean-variance investment opportunity set, where there are $N_2$ individual assets and $N_1$ spanning portfolios. A general expression for the conditional alphas based on some known information instrument $Z_{t-1}$ is,

$$\alpha_{2|z_{t-1}} = \mu_{2|z_{t-1}} - \beta_{2|z_{t-1}} \mu_{1|z_{t-1}}$$ (1)

where $\mu_{2|z_{t-1}}$ and $\mu_{1|z_{t-1}}$ are $N_2 \times 1$ and $N_1 \times 1$ conditional means of excess returns; and $\beta_{2|z_{t-1}}$ is an $N_2 \times N_1$ matrix of conditional betas, given by,

$$\beta_{2|z_{t-1}} = \Sigma_{21|z_{t-1}} \Sigma_{11|z_{t-1}}^{-1}$$ (2)

where $\Sigma_{21|z_{t-1}}$ and $\Sigma_{11|z_{t-1}}$ are $N_2 \times N_1$ and $N_1 \times N_1$ conditional covariance matrices, respectively.

24 Our model may be readily extended to incorporate macroeconomic information instruments that impact all assets in the opportunity set.
Suppose the information instruments for conditional means and conditional covariances are given by $Z_{t-1}$ and $Z_{\sigma,t-1}$, respectively. If we assume that the conditional mean for each asset is a linear function of its own information instrument, it follows that,

$$
\begin{cases}
\mu_1|Z_{t-1} = a_{10} + a_{11} \otimes Z_{1,t-1} \\
\mu_2|Z_{t-1} = a_{20} + a_{21} \otimes Z_{2,t-1}
\end{cases}
$$

(3)

where $a_{10}, a_{11}$ and $a_{20}, a_{21}$ are $N_1 \times 1$ and $N_2 \times 1$ coefficient vectors; $Z_{1,t-1}$ and $Z_{2,t-1}$ are $N_1 \times 1$ and $N_2 \times 1$ information instruments for conditional means; and $\otimes$ denotes the Hadamard matrix product operator (i.e., elementwise multiplication).

Substituting (2) and (3) in (1) results in the following expression for the conditional alphas,

$$
\alpha_{2}|Z_{t-1} = a_{20} + a_{21} \otimes Z_{2,t-1} - (a_{10} + a_{11} \otimes Z_{1,t-1}) \Sigma_{21} \sigma_{\sigma,t-1}^{-1} \Sigma_{11}^{-1}
$$

(4)

In the simplest case of only one asset and one factor portfolio (i.e., $N_2 = 1$ and $N_1 = 1$), equation (4) becomes,

$$
\alpha_{2}|Z_{t-1} = a_{20} + a_{21} Z_{2,t-1} - (a_{10} + a_{11} Z_{1,t-1}) \sigma_{21} \sigma_{\sigma,t-1}^{-1}
$$

(5)

Following Turtle and Wang (2011), the conditional covariance matrix, $\Sigma_{|Z_{\sigma,t-1}}$, is given by,

$$
\Sigma_{|Z_{\sigma,t-1}} = L_{|Z_{\sigma,t-1}} L_{|Z_{\sigma,t-1}}'
$$

(6)

where $L_{|Z_{t-1}}$ is a lower triangular matrix; and each element in $L_{|Z_{\sigma,t-1}} = [l_{ij}|Z_{\sigma,t-1}]$ is a linear function of the known information instrument for conditional covariances, $Z_{\sigma,t-1} = [Z_{1,\sigma,t-1}, Z_{2,\sigma,t-1}]'$.

Consider the following functional form for $l_{ij}|Z_{\sigma,t-1}$,

$$
l_{ij}|Z_{\sigma,t-1} = b_{ij0} + b_{ij1} Z_{i,\sigma,t-1}
$$

(7)

for $i = 1$ and $2$; and model parameters $b_{ij0}$ and $b_{ij1}$.

Substituting (6) and (7) in (5), we get the following expression for the conditional alpha,

$$
\alpha_{2}|Z_{t-1} = a_{20} + a_{21} Z_{2,t-1} - (a_{10} + a_{11} Z_{1,t-1}) \frac{b_{210} + b_{211} Z_{2,\sigma,t-1}}{b_{110} + b_{111} Z_{1,\sigma,t-1}}
$$

(8)
The development details are given in Appendix 1. The conditional alpha may be evaluated at specific instrument values for any given period.25

To estimate equation (8), consider the following system of equations,

\[
\begin{align*}
    r_1 &= \mu_1|z_{t-1} + e_1|z_{t-1} \\
    r_2 &= \mu_2|z_{t-1} + e_2|z_{t-1}
\end{align*}
\]

(9)

where \( r_1 \) and \( r_2 \) are the excess return for the factor portfolio and the individual asset, respectively; and the covariance matrix of the model disturbances is given by

\[
\Sigma|z_{\sigma,t-1} = \begin{bmatrix}
\sigma_{11}|z_{\sigma,t-1} & \sigma_{12}|z_{\sigma,t-1} \\
\sigma_{21}|z_{\sigma,t-1} & \sigma_{22}|z_{\sigma,t-1}
\end{bmatrix}
\]

(8)

If we assume that the model disturbances follow a multivariate normal distribution, we may construct the log likelihood function as,

\[
\log(\theta|r_{z_{t-1}}) = -\frac{1}{2} \sum_{t=1}^{N} \left[ N \log(2\pi) + \log|\Sigma|z_{\sigma,t-1}| + e_{t}|z_{t-1} \right]^T \Sigma^{-1}|z_{\sigma,t-1}^{-1} e_{t}|z_{t-1} \right]
\]

(10)

where \( e = \begin{bmatrix} e_1|z_{t-1} \\ e_2|z_{t-1} \end{bmatrix} = \begin{bmatrix} r_1 - \mu_1|z_{t-1} \\ r_2 - \mu_2|z_{t-1} \end{bmatrix} \).

The log likelihood function can be used to obtain parameter estimates using ML. Under standard regularity conditions, parameter estimates converge to population values. According to the functional invariance property of ML estimation, the resultant conditional covariance matrices are also ML and therefore are consistent estimators of the true conditional covariances.

