CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Cardiff Economics
Working Papers

Woon K. Wong

A Unique Orthogonal Variance Decomposition

E2008/10

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
April 2008

This working paper is produced for discussion purpose only. These working papers are expected to be published in due course, in revised form, and should not be quoted or cited without the author's written permission.
Cardiff Economics Working Papers are available online from: http://www.cardiff.ac.uk/carbs/econ/workingpapers
Enquiries: EconWP@cardiff.ac.uk

A Unique Orthogonal Variance Decomposition

Woon K. Wong*

Investment Management Research Unit

Cardiff Business School

13 March 2008

* Cardiff University, Aberconway Building, Colum Drive, Cardiff CF10 3EU. Tel: +44 29 28075079; Fax: +44 29 28074419; Email: wongwk3@cardiff.ac.uk.

A UNIQUE ORTHOGONAL VARIANCE DECOMPOSITION

ABSTRACT

Let e and Σ be respectively the vector of shocks and its variance covariance matrix in a linear system of equations in reduced form. This article shows that a unique orthogonal variance decomposition can be obtained if we impose a restriction that maximizes the trace of A, a positive definite matrix such that $Az = e$ where z is vector of uncorrelated shocks with unit variance. Such a restriction is meaningful in that it associates the largest possible weight for each element in e with its corresponding element in z. It turns out that $A = \Sigma^{1/2}$, the square root of Σ.

KEYWORDS: Variance decomposition, Cholesky decomposition, unique orthogonal decomposition and square root matrix.

JEL Classification: C01

1. INTRODUCTION

VARIANCE DECOMPOSITION IS OFTEN CARRIED OUT in an econometric analysis. For example, Shorrocks (1982) considers the issue of inequality decomposition by factor components. In a structural VAR system, economic theory is often employed in order to construct the structural shocks that are uncorrelated with each other; see, for instance, Sims (1986), Bernanke (1986) and Blanchard and Quah (1989). However, it is well known that the variance decomposition for a single equation system is not unique. In the case of a structural VAR analysis, the selection of ordering in Cholesky decomposition is generally ad hoc, and convincing identifying assumptions are hard to come by.[1] This article proposes a unique orthogonal variance decomposition that can be applied to both single as well as multiple equation system.

[1] This is evidenced from the remark by Hamilton (1994, p. 335) that "if there were compelling identifying assumptions for such a system, the fierce debates among macroeconomists would have been settled long ago!"

Let e and Σ be respectively the vector of shocks and its variance covariance matrix in the equation system in reduced form. Let A be a decomposition matrix such that $AA' = \Sigma$ and $Az = e$ where z is vector of uncorrelated shocks with unit variance. It is shown that if we restrict A to be positive definite and its trace be maximized, a unique decomposition matrix given by $A = \Sigma^{1/2}$ is obtained. While A being positive definite is a very general condition, the maximization of its trace is intuitively appealing. For the sake of argument, let's consider z as the underlying structural shock.[2] Then, the higher is the trace of A, the less is, say the i-th component shock in e, can be linearly explained by the other components of structural shocks in z. In this sense, the trace of A measures the extent for each individual component of e, which may be regarded as observable, to be explained by its own corresponding structural shock component. This is particularly meaningful if there is no economic theory available to identify the system.

This paper is organized as follows. Section 2 provides the motivation and proof of a unique orthogonal decomposition for a simple two-variable system. Section 3 generalizes the result to the n-variable case. Numerical examples are given in the Section 4 and Section 5 concludes with some remarks.

2. The Two-Variable Case

2.1 Motivation

Without loss of generality, let us consider the following simple case: $y = e_1 + e_2$, where $e = (e_1\ e_2)'$ is observable, serially uncorrelated with[3]

$$\mathrm{var}(e) = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}.$$

In a structural VAR analysis, e may be regarded as driven by underlying uncorrelated structural shocks $\varepsilon = (\varepsilon_1\ \varepsilon_2)'$ in the following manner

(1) $$e_1 = \gamma_{11}\varepsilon_1 + \gamma_{12}\varepsilon_2,$$

(2) $$e_2 = \gamma_{21}\varepsilon_1 + \gamma_{22}\varepsilon_2.$$

---

[2] If we are interested in determining the contribution of a component towards the total variation, then the variance of the corresponding structural shock can be arbitrarily set to one; see Remark 3.

[3] If e is not observable, they are normally estimated as the residuals of a VAR in reduced form.