---

Equation (8) may be rewritten to show the relationship between the conditional alphas and the traditional unconditional alphas. In particular, the conditional alpha is given as,

\[
\alpha_{2|z_{t-1}} = \alpha_2 + \left( \frac{b_{210} + b_{211}\mu_{z_1}}{b_{110} + b_{111}\mu_{z_1}} \right) \left( \frac{b_{210} + b_{211}\mu_{z_1}}{b_{110} + b_{111}\mu_{z_1}} \right) + b_{111}^2 \sigma_{z_1,z_2} \]

\[
+ \left( Z_{z_{t-1}} - \mu_{z_2} \right) \left( b_{110} + b_{111}Z_{1_{t-1}} \right) - a_{11} \left( Z_{1_{t-1}} - \mu_{z_1} \right) \left( b_{210} + b_{211}Z_{2_{t-1}} \right)
\]

where \( \alpha_2 \) is the unconditional alpha of the individual asset; \( \mu_1 \) is the unconditional mean of the spanning portfolio returns; \( \mu_{z_1} \) and \( \mu_{z_2} \) are the unconditional mean of the information variable for the individual asset and the spanning portfolio, respectively; \( \Sigma_{z_1,z_1} \) is the unconditional variance of the spanning portfolio; and \( \Sigma_{z_1,z_2} \) is the unconditional covariance of the individual asset and the spanning portfolio. This result makes use of the double expectation theorem.
Jha et al. (2009) develop a conditional alpha model that admits changes in conditional means with the economic fundamentals. They show that given an independent identically distributed disturbance vector with homoscedastic covariance structure, equations (1) and (2) result in a conditional alpha given by,

\[
\alpha_t = \alpha_{j0} + \alpha_{j1} z_{jt-1}.
\]  

We denote this alpha as the homoscedastic conditional alpha and we estimate this alpha as the special homoscedastic case of the more general specification in equation (1).

3. Empirical Analysis

We present empirical results for the risk-adjusted performance of size portfolios using both the unconditional and conditional models. For conditional models, we use the previous month’s scaled fund flow as the information instrument. We compare our model with popular unconditional models as well as with the homoscedastic model. To examine the significance of conditional alphas, we develop a conditional bootstrap procedure for both our model and the homoscedastic special case.

3.1 Data

Our mutual fund data are from the Morningstar Direct database, which contains survivorship-bias free data on historical mutual fund characteristics, fund flow, and performance. We limit our analysis to US domestic equity funds that are not closed to all investments, or to new investments. We classify mutual funds into domestic funds if the percentage of holdings in US stocks is equal to or greater than 80 percent. Following Del Guercio and Tkac (2002), we require each fund in our sample to have an initial minimum purchase amount of less than $20,000. When a fund offers multiple share classes, we retain only the largest fund class as measured by net assets at the end of December 2011. Our final sample includes 2,036 distinct domestic equity mutual funds and 280,894 fund-month observations over the years 1991 to 2011. The number of funds in each month varies from 289 in January 1991 to 1988 in November 2011. We extract monthly total returns for all funds in the sample period. Although the total returns from Morningstar do not adjust for sales charges such as loads and redemption fees, they do account for
management, administrative, 12b-1 fees and other costs taken out of fund assets. Monthly excess returns are calculated by subtracting the CRSP-Sift riskless rate for one month from the total returns. We obtain fund characteristics including fund size, fund flow, and fund portfolio market capitalization for each mutual fund and for each month. In particular, fund size at month $t$ is measured by the total net assets across classes at month $t$, whereas scaled fund flow at month $t$ is computed as dividing total net flow across classes at month $t$ by fund size at month $t-1$. The portfolio market capitalization at month $t$ is computed as the geometric mean of the market capitalization for all of the stocks the fund owns at month $t$. To reduce the impact of outliers on our results, all extreme values of total net assets and total net flow are trimmed at the 99% level.

We form size portfolios at month $t$ based on each fund’s size at the end of month $t-1$. For each month, we create five size portfolios. We update these portfolios at each month over the period from 1991 to 2011 and construct both equally and value weighted portfolio returns and portfolio fund flow measures.

3.2 Descriptive statistics

Table 1 reports summary statistics for our sample. We report the means, standard deviations, and common quantile values for the variables of interest for all funds, and for funds in the smallest, mid, and largest size quintiles. In Panel A, we note that the average total net asset of $918.65 million is dramatically larger than the median total net assets of $213.42 million, suggesting a familiar right skewness in the original data. Similar results are found for total net flow and scaled fund flow. The expense ratio yields a sample mean and median of 1.17% per annum. Panels B, C, and D report the analogous numbers for funds in the smallest, mid, and largest size quintiles, respectively. We find that the average total net assets of the large fund sample is about $3.6 billion larger than the average small fund. In contrast, the difference of total net assets between the average mid and small fund sample is approximately $210 million. This result is consistent with a highly right-skewed distribution in total net assets across the sample of funds. In addition, we notice that the total average net flow increases with fund size, whereas the scaled fund flow decreases with fund size. Intuitively, larger funds are more attractive to institutional investors, which account for most of the market’s trading activities. So it is not surprising to see that on average total net flow increases with fund size. To adjust fund
flow for size, we use scaled fund flow as our information instrument. Moreover, we find that the small fund sample has an average expense ratio of 1.49% per annum, which is 0.60% higher than its larger size counterpart. This result is consistent with Tufano and Sevick (1997) that fund fees decrease with fund size. Last but not least, we find that the average portfolio market capitalization increases with fund size, suggesting that smaller funds tend to invest in smaller stocks.

In the following empirical analysis, we provide results for the risk-adjusted performance of the small, mid, and large fund samples. We use a portfolio approach to study both the unconditional and conditional fund performance. In particular, we form five size portfolios based on fund’s total net assets at the end of each month and update them monthly over the entire sample period. The portfolios are formed either equally or value weighted. All fund returns are net of management and administrative fees.

*** Insert Table 1 about here ***

3.3 Unconditional performance analysis

We use four unconditional models to analyze the performance of the small, mid, and large size portfolios. The first model is the market-adjusted fund returns calculated as the difference between total fund returns and market returns. The market returns are obtained either from a value-weighted CRSP index or from a value-weighted portfolio including all funds in the sample.26 The second model is the unconditional Jensen’s alpha from the Capital Asset Pricing Model (CAPM), which is estimable from an unconditional regression with the market excess return as the sole risk factor. The third model is the Fama-French (FF) three factor model (Fama and French (1993)) and is estimable from an unconditional regression with the three risk factors as regressors. The fourth model is the Carhart four factor model (Carhart (1997)) and is estimable from an unconditional regression with the FF three factors and the momentum factor.

Table 2 presents the unconditional results for the small, mid, and large size portfolios. We find that for both equal- and value-weighted portfolio formation strategies, the CRSP index benchmark yields smaller and insignificant market-adjusted returns, and larger but less significant unconditional alphas, than the sample

---

26 Because our primary interest is in the correlation between fund size and performance, our choice of the market return has little effect on the results.
fund index. This result suggests that on average, our fund portfolios outperform the CRSP index to a larger extent than the sample fund index. When using the sample fund index as market benchmark, all four unconditional measures produce similar results. In particular, small funds outperform larger funds by approximately 0.23% per month, or 2.80% per annum, for equal-weighted portfolios, and 0.25% per month, or 2.92% per annum, for value-weighted portfolios, after management and administrative fees. The $t$-test of the difference in unconditional measures between the small and large size portfolio suggests that small funds perform significantly better than large funds. The outperformance persists even after we control for additional systematic risk factors such as size, book to market, and momentum. In addition, we observe that the equal-weighted portfolios provide better unconditional performance than value-weighted portfolios. For instance, using the sample fund index as market benchmark, the unconditional performance of equal-weighted large portfolio is indifferent from zero across all models, whereas the performance of value-weighted large portfolio is all significantly negative at 5% level. This result occurs because of the extra weight given to smaller funds in an equal-weighted portfolio. In sum, the numbers in Table 2 confirm some of the findings in the extant literature that fund performance decreases with fund size. In the next section, we examine if these unconditional results persist in a conditional setting.

3.4 Conditional performance analysis

One of the benefits of conditional performance evaluation is that we are able to identify risk-adjusted performance on a period by period basis. In addition to our general case, we also present a homoscedastic special case for the conditional alpha. By construction, both models consider time-varying changes in asset means; however, the more general heteroscedastic model also has the potential to fit time-varying covariances that may induce nonlinear relationships between asset returns and the underlying information instrument. The information instrument for conditional means and conditional covariances is given by $Z_{t-1} = [Z_{t,t-1}]$ and $Z_{\sigma,t-1} = [Z_{\sigma,t,t-1}]$, respectively. In particular, $Z_{t,t-1}$ is computed as the total net flow at month $t-1$. **Insert Table 2 about here**

55
divided by the total net assets at month $t-2$; and $Z_{t-1}$ is obtained as the absolute value of $Z_{t-1}$. Turtle and Wang (2011) find that an absolute value of their information instrument behaves better than the related raw information instrument using standard measures of model performance. Because we do not have the fund flow data for the CRSP index, in the following conditional performance analysis, we compute the market returns as the total returns for the value-weighted portfolio including all funds in our sample.

To examine how the general heteroscedastic model is different from the homoscedastic model in modeling variances, we plot the conditional variances with the constant variances for the small and large size portfolios in Figure 1. We report the results for the equal- and value-weighted portfolios in Panels A, C and B, D, respectively. In all cases, we plot the conditional variances using solid lines and the constant variances using dotted lines. We observe that the variances from the heteroscedastic model are time-varying, whereas those from the homoscedastic model stay constant. For the homoscedastic model, small funds exhibit smaller variance than large funds for both equal- and value-weighted portfolio formation strategies. In contrast, for the heteroscedastic model and for the equal-weighted size portfolios, small funds produce larger variance than large funds during the early portion of the sample period. As for the value-weighted size portfolios, small funds have consistently smaller variance than large funds but the differences in variance change over time. In sum, Figure 1 suggests that the general heteroscedastic model provides time-varying variances which may be beneficial in conditional performance analysis.

Table 3 reports the ML parameter estimates for the two conditional alpha models for equal-weighted size portfolios. Panel A presents the parameter estimates for the heteroscedastic model. We observe that the majority of the coefficients are highly significant. The significant coefficients for the information instruments in the covariance specification suggest a potentially important nonlinear relationship between asset returns

---

27 Initial values of the parameters for the heteroscedastic conditional alpha model are obtained as follows. For conditional means, we run an ordinary least squares (OLS) regression of $r_{it}$ on $Z_{t-1}$ to obtain the starting parameter values. For conditional covariances, at each point of time $t$, we construct a square matrix in which the $i,j$th element is $e_{i,t}e_{j,t}$, where $e_{i,t} = \frac{1}{20} \sum_{k=0}^{20} (r_{i,t-k} - \overline{r}_t)$ and $\overline{r}_t$ is the time-series average of $r_{i,t}$ over the entire sample period. We take the Cholesky decomposition of the square matrix $[e_{i,t}e_{j,t}]$ at each time $t$. For each element in the resultant lower triangular matrix after the decomposition, we run an ordinary least squares (OLS) regression on $Z_{t-1}$ to obtain the starting parameter values. Robustness checks with alternative starting values suggest that our conditional alpha estimates are not particularly sensitive to the choice of the initial values.
and information instruments. In Panel B, we provide the parameter estimates for homoscedastic model. We find that the homoscedastic model has fewer parameters, and smaller number of significant coefficients. By design, the heteroscedastic model would be equivalent to the homoscedastic model if the coefficients for the conditional covariance instruments are all zero. We report Wald and Likelihood Ratio test statistics to examine the null hypothesis that the coefficients for conditional covariance instruments are all zero (i.e., $b_{ij1} = 0$ for $i \geq j$). Panel C presents the test statistics along with $p$-values in brackets. All tests statistics are highly significant, suggesting the heteroscedastic model is superior to the special case model under homoscedasticity.

*** Insert Table 3 about here ***

Table 4 reports analogous results for value-weighted size portfolios. In Panel A, we find that at least half of the coefficients for information instruments are significant. In Panel C, we notice that both the Wald test and the Likelihood Ratio test reject the null hypothesis that the heteroscedastic model has equivalent specification to the homoscedastic model. These results are highly comparable to those in Table 3.

*** Insert Table 4 about here ***

Table 5 presents the monthly conditional alphas for the small, mid, and large size portfolios. We report the results for equal-weighted size portfolios in Panel A. We find that for each size portfolio, both conditional models yield similar average conditional alphas. In addition, we find that the average conditional alphas are highly comparable to the unconditional alphas as documented in Table 2. Similar to the unconditional results, there is a strong monotonic pattern of the average conditional alphas declining with fund size. For example, according to the heteroscedastic model, the small funds have an average return of 0.25% per month or 3.00% per annum, after appropriate risk adjustment. The similar statistic for large funds is only 0.02% per month or 0.24% per annum. The analogous results for value-weighted size portfolios are provided in Panel B. We observe that for small and mid-size funds, the value-weighted results are similar to the equal-weighted findings. However, for large funds, the value-weighted portfolio produces smaller conditional alphas than their equal-weighted counterpart. In Table 5, we also report the standard deviation of conditional alphas for each size portfolio and for each conditional model. We find that the heteroscedastic alphas are more volatile.
than the homoscedastic alphas for equal-weighted size portfolios. In contrast, we do not see much difference in conditional alpha volatilities for value-weighted portfolios.

Because the conditional alphas are time-varying, it is natural to present them graphically. Figure 2 plots the evolution of conditional alphas for the small, mid, and large size portfolios over the sample period from 1991 to 2011. We observe strong differences in risk-adjusted performance between the small and large funds for virtually all periods. Similar results are found for the homoscedastic model in the lower two panels. Comparing the heteroscedastic model to the homoscedastic model, we find that both models produce similar pattern of conditional alphas but with different volatilities. In particular, the equal-weighted small and mid fund portfolios exhibit noticeably smaller volatilities of conditional alphas using the homoscedastic model than the heteroscedastic model. In contrast, all value-weighted size portfolios show similar level of volatilities across the two conditional models.

To assess the importance of changes in the conditional alphas with underlying instrument realizations, Table 6 provides the conditional alpha estimates given the 5, 25, 50, 75, and 95 percentile values of the information instruments within the sample period.\textsuperscript{28} We report the conditional alphas for the heteroscedastic model and the homoscedastic model in Panel A and B, respectively. The conditional alphas are estimated using the parameter estimates from Table 4 and 5.