One general interest is to find out the contribution of, say ε_1, in the variation of *y*. From the above system of equations, the required variance contribution is $(\gamma_{11} + \gamma_{21})^2 \text{ var}(\varepsilon_1)$. But ε is unobservable and γ's are unknown. In the absence of economic theory, this paper proposes to decompose the variation in $y = e_1 + e_2$ by considering a positive definite decomposition matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

such that $e = Az$, which can be written as

(3) $$e_1 = a_{11}z_1 + a_{12}z_2 ,$$

(4) $$e_2 = a_{21}z_1 + a_{22}z_2 .$$

$z = (z_1 \ z_2)'$ is a serially uncorrelated random vector with identity covariance matrix. We may regard the random variable z_1 as unit structural shocks associated with ε_1. The variance contribution of interest can then be regarded as the variation contributed by z_1, which is $(a_{11} + a_{21})^2$. Now to obtain *A*, we make use of the fact that $AA' = \Sigma$, which implies

(5) $$a_{11}^2 + a_{12}^2 = \sigma_{11} ,$$

(6) $$a_{22}^2 + a_{21}^2 = \sigma_{22} ,$$

(7) $$a_{11}a_{21} + a_{22}a_{12} = \sigma_{12} .$$

Since there are only three equations available to solve for four unknowns in *A*, the decomposition matrix is not unique, which is a well known fact. One approach is to apply the Cholesky decomposition, for which we may choose to restrict a_{12} to be zero, assuming that z_2 does not contribute towards e_1 in (3). Since *A* is positive definite, both a_{11} and a_{22} are positive. Then restricting $a_{12} = 0$ implies that a_{11} attains its largest possible value, the standard deviation of e_1, and z_1 contributes to the variation in *y* by $(a_{11} + a_{21})^2$. Alternatively, one could vary a_{12} so that $a_{21} = 0$. This is equivalent to using the other ordering choice in the Cholesky decomposition, resulting in the assumption that z_1 does not contribute towards e_2 at all. We have the opposite effect in this case: the variance contribution of z_1 is simply a_{11}^2 with the magnitude of a_{11} reduced, and a_{22} attains its maximum value, which is the standard deviation of e_2.

The above example illustrates why the selection of ordering in the Cholesky decomposition could affect the outcome drastically. However, it does offer hints on how we could decompose a variance if there is no economic theory available to identify the system. First, changes in an element of *A*, say a_{12}, causes corresponding changes in all other elements of *A* so that the relationship, $AA' = \Sigma$, is maintained. Second, in the absence of economic theory, it is meaningful to choose a value of a_{12} such that the trace of *A*, $a_{11} + a_{22}$, are maximized. Doing so is equivalent to the system being identified with each e_i associated with maximum weight given to its corresponding z_i; the largest possible association between each e_i and its underlying structural shock. Indeed, this paper shows that when the choice of elements of *A* is restricted to maximizing the trace of the matrix, *A* is uniquely determined to be $\Sigma^{1/2}$.

2.2 Orthogonal Decomposition for the 2×2 Case

We shall now proceed to obtain the unique orthogonal variance decomposition for the two variable system considered above. First, we assume that the decomposition matrix *A* is positive definite. Then setting the derivative of *tr*(*A*), the trace of *A*, to zero leads us to $A = \Sigma^{1/2}$. The proof is completed by showing that the second derivative of *tr*(*A*) at the turning point is negative.

Now, since $tr(A) = a_{11} + a_{22}$, differentiating *tr*(*A*) with respect to a_{12} gives

(8) $$d\,tr(A) = da_{11} + da_{22}.$$

Since Σ is constant, differentiate (5) – (7) with respect to a_{12} yields

(9) $$2a_{11} \cdot da_{11} + 2a_{12} = 0$$

(10) $$2a_{22} \cdot da_{22} + 2a_{21} \cdot da_{21} = 0$$

(11) $$a_{11} \cdot da_{21} + da_{11} \cdot a_{21} + a_{22} + a_{12} \cdot da_{22} = 0$$

From (9), we have

(12) $$da_{11} = -a_{11}^{-1} a_{12}.$$

From (10), the derivative of a_{22} can be written as

(13) $$da_{22} = -a_{22}^{-1}a_{21} \cdot da_{21}.$$

Note that *A* is positive definite, so a_{11} and a_{22} are positive. Substitute the results of (12) and (13) into (11), we have

(14) $$da_{21} = \frac{a_{11}^{-1}(a_{12}a_{21} - a_{11}a_{22})}{a_{22}^{-1}(a_{11}a_{22} - a_{12}a_{21})} = -a_{11}^{-1}a_{22},$$

since, by virtue of positive definiteness of *A*, $a_{11}a_{22} - a_{12}a_{21} > 0$. The result in (14) above enables us to obtain da_{22} in (13) in terms of a_{11} and a_{21}, which is

(15) $$da_{22} = a_{11}^{-1}a_{21}.$$

Now substituting (12) and (15) into (8), we have at a turning point when $d\,tr(A) = 0$,

$$-a_{11}^{-1}a_{12} + a_{11}^{-1}a_{21} = 0,$$

which implies $a_{12} = a_{21}$. That is, $AA' = A^2 = \Sigma$. Since both *A* and Σ are real, symmetric and positive definite, $A = \Sigma^{1/2}$ is unique.

To prove that $A = \Sigma^{1/2}$ is a maximum turning point, we need to show that

$$d^2tr(A) = d^2a_{11} + d^2a_{22}$$

is negative when $a_{12} = a_{21}$. First differentiate (12) with respect to a_{12}, we have

$$d^2a_{11} = -a_{11}^{-1} + a_{11}^{-2}a_{12} \cdot da_{11} = -a_{11}^{-1} - a_{11}^{-3}a_{12}^2 < 0.$$

Since $a_{21} = a_{12}$, differentiate (13) with respect to a_{12} yields

$$d^2a_{22} = -a_{22}^{-1} + a_{22}^{-2}a_{12} \cdot da_{22} = -a_{22}^{-1} - a_{22}^{-3}a_{12}^2 < 0.$$

REMARK 1: Maximizing the trace of *A* leads us to the symmetrical restriction, which is the required additional equation to identify the four unknowns in (5) – (7).

REMARK 2: The symmetrical restriction $a_{21} = a_{12}$ is not the same as restricting the γ_{12} and γ_{21} in (1) and (2) to be the same. If we assume that $\text{var}(\varepsilon_i) = \text{var}(e_i)$, then the structural decomposition can be expressed as $\Sigma^{1/2}\Lambda^{-1/2}\varepsilon = e$, where Λ is a diagonal matrix with *i*-th diagonal entry equals to $\text{var}(e_i)$.

REMARK 3: In measuring the variance contribution of each component, it makes no difference which form of decomposition, $\Sigma^{1/2} z = e$ or $\Sigma^{1/2}\Lambda^{-1/2} \cdot \varepsilon = e$, is used.

3 THE GENERAL *n*-VARIABLE CASE

3.1 Orthogonal Decomposition for the $n \times n$ Case

The results in the above two-variable case can be generalized into the general multivariate *n*-vector case. Generally speaking, since Σ is symmetric, it has $(n^2 + n)/2$ distinct elements. However, *A* has n^2 unknown parameters to be identified. To determine the system, it is thus necessary to impose $(n^2 - n)/2$ restrictions. Similar to the two-variable case, maximizing the trace of an $n \times n$ positive definite matrix *A* leads us to the same symmetrical restriction for *A*, which provides the additional $(n^2 - n)/2$ restrictions. We shall now formally state the results in Theorem 1 below.

THEOREM 1: *Let e be a serially uncorrelated random n-vector with* $\mathrm{var}(e) = \Sigma$, *which is symmetrical and positive definite. Let A be a positive definite matrix such that* $Az = e$, $AA' = \Sigma$ *and* $\mathrm{var}(z) = I$. *Then maximizing the trace of A lead us to* $A = \Sigma^{1/2}$, *which is the unique positive definite square root of* Σ.

PROOF: The proof comprises three steps. First, we demonstrate that a symmetrical *A* gives rise to a zero derivative of *tr*(*A*). It is then shown that the second derivative of *tr*(*A*) is negative when *A* is symmetrical. Finally, we prove that $A = \Sigma^{1/2}$ is the global maximum point. Throughout the proof, the identity of $AA' = \Sigma$ is repeatedly used.