Table 6 shows that the variability in the estimated conditional alphas is economically important. For example, any conditional alpha based on the 95 percentile is at least 3 times larger in absolute value than the alpha based on the 5 percentile instrument realization. In addition, for both equal- and value-weighted size portfolios, we find a clear pattern of conditional alphas increasing with instrument values for small and mid funds, and decreasing with instrument values for large funds. This result is consistent with the implication of the model by Berk and Green (2004). As a fund grows larger, managers become less flexible in taking on

\textsuperscript{28} We compute the percentile values for conditional mean instruments. The percentile values for conditional covariance instruments are then obtained as the absolute value of those for conditional mean instruments.
optimal positions of stocks, which worsens performance. Another interesting finding in Table 6 is that the heteroscedastic model appears to be more responsive to information instruments than the homoscedastic model for large funds. For example, for the equal-weighted large fund portfolio, the range of the heteroscedastic alpha based on the 5 and 95 percentile of instruments is 0.17% (=0.0934%−(−0.0743%)), which is about 6 times larger than the range of the homoscedastic alpha of 0.03% (=0.0231%−(−0.0049%)). Comparable results are also found for the value-weighted large fund portfolios.

3.5 Inferences on conditional performance measures

To examine the statistical significance of conditional alphas, we propose a conditional bootstrap approach for both heteroscedastic and homoscedastic models. We extend the unconditional bootstrapping procedure (e.g., Kosowski et al. (2006)) to the conditional regression context in which our conditional alphas are derived. Development details for the conditional bootstrap are provided in Appendix 2. For both heteroscedastic and homoscedastic models, we run 1,000 bootstrap simulations.

We plot the time-varying conditional alphas along with the 90 percent non-rejection regions for the small and large size portfolios. We report the results for the equal- and value-weighted portfolios in Figure 3 and 4, respectively. In both figures, we plot the conditional alphas using solid lines and the 90 percent non-rejections using dotted lines. The results for small fund portfolios are given in the upper two panels, whereas those for large fund portfolios are provided in the lower two panels. We observe that for the small fund portfolio, both heteroscedastic and homoscedastic models generate strong positive significance for much of the sample period, and this result seems persistent across both equal- and value-weighted portfolio formation strategies. Moreover, we find that for the equal-weighted large fund portfolio, the heteroscedastic model produces significantly positive conditional alphas for much of the time period. In contrast, no significance is found for the unconditional models for the same portfolio. The differences in reported significance between the conditional and unconditional models suggest the economic importance of conditional performance evaluation. In Figures 2 and 3, we also show that the heteroscedastic alphas produce different time series significance than the homoscedastic alphas. For example, for the equal-weighted large fund portfolio and out
of the 251 time series conditional alphas, the heteroscedastic model shows 197 significant observations, whereas the homoscedastic model identifies only 71. In contrast, for the value-weighted large fund portfolio, the heteroscedastic model results in 212 significant conditional alphas, which are 28 less than the homoscedastic model.

4. Concluding Remarks

We propose a new conditional alpha model in which return moments rationally evolve with predetermined economic information instruments. Unlike other homoscedastic performance evaluation methods, our new conditional alpha model captures the interesting nonlinear relationship between asset returns and information instruments. As an application, we use our approach to examine the conditional relationship between mutual fund size and performance after corrections for risk differences. We use the lagged fund flow as an information instruments for both conditional means and covariances as fund flow represents public information about a funds’ future performance. Consistent with Berk and Green (2004), we find strong evidence that fund performance declines with fund size. In particular, we find that small funds generate significantly larger risk-adjusted returns than large funds. The identified outperformance of small funds relative to large funds is robust to different unconditional and conditional models, and to different portfolio formation strategies. Moreover, we find that the lagged fund flow is positively related to the performance of small and mid-sized funds, but negatively related to larger fund performance.

To make statistical inferences about the conditional alphas, we develop a new conditional bootstrap procedure. Using this new inference procedure, we find that although the heteroscedastic model produces different time series significance than the homoscedastic model, small funds outperform large funds for virtually all time periods in our sample. This new framework of conditional alphas and conditional inferences can be used to study the risk-adjusted performance of any financial asset whose returns rationally evolve with economic fundamentals.
References


Fama, E., and K. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of
Financial Economics* 33, 3-56.

Ferson, W., and R. Schadt, 1996, Measuring Fund Strategy and Performance in Changing

56, 967-982.

Grinblatt, M., and S. Titman, 1989, Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings,
*Journal of Business* 62, 393-416.

Gruber, M., 1996, Another Puzzle: The Growth in Actively Managed Mutual Funds, *Journal of Finance* 51, 783-
810.

Hendricks, D., J. Patel and R. Zeckhauser, 1993, Hot Hands in Mutual Funds: Short-run Persistence of


Jha, R., B. Korkie, and H. Turtle, 2009, Measuring Performance in a Dynamic World: Conditional

Administrative Sciences* 15, 213-229.


Appendix 1 – Conditional alpha with conditional covariances

In the simplest case of only one asset and one factor portfolio, the conditional alpha is given by,

\[
\alpha_{2|z_{t-1}} = a_{20} + a_{21}Z_{2,t-1} - (a_{10} + a_{11}Z_{1,t-1})\sigma_{21|z_{\sigma,t-1}}\sigma_{11|z_{\sigma,t-1}}^{-1} \tag{A1.1}
\]

Suppose the conditional covariance matrix, \(\Sigma|z_{\sigma,t-1}\), has the following specification with respect to information instruments, \(Z_{\sigma,t-1}\),

\[
\Sigma|z_{\sigma,t-1} = L|z_{\sigma,t-1}L' \tag{A1.2}
\]

where \(L|z_{\sigma,t-1} = [l_{ij}|z_{\sigma,t-1}]\); and where \(l_{ij}|z_{\sigma,t-1} = b_{ij0} + b_{ij1}Z_{i\omega,t-1}\).