To prove that symmetry of *A* implies a zero derivative of *tr*(*A*), we differentiate $AA' = \Sigma$ and obtain

(16) $$A \cdot dA' + dA \cdot A' = 0.$$

Since *A* is positive definite, A^{-1} exists. Left multiply (16) by A^{-1} and take trace,

(17) $$tr(dA') = -tr(A^{-1} \cdot dA \cdot A').$$

Since $tr(dA) = tr(dA')$ and $tr(AB) = tr(BA)$, (17) can be rewritten as

(18) $$tr(dA) = -tr(dA \cdot A'A^{-1}).$$

Now $A = A'$, so $A'A^{-1} = I$. Therefore (18) can hold only when $d\,tr(A) = tr(dA) = 0$.

To show that it is a maximum turning point, we need to prove that $d^2\,tr(A) < 0$ when A is symmetrical. The second order differential of $A^2 = \Sigma$ can be written as

(19) $$A \cdot d^2A + d^2A \cdot A + 2(dA)^2 = 0 .$$

Left multiply (19) by A^{-1} and take trace,

(20) $$tr(d^2A) + tr(A^{-1} \cdot d^2A \cdot A) + 2tr\left(A^{-1}(dA)^2\right) = 0 .$$

Because $tr(A^{-1} \cdot d^2A \cdot A) = tr(d^2A)$ and positive definiteness of A implies that $A^{-1/2}$ exists, we can write (20) as

$$d^2tr(A) = tr(d^2A) = -tr\left(A^{-1}(dA)^2\right) = -tr(dA \cdot A^{-1/2}A^{-1/2} \cdot dA),$$

which, unless A is a null matrix, is negative.

We have demonstrated in the above that $A = \Sigma^{1/2}$ is a maximum turning point. To verify that it is a global maximum, let $A = \Sigma^{1/2} + D$ where D is any arbitrary matrix. The condition $AA' = \Sigma$ implies that

(21) $$\Sigma^{1/2}D' + D\Sigma^{1/2} + DD' = 0 .$$

Left multiply (21) by $\Sigma^{-1/2}$ and take trace,

(22) $$2\,tr(D) + tr(\Sigma^{-1/2}DD') = 0 .$$

But $tr(\Sigma^{-1/2}DD') = tr(D'\Sigma^{-1/4} \cdot \Sigma^{-1/4}D) \geq 0$ with equality if and only if $D = 0$, in which case, $A = \Sigma^{1/2}$. For the other case of $tr(\Sigma^{-1/2}DD') > 0$, (22) implies that $tr(D) < 0$. Since $A = \Sigma^{1/2} + D$, we have $tr(\Sigma^{1/2}) > tr(A)$.

Q.E.D.

3.2 Asymptotic Distribution of $\hat{\Sigma}^{1/2}$

Here we derive the asymptotic distribution of $\hat{\Sigma}^{1/2}$ when it is estimated from a real-valued, n-vector sample of $(e_1,\ldots,e_T)$. We assume that e_t is distributed IID $N(0,\Sigma)$, where Σ is a positive definite symmetric matrix. $\hat{\Sigma}^{1/2}$ is estimated by taking square root of $\hat{\Sigma}$, which is given by $\hat{\Sigma} = T^{-1}\sum_{t=1}^{T}(e_t - \bar{e})(e_t - \bar{e})'$ where $\bar{e}$ is the sample mean of e_t.

By the Central Limit Theorem, it is established that $\hat{\Sigma}$ has an asymptotic normal distribution given by

(23) $$\sqrt{T}\left(\text{vech}\hat{\Sigma} - \text{vech}\Sigma\right) \xrightarrow{d} N(0, V),$$

where vech is an operator that stacks distinct elements of a symmetric matrix into a vector (the stacking rules given by Magnus and Neudecker (1999, p. 49) are adopted here). Let D_n be the duplication matrix such that $D_n \text{vech}(\Sigma) = \text{vec}(\Sigma)$, and D_n^+ be the Moore-Penrose inverse of D_n that reverses the operation, that is, $D_n^+ \text{vec}(\Sigma) = \text{vech}(\Sigma)$. The variance covariance matrix in (23) can be written as

(24) $$V = 2D_n^+ (\Sigma \otimes \Sigma)(D_n^+)'.$$

Now, we state the second result of this article and provide its proof below.

THEOREM 2: *Given a real-valued, n-vector IID Gaussian sample of* $(e_1, \ldots, e_T)$ *with zero mean and* $\text{var}(e_t) = \Sigma$, *a positive definite symmetric matrix. The estimator of* $\Sigma^{1/2}$ *obtained by taking square root of the maximum likelihood estimator,* $\hat{\Sigma}$, *is asymptotically distributed as*

(25) $$\sqrt{T}\left(\text{vech}\hat{\Sigma}^{1/2} - \text{vech}\Sigma^{1/2}\right) \xrightarrow{d} N(0, V_{1/2}),$$

where

(26) $$V_{1/2} = 2^{-1}\left(D_n'(I \otimes \Sigma^{1/2})D_n\right)^{-1} D_n'(\Sigma \otimes \Sigma)D_n\left(D_n'(I \otimes \Sigma^{1/2})D_n\right)^{-1}.$$

PROOF: Since $\Sigma^{1/2}$ is a continuous matrix function of Σ, applying the delta method to (23) yields the result of (25). So we just need to prove (26); derive the Jacobian matrix of $\Sigma^{1/2}$, and obtain the required variance covariance matrix. We begin this by considering the differential of $\Sigma^{1/2} \cdot \Sigma^{1/2} = \Sigma$, which is given by

(27) $$\Sigma^{1/2} \cdot d\Sigma^{1/2} + d\Sigma^{1/2} \cdot \Sigma^{1/2} = d\Sigma.$$

Applying the vec operator to (27), we have

(28) $$(I \otimes \Sigma^{1/2}) \cdot \text{vec}(d\Sigma^{1/2}) + (\Sigma^{1/2} \otimes I) \cdot \text{vec}(d\Sigma^{1/2}) = \text{vec}(d\Sigma).$$

Let K_{nn} be the permutation matrix such the term $\Sigma^{1/2} \otimes I$ in (28) can be written as $K_{nn}(I \otimes \Sigma^{1/2})K_{nn}$. Due to symmetry of $\Sigma^{1/2}$, $K_{nn}\text{vec}(d\Sigma^{1/2}) = d\,\text{vec}\Sigma^{1/2}$. Thus we have from (28),

(29) $$(I + K_{nn})(I \otimes \Sigma^{1/2}) \cdot d\,\text{vec}\Sigma^{1/2} = d\,\text{vec}\Sigma.$$

Now, $I + K_{nn} = 2D_n D_n^+$, where the Moore-Penrose inverse is $D_n^+ = (D_n' D_n)^{-1} D_n'$. Left multiply (29) by D_n' and simplify the duplication matrices,

(30) $$2D_n'(I \otimes \Sigma^{1/2})D_n \cdot d\,\text{vech}\Sigma^{1/2} = D_n' D_n \cdot d\,\text{vech}\Sigma.$$

Note that the dimension of D_n is $n^2 \times n(n+1)/2$ and has full column rank, so the inverse of $(D_n'(I \otimes \Sigma^{1/2})D_n)$ exists. Therefore, the required differential, $d\,\text{vech}\Sigma^{1/2}$, in (30) can be expressed as

$$d\,\text{vech}\Sigma^{1/2} = 2^{-1}\left(D_n'(I \otimes \Sigma^{1/2})D_n\right)^{-1} D_n' D_n \cdot d\,\text{vech}\Sigma,$$

which, by the identification theorem for matrix functions (Magnus and Neudecker (1999, p. 96)), yields the required Jacobian matrix

(31) $$\nabla\Sigma^{1/2} = 2^{-1}\left(D_n'(I \otimes \Sigma^{1/2})D_n\right)^{-1} D_n' D_n.$$

By the delta method, $V_{1/2} = \nabla\Sigma^{1/2} \cdot V \cdot (\nabla\Sigma^{1/2})'$. Substituting (31) for the Jacobian matrix and simplifying the duplication matrices, we arrive at (26) in Theorem 2.

Q.E.D.

REMARK 4: Theorem 2 also holds if e_t is non-Gaussian but distributed $\text{IID}(0,\Sigma)$ with zero fourth order cumulants.

4. NUMERICAL EXAMPLES

In this section, we provide three numerical examples based on the work of Campbell and Ammer (1993). They use a VAR model to decompose the excess stock return (e_{t+1}), excess 10-year bond return (b_{t+1}), and unexpected yield spread innovation (s_{t+1}) into changes in expectations of future stock dividend, inflation, short-term real interest rate, and changes in expectations of future excess stock and bond returns. The three variables of interest can be written as

$$e_{t+1} = e_{d,t+1} - e_{r,t+1} - e_{x,t+1},$$

$$b_{t+1} = -b_{\pi,t+1} - b_{r,t+1} - b_{x,t+1},$$

$$s_{t+1} = s_{\pi,t+1} + s_{r,t+1} + s_{x,t+1},$$

where the subscripts d, r, π and x stand for dividend, real interest rate, inflation and excess return respectively. So, for instance, $e_{d,t+1}$ can be interpreted as the news about future dividend for the unexpected excess stock return, $b_{x,t+1}$ refers to the news on future excess bond return, whereas $s_{\pi,t+1}$ is the news about future inflation for the unexpected yield spread innovation.