Expanding equation (A1.2), we have,

\[
\Sigma|z_{\sigma,t-1} = \begin{bmatrix}
    b_{110} + b_{111}Z_{1\sigma,t-1} & 0 \\
    b_{210} + b_{211}Z_{2\sigma,t-1} & b_{220} + b_{221}Z_{2\sigma,t-1} \\
    b_{110} + b_{111}Z_{1\sigma,t-1} & b_{210} + b_{211}Z_{2\sigma,t-1} & b_{220} + b_{221}Z_{2\sigma,t-1}
\end{bmatrix} \tag{A1.3}
\]

Evaluating term by term, we get,

\[
\sigma_{21|z_{\sigma,t-1}}\sigma_{11|z_{\sigma,t-1}}^{-1} = \frac{b_{110} + b_{111}Z_{1\sigma,t-1}}{b_{110} + b_{111}Z_{1\sigma,t-1}} \tag{A1.4}
\]

Substituting (A1.4) in (A1.1), we have,

\[
\alpha_{2|z_{t-1}} = a_{20} + a_{21}Z_{2,t-1} - (a_{10} + a_{11}Z_{1,t-1})\frac{b_{110} + b_{111}Z_{1\sigma,t-1}}{b_{110} + b_{111}Z_{1\sigma,t-1}} \tag{A1.5}
\]
Appendix 2 – Inferences with conditional bootstrap

One important implication from the conditional alpha expression is that it is equivalent to a multivariate regression intercept obtained using conditional moments and returns given by,

\[ r_{2|zt-1} = \alpha_{2|zt-1} + \beta_{2|zt-1} r_{1|zt-1} + e_{2|zt-1} \]  
(A2.1)

where \( \alpha_{2|zt-1} \) and \( \beta_{2|zt-1} \) are given by (1) and (2); and \( r_{2|zt-1} \) and \( r_{1|zt-1} \) are conditional excess returns for \( N_2 \) individual assets and \( N_1 \) spanning portfolios. Following Anderson (2003), \( r_{2|zt-1} \) and \( r_{1|zt-1} \) are given by,

\[ r_{2|zt-1} = r_2 - \mu_{2|zt-1} \]  
(A2.2)

\[ r_{1|zt-1} = r_1 - \mu_{1|zt-1} \]  
(A2.3)

For ease of demonstration, consider a simplest case where there are one individual asset and one spanning portfolio. To make statistical inferences about \( \alpha_{2|zt-1} \), we use a conditional bootstrap procedure described as follows:

1) Estimate \( \alpha_{2|zt-1} \) using equation (8) and obtain the model residuals \( e_{2|zt-1} \) in (A2.1).

2) For each bootstrap replication, resample \( e_{2|zt-1} \) without replacement to get a pseudo time series of residuals \( e^b_{2|zt-1} \), and construct a time-series of bootstrapped excess returns under the null hypothesis of zero conditional alphas, \( \alpha_{2|zt-1} = 0 \). The resultant bootstrapped excess returns are given by,

\[ r^b_{2|zt-1} = \tilde{\beta}_{2|zt-1} \tilde{r}_{1|zt-1} + e^b_{2|zt-1} \]  
(A2.4)

Substituting (A2.2) and (A2.3) in (A2.4), we have,

\[ r^b_{2|zt-1} = \tilde{\beta}_{2|zt-1} [\tilde{r}_{1|zt-1} - \tilde{\mu}_{1|zt-1}] + e^b_{2|zt-1} \]  
(A2.5)

where \( \tilde{\beta}_{2|zt-1} = \tilde{\beta}_{21|zt-1} \tilde{\alpha}_{11|zt-1} = \frac{\tilde{\beta}_{210} + \tilde{\beta}_{211} Z_{2|t-1}}{\tilde{\beta}_{110} + \tilde{\beta}_{111} Z_{1|t-1}} \); and where \( \tilde{\beta}_{2|zt-1} = \tilde{\alpha}_{20} + \tilde{\alpha}_{21} Z_{2|t-1} \) and \( \tilde{\mu}_{1|zt-1} = \tilde{\alpha}_{10} + \tilde{\alpha}_{11} Z_{1|t-1} \). Similar to Kosowski et al. (2006), we do not randomize over \( r_1 \) and \( Z_{t-1} \) when constructing \( r^b_{2|zt-1} \).

3) For each of \( b = 1, 2, ..., B \), bootstrap replications, estimate the conditional alphas using equation (8) and obtain \( \tilde{\alpha}^b_{2|zt-1} \).

4) The observed empirical distribution of \( \tilde{\alpha}^b_{2|zt-1} \) may be used to determine the non-rejection region for given test sizes for each point in time and for every instrument realization.

---

\(^29\) We resample the residuals for all assets in the cross-section to maintain any potential correlation in the cross-section.
### Table 1
Descriptive statistics for the mutual fund sample

#### Panel A: All funds

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>25th Pctl.</th>
<th>Median</th>
<th>75 Pctl.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of funds</td>
<td>1119.10</td>
<td>534.92</td>
<td>289.00</td>
<td>602.00</td>
<td>1139.00</td>
<td>1629.00</td>
<td>1991.00</td>
</tr>
<tr>
<td>Total net assets (mil)</td>
<td>918.65</td>
<td>2225.75</td>
<td>0.74</td>
<td>54.35</td>
<td>213.42</td>
<td>799.32</td>
<td>30069.74</td>
</tr>
<tr>
<td>Total net flow (mil)</td>
<td>2.02</td>
<td>33.24</td>
<td>-230.77</td>
<td>-3.02</td>
<td>0.01</td>
<td>3.44</td>
<td>299.59</td>
</tr>
<tr>
<td>Scaled fund flow (%)</td>
<td>1.46</td>
<td>7.95</td>
<td>-29.61</td>
<td>-1.13</td>
<td>0.03</td>
<td>1.93</td>
<td>100.43</td>
</tr>
<tr>
<td>Expense ratio (%)</td>
<td>1.17</td>
<td>0.45</td>
<td>-0.51</td>
<td>0.92</td>
<td>1.17</td>
<td>1.40</td>
<td>13.48</td>
</tr>
<tr>
<td>Portfolio market cap (mil)</td>
<td>17947.90</td>
<td>18930.36</td>
<td>29.53</td>
<td>2265.72</td>
<td>9817.21</td>
<td>30798.09</td>
<td>176891.35</td>
</tr>
</tbody>
</table>

#### Panel B: Small funds

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>25th Pctl.</th>
<th>Median</th>
<th>75 Pctl.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total net assets (mil)</td>
<td>17.75</td>
<td>12.47</td>
<td>0.74</td>
<td>7.15</td>
<td>15.80</td>
<td>26.08</td>
<td>60.55</td>
</tr>
<tr>
<td>Total net flow (mil)</td>
<td>0.31</td>
<td>1.75</td>
<td>-25.96</td>
<td>-0.12</td>
<td>0.04</td>
<td>0.46</td>
<td>25.93</td>
</tr>
<tr>
<td>Scaled fund flow (%)</td>
<td>3.38</td>
<td>11.96</td>
<td>-29.61</td>
<td>-0.84</td>
<td>0.43</td>
<td>3.77</td>
<td>100.43</td>
</tr>
<tr>
<td>Expense ratio (%)</td>
<td>1.49</td>
<td>0.56</td>
<td>0.00</td>
<td>1.20</td>
<td>1.45</td>
<td>1.74</td>
<td>13.48</td>
</tr>
<tr>
<td>Portfolio market cap (mil)</td>
<td>14397.86</td>
<td>16919.82</td>
<td>30.81</td>
<td>1611.94</td>
<td>6602.64</td>
<td>23099.30</td>
<td>176891.35</td>
</tr>
</tbody>
</table>