As different orderings in the Cholesky decomposition yield vastly different results, Campbell and Ammer report all 6 variance-covariance terms (which are standardized to sum to equal one) and R^2 statistics from simple regressions of the variable of interest on each of their corresponding components. These statistics are provided in Table I, II and III. In each table, we also provide the square root decomposition matrix ($\Sigma^{1/2}$), the associated variance contributions (*vc*), the Chi-squared statistics for testing equality of variance contributions, as well as the variance contributions obtained using different ordering choices in the Cholesky decomposition.

The Chi-squared tests are constructed as follows. First, let c_j be the sum of the *j*-th column elements of $\Sigma^{1/2}$. Since the variance of variable of interest is standardized, the orthogonal variance contribution from component *j* is simply given by $vc_j = c_j^2$. To carry out the hypothesis testing of $vc_j = vc_k$, $j \neq k$, we make use of the fact that equality implies $c_j = \pm c_k$ and calculate

$$Csq = \begin{cases} (c_j - c_k)^2 \big/ \mathrm{var}(c_j - c_k) & \text{if } c_j \cdot c_k \geq 0, \\ (c_j + c_k)^2 \big/ \mathrm{var}(c_j + c_k) & \text{if } c_j \cdot c_k < 0. \end{cases} \tag{32}$$

The variance terms in (32) are calculated using (26) in Theorem 2 with $T = 442$.[4] Under the null hypothesis that the two variance contributions are equal, *Csq* is distributed as Chi-squared with one degree of freedom.

< Insert Table I >

From Table I, we can see that correlations between the three components that explain excess stock return are low. As a result, raw variances in Σ , R^2 statistics and variance contributions (*vc*) are of consistent magnitudes. Except for the first and third selections of ordering, Cholesky method yields relatively similar results too. Tests on equality of square root variance contributions reveal that they are significantly different from each other.

< Insert Table II >

Next, we look at the variance decomposition for excess bond return given in Table II. It can be seen that both the inflation (b_π) and excess return (b_x) components have large variances, and that correlations between the three components are fairly high. Except for the Cholesky decomposition, the first three measures of variance contributions are consistent with each other. Though *vc* suggests b_x has the largest variance contribution whereas the raw variance in Σ suggests b_π contributes the most, the Chi-squared test reveals that the difference is insignificant. The test, however, confirms that the real interest component (b_r) has the least contribution to the variation in excess bond return. For the Cholesky decomposition, variation in the measures of contribution is huge for different orderings. For example, selecting the second ordering yields 0.207 and 0.775 for b_π and b_x respectively, whereas the fourth selection choice yields hugely contrasting results of 0.754 and 0.174 for b_π and b_x respectively.

< Insert Table III >

Table III provides the variance decomposition for unexpected yield spread innovation. First, it is noticeable that both the inflation (s_π) and real interest rate (s_r) components have large variances and are highly negatively correlated (correlation equals -0.929). Despite their large variances, simple regressions of unexpected yield spread innovation

[4] Campbell and Ammer (1993) use monthly data from January 1952 to February 1987, a total of 442 months.

on each of both components yield very low R^2 statistics of 0.072 and 0.003. Square root decomposition reveals that the contributions of s_π, s_r and s_x to the total variation are 0.434, 0.277 and 0.289 respectively. While these figures are more sensible than the R^2 statistics, one point merits further discussions. That is, in spite of the fact that $\text{var}(s_r) \approx \text{var}(s_\pi)$ and that $\text{var}(s_r)$ is much larger than $\text{var}(s_x)$, vc_r (vc of s_r) is the smallest at 0.277. However, a more careful analysis reveals that it is a plausible outcome. First, as noted above, correlation between s_r and s_π is highly negative. Second, $\text{cov}(s_r, s_x)$ is negative whereas $\text{cov}(s_\pi, s_x)$ is positive. The resulting $\Sigma^{1/2}$ implies that while the associated unit-variance uncorrelated structural shock, z_r, 'innovates' s_r with an impact coefficient as high as 1.778, this effect is greatly reduced by its opposite effect on s_π with a negative impact coefficient of -1.225. Moreover, negative $\text{cov}(s_r, s_x)$ implies that the variance contribution by s_r is further reduced, albeit by a small negative impact coefficient of -0.027 on s_x.