#### Panel C: Mid funds

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>25th Pctl.</th>
<th>Median</th>
<th>75 Pctl.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total net assets (mil)</td>
<td>230.30</td>
<td>86.58</td>
<td>48.06</td>
<td>163.71</td>
<td>214.55</td>
<td>285.16</td>
<td>544.94</td>
</tr>
<tr>
<td>Total net flow (mil)</td>
<td>1.44</td>
<td>12.63</td>
<td>-155.68</td>
<td>-2.97</td>
<td>-0.14</td>
<td>3.84</td>
<td>176.13</td>
</tr>
<tr>
<td>Scaled fund flow (%)</td>
<td>1.12</td>
<td>6.94</td>
<td>-29.56</td>
<td>-1.32</td>
<td>-0.07</td>
<td>1.89</td>
<td>100.30</td>
</tr>
<tr>
<td>Expense ratio (%)</td>
<td>1.16</td>
<td>0.36</td>
<td>0.00</td>
<td>0.95</td>
<td>1.19</td>
<td>1.38</td>
<td>3.59</td>
</tr>
<tr>
<td>Portfolio market cap (mil)</td>
<td>16555.93</td>
<td>18803.93</td>
<td>82.26</td>
<td>1658.25</td>
<td>7370.30</td>
<td>28366.32</td>
<td>131011.29</td>
</tr>
</tbody>
</table>
## Panel D: Large funds

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>25th Pctl.</th>
<th>Median</th>
<th>75th Pctl.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total net assets ($mil)</td>
<td>3620.64</td>
<td>3919.00</td>
<td>461.88</td>
<td>1454.64</td>
<td>2181.13</td>
<td>4070.35</td>
<td>30069.74</td>
</tr>
<tr>
<td>Total net flow ($mil)</td>
<td>5.02</td>
<td>67.73</td>
<td>-230.77</td>
<td>-23.68</td>
<td>-2.40</td>
<td>25.39</td>
<td>299.59</td>
</tr>
<tr>
<td>Scaled fund flow (%)</td>
<td>0.25</td>
<td>2.62</td>
<td>-16.82</td>
<td>-0.97</td>
<td>-0.11</td>
<td>1.01</td>
<td>58.97</td>
</tr>
<tr>
<td>Expense ratio (%)</td>
<td>0.90</td>
<td>0.36</td>
<td>0.00</td>
<td>0.71</td>
<td>0.93</td>
<td>1.13</td>
<td>3.04</td>
</tr>
<tr>
<td>Portfolio market cap ($mil)</td>
<td>23676.28</td>
<td>19868.99</td>
<td>66.00</td>
<td>5519.06</td>
<td>20487.71</td>
<td>38196.22</td>
<td>143495.67</td>
</tr>
</tbody>
</table>

We report the summary statistics for fund characteristics. The sample includes 2,036 open-end domestic equity mutual funds with 280,894 monthly observations over the period from 1991 to 2011. We exclude the mutual funds that have an initial minimum purchase amount of more than $20,000. We classify mutual funds into domestic funds if the percentage of US stock holdings is equal to or greater than 80 percent. For mutual funds with multiple share classes, we retain only the largest mutual fund class as measured by net assets at the end of December 2011. Fund returns and characteristics are from the Morningstar Direct database. The total net assets, total net flow, and portfolio market cap are obtained at the fund level, whereas the expense ratio is for the largest fund class and is changed on a per annum basis. The scaled fund flow is computed as dividing total net flow by previous month’s total net assets. All extreme values of total net assets and total net flow are trimmed at the 99% level when calculating summary statistics.
Table 2
Monthly market-adjusted returns and unconditional alphas for size portfolios

Panel A: Equal-weighted size portfolios, CRSP index

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Small vs. Large t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-adjusted return (%)</td>
<td>0.1541 (0.6800)</td>
<td>0.0283 (0.9406)</td>
<td>-0.0593 (0.8778)</td>
<td>0.2134 (0.6909)</td>
</tr>
<tr>
<td>CAPM model alphas (%)</td>
<td>0.6465 (0.0228)</td>
<td>0.5244 (0.0717)</td>
<td>0.4484 (0.1256)</td>
<td>0.1981 (0.0002)</td>
</tr>
<tr>
<td>FF 3-factor model alphas (%)</td>
<td>0.6744 (0.0193)</td>
<td>0.5601 (0.0581)</td>
<td>0.4872 (0.1009)</td>
<td>0.1872 (0.0005)</td>
</tr>
<tr>
<td>Carhart 4-factor model alphas (%)</td>
<td>0.7195 (0.0138)</td>
<td>0.6072 (0.0426)</td>
<td>0.5222 (0.0828)</td>
<td>0.1973 (0.0003)</td>
</tr>
</tbody>
</table>

Panel B: Equal-weighted size portfolios, sample fund index

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Small vs. Large t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-adjusted return (%)</td>
<td>0.2219 (&lt;0.0001)</td>
<td>0.0961 (0.0166)</td>
<td>0.0085 (0.3787)</td>
<td>0.2134 (0.0002)</td>
</tr>
<tr>
<td>CAPM model alphas (%)</td>
<td>0.2469 (&lt;0.0001)</td>
<td>0.1042 (0.0096)</td>
<td>0.0104 (0.2845)</td>
<td>0.2365 (&lt;0.0001)</td>
</tr>
<tr>
<td>FF 3-factor model alphas (%)</td>
<td>0.2424 (&lt;0.0001)</td>
<td>0.1066 (0.0090)</td>
<td>0.0128 (0.1890)</td>
<td>0.2296 (&lt;0.0001)</td>
</tr>
<tr>
<td>Carhart 4-factor model alphas (%)</td>
<td>0.2591 (&lt;0.0001)</td>
<td>0.1226 (0.0024)</td>
<td>0.0155 (0.1132)</td>
<td>0.2436 (&lt;0.0001)</td>
</tr>
</tbody>
</table>

Panel C: Value-weighted size portfolios, CRSP index

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Small vs. Large t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-adjusted return (%)</td>
<td>0.1371 (0.7135)</td>
<td>0.0301 (0.9369)</td>
<td>-0.0878 (0.8211)</td>
<td>0.2249 (0.6761)</td>
</tr>
<tr>
<td>CAPM model alphas (%)</td>
<td>0.6312 (0.0257)</td>
<td>0.5260 (0.0706)</td>
<td>0.4247 (0.1484)</td>
<td>0.2065 (0.0008)</td>
</tr>
<tr>
<td>FF 3-factor model alphas (%)</td>
<td>0.6598 (0.0216)</td>
<td>0.5622 (0.0569)</td>
<td>0.4600 (0.1229)</td>
<td>0.1998 (0.0013)</td>
</tr>
<tr>
<td>Carhart 4-factor model alphas (%)</td>
<td>0.7029 (0.0158)</td>
<td>0.6104 (0.0413)</td>
<td>0.4896 (0.1054)</td>
<td>0.2133 (0.0007)</td>
</tr>
</tbody>
</table>
Panel D: Value-weighted size portfolios, sample fund index

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Small vs. Large t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market-adjusted return (%)</td>
<td>0.2049 (0.0002)</td>
<td>0.0979 (0.0161)</td>
<td>-0.0200 (0.0137)</td>
<td>0.2249 (&lt;0.0001)</td>
</tr>
<tr>
<td>CAPM model alphas (%)</td>
<td>0.2312 (&lt;0.0001)</td>
<td>0.1066 (0.0089)</td>
<td>-0.0198 (0.0154)</td>
<td>0.2510 (&lt;0.0001)</td>
</tr>
<tr>
<td>FF 3-factor model alphas (%)</td>
<td>0.2273 (&lt;0.0001)</td>
<td>0.1096 (0.0080)</td>
<td>-0.0212 (0.0107)</td>
<td>0.2485 (&lt;0.0001)</td>
</tr>
<tr>
<td>Carhart 4-factor model alphas (%)</td>
<td>0.2422 (&lt;0.0001)</td>
<td>0.1264 (0.0020)</td>
<td>-0.0239 (0.0037)</td>
<td>0.2661 (&lt;0.0001)</td>
</tr>
</tbody>
</table>

We report the unconditional results for equal- and value-weighted size portfolios over the sample period from 1991 to 2011. Size portfolios are formed and updated according to each fund’s total net assets at the end of each month. The market-adjusted returns are calculated as subtracting market returns from fund total returns, where the market returns are obtained either from a value-weighted CRSP index or from a value-weighted portfolio including all funds in the sample. All fund returns are net of administrative and management fees. We present the unconditional results using a value-weighted CRSP index in Panels A and C and a value-weighted sample fund index in Panels B and D, respectively. We report the time series average of market-adjusted returns with the $p$-values given in parentheses. The CAPM model alpha estimates are based on the unconditional regression with the market excess return as the sole risk factor. The FF 3-factor model alpha estimates are derived from the unconditional regression with Fama-French three factors as regressors. For CAPM and the FF 3-factor model, we report the estimated alphas along with the $p$-values in parentheses.
Table 3
Parameter estimates for equal-weighted size portfolios

### Panel A: Heteroscedastic model estimates

<table>
<thead>
<tr>
<th>Fund Size Portfolios</th>
<th>Market</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Conditional Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b_{l0}$</td>
<td>$b_{l1}$</td>
<td>$b_{l20}$</td>
<td>$b_{l21}$</td>
<td>$b_{l30}$</td>
</tr>
<tr>
<td>Market</td>
<td>4.6673</td>
<td>-11.3538</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.3151)</td>
<td>(5.2139)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>4.3955</td>
<td>-0.0939</td>
<td>0.6751</td>
<td>4.1357</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.3036)</td>
<td>(1.6210)</td>
<td>(0.0693)</td>
<td>(1.5540)</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>4.5457</td>
<td>-1.3735</td>
<td>0.5838</td>
<td>-3.6578</td>
<td>0.2167</td>
</tr>
<tr>
<td></td>
<td>(0.3100)</td>
<td>(2.9597)</td>
<td>(0.0561)</td>
<td>(2.5988)</td>
<td>(0.0234)</td>
</tr>
<tr>
<td>Large</td>
<td>4.6547</td>
<td>-13.5472</td>
<td>0.0132</td>
<td>8.5360</td>
<td>-0.0232</td>
</tr>
<tr>
<td></td>
<td>(0.3140)</td>
<td>(5.6807)</td>
<td>(0.0114)</td>
<td>(2.2454)</td>
<td>(0.0125)</td>
</tr>
</tbody>
</table>

### Panel B: Homoscedastic model estimates

<table>
<thead>
<tr>
<th>Fund Size Portfolios</th>
<th>$a_{l0}$</th>
<th>$a_{l1}$</th>
<th>$\beta_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.1362 (0.0891)</td>
<td>2.7343 (1.7474)</td>
<td>0.9508 (0.0117)</td>
</tr>
<tr>
<td>Mid</td>
<td>0.0896 (0.0513)</td>
<td>0.9127 (2.0251)</td>
<td>0.9839 (0.0086)</td>
</tr>
<tr>
<td>Large</td>
<td>0.0173 (0.0118)</td>
<td>-1.4240 (1.3989)</td>
<td>0.9963 (0.0021)</td>
</tr>
</tbody>
</table>

### Panel C: Model comparison

<table>
<thead>
<tr>
<th>Statistical Test</th>
<th>Test Statistic</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald Test</td>
<td>256.80 [&lt;0.0001]</td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio Test</td>
<td>453.86 [&lt;0.0001]</td>
<td></td>
</tr>
</tbody>
</table>

We report estimation results for equal-weighted size portfolios using both heteroscedastic and homoscedastic models over the sample period from 1991 to 2011. For the heteroscedastic model, the conditional mean is given by $\mu_{i|\mathbf{z}_{it-1}} = a_{i0} + a_{i1}Z_{it-1}$; and the conditional covariance matrix $\Sigma_{i|\mathbf{z}_{it-1}}$ is modeled as $\Sigma_{i|\mathbf{z}_{it-1}} = \mathbf{L}_t\mathbf{U}_t$, where $\mathbf{L}_t = [l_{ij}]$ is the lower triangular matrix at $t$, and $l_{ij} = b_{i0} + b_{i1}Z_{i,t-1}$ for predetermined information instrument $Z_{i,t-1}$. For the homoscedastic model, the unconditional regression for portfolio $i$ is given by $r_{it} = a_{i0} + a_{i1}Z_{i,t-1} + \beta_l r_{mt} + \epsilon_{it}$, where $r_{it}$ and $r_{mt}$ are the excess returns for the fund size portfolio $i$ and for the market index, respectively. In Panels A and B, we report the parameter estimates with standard errors in parentheses for the heteroscedastic and homoscedastic model, respectively. For both models, the coefficients are obtained by the ML method. In Panel C, we provide test statistics for model comparison using both the Wald test and the Likelihood Ratio test. The null hypothesis for both tests is $b_{i1t} = 0$ for $i \geq j$. The $p$-values are given in brackets.
Table 4

Parameter estimates for value-weighted size portfolios

Panel A: Heteroscedastic model estimates

<table>
<thead>
<tr>
<th>Fund Size Portfolios</th>
<th>Market</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
<th>Conditional Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_{i10}$</td>
<td>$b_{i11}$</td>
<td>$b_{i20}$</td>
<td>$b_{i21}$</td>
<td>$b_{i30}$</td>
</tr>
<tr>
<td>Market</td>
<td>4.5942</td>
<td>2.4766</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(0.0433)</td>
<td>(2.2856)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>4.3792</td>
<td>-0.0727</td>
<td>0.6709</td>
<td>3.9198</td>
<td>-</td>
</tr>
<tr>
<td>(0.0459)</td>
<td>(1.6234)</td>
<td>(0.0532)</td>
<td>(1.2238)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>4.5305</td>
<td>0.1447</td>
<td>0.5020</td>
<td>1.6881</td>
<td>0.2602</td>
</tr>
<tr>
<td>(0.0055)</td>
<td>(2.1685)</td>
<td>(0.0408)</td>
<td>(1.4714)</td>
<td>(0.0220)</td>
<td>(1.0369)</td>
</tr>
<tr>
<td>Large</td>
<td>4.6024</td>
<td>0.9417</td>
<td>-0.0824</td>
<td>-3.8763</td>
<td>-0.0436</td>
</tr>
<tr>
<td>(0.0435)</td>
<td>(2.8269)</td>
<td>(0.0064)</td>
<td>(0.9426)</td>
<td>(0.0045)</td>
<td>(0.8662)</td>
</tr>
</tbody>
</table>