Though $\text{var}(s_x)$ is low, less than 3% of $\text{var}(s_\pi) + \text{var}(s_r)$, its variance contribution according to the square root method is relatively high. This can be explained by the high negative correlation between s_π and s_x, which implies that their shocks are in opposite directions and the net effect becomes much smaller. Also, it is noted that the correlations between s_x and the other two components are small. Indeed, $vc_x = 0.289$ is consistent with the corresponding relatively high R^2 statistic of 0.325. The Chi-squared tests lend further credibility to the square root decomposition approach. Though vc_π seems considerably larger than vc_r, the test reveals an insignificant difference, an outcome consistent with their similar-size variances and large correlation. The test between vc_π and vc_x, however, confirms that s_π explains variation in the yield spread innovation significantly more than s_x.

Unlike the results in Table I, the Cholesky method generates vastly different figures of variance contributions if different selections of ordering are used. This is due to the fact that the components in Table I are relatively uncorrelated. Finally, we remark that because the numerical examples considered have only three variables, above analyses based on variance covariance matrix are feasible. However, in practice, we often have more than three variables to analyze. Indeed, n variables imply that there are $n(n+1)/2$ variance covariance terms to consider. Worse still, for the Cholesky decomposition, there are $n!$ selections of ordering.

4. CONCLUSION WITH SOME REMARKS

This article proposes an orthogonal variance decomposition that maximizes the trace of a positive definite decomposition matrix. When there is no economic theory one can rely upon to decompose the shocks, the trace of A has meaningful interpretation: a larger trace means a higher association between observable shocks and their corresponding structural shocks. It turns out that such a decomposition matrix is unique and equals to the square root of the variance covariance matrix. Limiting distribution of the estimator of square root decomposition matrix is derived, and numerical examples are provided to illustrate its usefulness.

Though this article considers the simple IID real-valued n-vector case, its results can be readily extended to, for example, the VARMA case of Mittnik and Zadrozny (1993). From the numerical examples, we can see that a different selection of ordering in the Cholesky method could yield a vastly different outcome. Therefore, the square root decomposition is a useful alternative for comparison. In particular, when there are many variables in the system, the proposed method is able to provide a concise analysis. Finally, since multiple regression and VAR models are ubiquitous in most social science studies, this article proposes a simple means of variance decomposition that is intuitive and requires no prior information.

REFERENCE

BERNANKE, B. (1986): “Alternative explanations of money-income correlation,” *Carnegie-Rochester Conference Series on Public Policy*, 25, 49 – 100.

BLANCHARD, O. J., AND D. QUAH (1989): “The dynamic effects of aggregate demand and supply disturbances,” *American Economic Review*, 79, 655 – 673.

CAMPBELL, J. Y., AND J. AMMER (1993): “What moves the stock and bond markets? A variance decomposition for long-term asset returns,” *Journal of Finance*, 48, 3 – 37.

HAMILTON, J. D. (1994): *Time Series Analysis*, New Jersey: Princeton.

MAGNUS, J. R., AND H. NEUDECKER (1999): *Matrix differential calculus with applications in statistics and econometrics*. New York: John Wiley.

MITTNIK, S., AND P. A. ZADROZNY (1993): "Asymptotic distributions of impulse responses, step responses, and variance decompositions of estimated linear dynamic models," *Econometrica*, 61, 857 – 870.

SIMS, C. (1986): "Are forecasting models usable for policy analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 3 – 16.

SHORROCKS, A. F. (1982): "Inequality decomposition by factor components," *Econometrica*, 50, 193 – 211.