Panel B: Homoscedastic model estimates

<table>
<thead>
<tr>
<th>Fund Size Portfolios</th>
<th>$a_{i0}$</th>
<th>$a_{i1}$</th>
<th>$\beta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.0726 (0.0406)</td>
<td>4.4724 (1.4240)</td>
<td>0.9486 (0.0114)</td>
</tr>
<tr>
<td>Mid</td>
<td>0.0307 (0.0457)</td>
<td>4.8771 (1.3899)</td>
<td>0.9829 (0.0087)</td>
</tr>
<tr>
<td>Large</td>
<td>-0.0154 (0.0087)</td>
<td>-1.0797 (0.7373)</td>
<td>0.9997 (0.0018)</td>
</tr>
</tbody>
</table>

Panel C: Model comparison

<table>
<thead>
<tr>
<th>Statistical Test</th>
<th>Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald Test</td>
<td>172.13 [&lt;0.0001]</td>
</tr>
<tr>
<td>Likelihood Ratio Test</td>
<td>473.54 [&lt;0.0001]</td>
</tr>
</tbody>
</table>

We report estimation results for value-weighted size portfolios using both heteroscedastic and homoscedastic models over the sample period from 1991 to 2011. For the heteroscedastic model, the conditional mean is given by $\mu_{i,t-1} = a_{i0} + a_{i1}Z_{i,t-1}$; and the conditional covariance matrix $\Sigma_{i,t-1}$ is modeled as $\Sigma_{i,t-1} = L_tL_t'$, where $L_t = [l_{i,t}]$ is the lower triangular matrix at $t$, and $l_{i,t} = b_{i10} + b_{i12}Z_{i,t-1}$ for predetermined information instrument $Z_{i,t-1}$. For the homoscedastic model, the unconditional regression for portfolio $i$ is given by $\sigma_{i,t} = a_{i0} + a_{i1}Z_{i,t-1} + \beta_i r_{mt} + \epsilon_{it}$, where $r_{it}$ and $r_{mt}$ are the excess returns for the fund size portfolio $i$ and for the market index, respectively. In Panels A and B, we report the parameter estimates with standard errors in parentheses for the heteroscedastic and homoscedastic model, respectively. For both models, the coefficients are obtained by the ML method. In Panel C, we provide test statistics for model comparison using both the Wald test and the Likelihood Ratio test. The null hypothesis for both tests is $b_{i1j} = 0$ for $i \geq j$. The $p$-values are given in brackets.
Table 5
Conditional alphas for size portfolios

Panel A: Equal-weighted size portfolios

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroscedastic conditional alpha (%)</td>
<td>0.2491 [0.1109]</td>
<td>0.1171 [0.0719]</td>
<td>0.0185 [0.0142]</td>
</tr>
<tr>
<td>Homoscedastic conditional alpha (%)</td>
<td>0.2465 [0.0588]</td>
<td>0.1042 [0.0121]</td>
<td>0.0104 [0.0091]</td>
</tr>
</tbody>
</table>

Panel B: Value-weighted size portfolios

<table>
<thead>
<tr>
<th>Mutual Fund Size Quintile</th>
<th>Small</th>
<th>Mid</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroscedastic conditional alpha (%)</td>
<td>0.2425 [0.0853]</td>
<td>0.1192 [0.0480]</td>
<td>-0.0249 [0.0078]</td>
</tr>
<tr>
<td>Homoscedastic conditional alpha (%)</td>
<td>0.2306 [0.1003]</td>
<td>0.1064 [0.0673]</td>
<td>-0.0198 [0.0061]</td>
</tr>
</tbody>
</table>

We report the conditional alphas for size portfolios over the sample period from 1991 to 2011. Size portfolios are formed and updated according to each fund’s total net assets at the end of each month. We use two portfolio formation strategies and present the results for the equal-weighted and value-weighted size portfolios in Panel A and B, respectively. We report the time series average of the conditional alphas for both heteroscedastic and homoscedastic models. The standard deviations are given in brackets.
Table 6
Conditional alphas based on instrument percentiles for size portfolios

Panel A: Heteroscedastic conditional alpha

<table>
<thead>
<tr>
<th>Instrument Percentiles</th>
<th>Equal-weighted Mutual Fund Size Quintile</th>
<th>Value-weighted Mutual Fund Size Quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small</td>
<td>Mid</td>
</tr>
<tr>
<td>5%</td>
<td>0.1567</td>
<td>0.0908</td>
</tr>
<tr>
<td>25%</td>
<td>0.1948</td>
<td>0.0999</td>
</tr>
<tr>
<td>50%</td>
<td>0.2359</td>
<td>0.1088</td>
</tr>
<tr>
<td>75%</td>
<td>0.2838</td>
<td>0.1195</td>
</tr>
<tr>
<td>95%</td>
<td>0.3638</td>
<td>0.1361</td>
</tr>
</tbody>
</table>

Panel B: Homoscedastic conditional alpha

<table>
<thead>
<tr>
<th>Instrument Percentiles</th>
<th>Equal-weighted Mutual Fund Size Quintile</th>
<th>Value-weighted Mutual Fund Size Quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small</td>
<td>Mid</td>
</tr>
<tr>
<td>5%</td>
<td>0.1645</td>
<td>0.0870</td>
</tr>
<tr>
<td>25%</td>
<td>0.2009</td>
<td>0.0954</td>
</tr>
<tr>
<td>50%</td>
<td>0.2401</td>
<td>0.1028</td>
</tr>
<tr>
<td>75%</td>
<td>0.2859</td>
<td>0.1117</td>
</tr>
<tr>
<td>95%</td>
<td>0.3623</td>
<td>0.1255</td>
</tr>
</tbody>
</table>

We present the conditional alpha estimates for the 5, 25, 50, 75, and 95 percentiles of the information instruments for both equal- and value-weighted size portfolios over the sample period from 1991 to 2011. The percentile values are computed for each information instrument across the sample period. The conditional alpha estimates are computed using the parameters given in Table 4 and 5.
Figure 1. Conditional and unconditional variances for small and large fund portfolios
Figure 2. Conditional alphas for small, mid, and large fund portfolios
Figure 3. Conditional alphas and 90% non-rejection regions by conditional bootstrap for equal-weighted small and large fund
Figure 4. Conditional alphas and 90% non-rejection regions by conditional bootstrap for value-weighted small and large fund portfolios.