TABLE I

VARIANCE DECOMPOSITION FOR EXCESS STOCK RETURNS

		e_d	$-e_r$	$-e_x$
	Panel A			
Σ	e_d	0.146	-0.007	0.036
	$-e_r$	***-0.172***	0.013	0.040
	$-e_x$	***0.112***	***0.413***	0.705
Simple regression	R^2	0.209	0.160	0.864
	Panel B			
$\Sigma^{1/2}$	e_d	0.380	-0.018	0.030
	$-e_r$	***195.12***	0.104	0.043
	$-e_x$	***251.48***	***577.29***	0.838
$\Sigma^{1/2}$ decomposition	vc	0.154	0.017	0.829
	Panel C			
Variance decomposition using Cholesky Method	$d \to r \to x$	0.209	0.231	0.561
	$d \to x \to r$	0.209	0.010	0.781
	$r \to d \to x$	0.284	0.156	0.561
	$r \to x \to d$	0.136	0.156	0.708
	$x \to d \to r$	0.126	0.010	0.864
	$x \to r \to d$	0.136	0.000	0.864

Excess stock returns are decomposed into e_d, $-e_r$ and $-e_x$, which are respectively news about future dividends, real interest rates and excess stock returns. Data in Panel A is obtained from Campbell and Ammer (1993, Table III). The figures are variance covariance matrix (correlations in bold italic fonts), and R^2 statistics obtained from simple regressions of excess stock returns on each component. In Panel B, $\Sigma^{1/2}$ is the square root decomposition matrix, *vc* is variance contribution, and test statistics for equality of *vc*'s are in bold italic fonts. Under the null hypothesis, the test statistic is distributed as Chi-squared with 1 degree of freedom; 3.841 and 6.635 are critical values at 5% and 1% significance levels respectively. Panel C provides the variance contributions using Cholesky method with different orderings.

TABLE II

VARIANCE DECOMPOSITION FOR EXCESS BOND RETURNS

		$-b_\pi$	$-b_r$	$-b_x$
	Panel A			
Σ	$-b_\pi$	1.084	-0.058	-0.552
	$-b_r$	***-0.367***	0.023	0.075
	$-b_x$	***-0.541***	***0.508***	0.962
Simple regression	R^2	0.207	0.072	0.245
	Panel B			
$\Sigma^{1/2}$	$-b_\pi$	1.001	-0.036	-0.284
	$-b_r$	***261.56***	0.134	0.061
	$-b_x$	***0.443***	***298.81***	0.937
$\Sigma^{1/2}$ decomposition	vc	0.464	0.025	0.510
	Panel C			
Variance decomposition using Cholesky Method	$\pi \rightarrow r \rightarrow x$	0.207	0.218	0.574
	$\pi \rightarrow x \rightarrow r$	0.207	0.017	0.775
	$r \rightarrow \pi \rightarrow x$	0.353	0.071	0.574
	$r \rightarrow x \rightarrow \pi$	0.754	0.071	0.174
	$x \rightarrow \pi \rightarrow r$	0.738	0.017	0.245
	$x \rightarrow r \rightarrow \pi$	0.754	0.000	0.245

Excess bond returns are decomposed into $-b_\pi$, $-b_r$ and $-b_x$, which are respectively news about future inflation, real interest rates and excess bond returns. Data from Panel A is obtained from Campbell and Ammer (1993, Table IV). Descriptions for figures in Panel A, B and C are similar to Table I. The 5% and 1% critical values of Chi-squared statistics under null hypothesis are 3.841 and 6.635 respectively.

TABLE III

VARIANCE DECOMPOSITION FOR YIELD SPREAD INNOVATIONS

		s_π	s_r	s_x
	Panel A			
Σ	s_π	4.864	-4.426	0.152
	s_r	***-0.929***	4.664	-0.124
	s_x	***0.133***	***-0.111***	0.267
Simple regression	R^2	0.072	0.003	0.325
	Panel B			
$\Sigma^{1/2}$	s_π	1.833	-1.225	0.051
	s_r	***1.570***	1.778	-0.027
	s_x	***9.744***	***0.074***	0.514
$\Sigma^{1/2}$ decomposition	vc	0.434	0.277	0.289
	Panel C			
Variance decomposition using Cholesky method	$\pi \to r \to x$	0.072	0.666	0.262
	$\pi \to x \to r$	0.072	0.637	0.292
	$r \to \pi \to x$	0.735	0.003	0.262
	$r \to x \to \pi$	0.660	0.003	0.337
	$x \to \pi \to r$	0.037	0.637	0.326
	$x \to r \to \pi$	0.660	0.014	0.326

Excess bond returns are decomposed into s_π, s_r and s_x, which are respectively news about future inflation, real interest rates and excess bond returns. Data from Panel A is obtained from Campbell and Ammer (1993, Table VIII). Descriptions for figures in Panel A, B and C are similar to Table I. The 5% and 1% critical values of Chi-squared statistics under null hypothesis are 3.841 and 6.635 respectively.