I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12920

VC - A Method For Estimating Time-Varying Coefficients in Linear Models

Ekkehart Schlicht

JANUARY 2020

DISCUSSION PAPER SERIES

IZA DP No. 12920

VC - A Method For Estimating Time-Varying Coefficients in Linear Models

Ekkehart Schlicht
University of Munich and IZA

JANUARY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

VC - A Method For Estimating Time-Varying Coefficients in Linear Models

This paper describes a moments estimator for a standard state-space model with coefficients generated by a random walk. A penalized least squares estimation is linked to the GLS (Aitken) estimates of the corresponding linear model with time-invariant parameters. The VC estimates are moments estimates. They do not require the disturbances to be Gaussian, but if they are, the estimates are asymptotically equivalent to maximum likelihood estimates. In contrast to Kalman filtering, no specification of an initial state or an initial covariance matrix is required. While the Kalman filter is one sided, the VC filter is two sided and therefore uses more of the available information for estimating intermediate states.. Further, the VC filter has a clear descriptive interpretation.

JEL Classification: C2, C22, C32, C51, C52

Keywords: time-series analysis, linear model, state-space estimation, time-varying coefficients, moments estimation, Kalman filtering, penalized least squares

Corresponding author:
Ekkehart Schlicht
Department of Economics
University of Munich
Hurtenstrasse 13
82346 Andechs
Germany
E-mail: schlicht@lmu.de

1 Introduction

This paper describes and discusses an estimator for a linear time series model with time-varying coefficients. Such a model, the variable coefficients model, or "VC model" for short, generalizes the standard linear model. The standard model assumes that the coefficients giving the influence of the independent variables on the dependent variable remain constant. In the VC model, these coefficients are permitted to change over time.

The VC model has been initially proposed for dealing empirically with economic theories that are subject to a *ceteris paribus* clause (Schlicht, 1977, ch.4). Schlicht (1989) has proposed an estimation method – the VC method – which has been embodied in some freely available software packages (Schlicht 2005b, 2005c Ludsteck 2004; 2018). Some simulations in Schlicht and Ludsteck (2006) have shown that the VC is preferable for studying the specific class of models for which it was designed.

In the meanwhile, VC has found a number of applications in various settings, mainly dealing with structural change, such as the recent decoupling of growth and pollution in the wake of global warming, the changes occurring in financial markets after the financial crisis of 2008, drifts in Okun's Law over time, and more. The references to contributions that have employed VC given at the end of the paper list some of these studies.

The following sections introduce the model and describe the "criteria" or "penalty" approach that permits to estimate the time-paths of the coefficients in a purely descriptive way (Sections 2 to 4). Based on that, a moments estimator will be proposed (Sections 5 to 8). If it is assumed additionally that the disturbances are normally distributed, a maximum likelihood estimator can be given (Sections 9 and 10). It is shown that this estimator coincides with the moments estimator for sufficiently long time series (Section 11).

2 The Linear Theoretical Model and its Empirical Application

Consider a theory stating that the dependent variable y as a linear function of some independent variables $x_1, x_2, \dots, x_n$:

$$y = a_1 x_1 + a_2 x_2 + \dots + a_n x_n \tag{2.1}$$

The coefficients $a_1, a_2, \dots, a_n$ give the influence of the independent variables.

If we have T observations $y_t, x_{1,t}, x_{2,t}, \dots x_{n,t}$ with $t = 1, 2, \dots T$ denoting the time of an observation, we can try to estimate the theoretical coefficients $a_1, a_2, \dots, a_n$ by a standard linear regression. In order to do that, we have to add an error term u_t to capture discrepancies of the empirical from the theoretical regularity due to measurement errors *etc.* and obtain

$$y_t = a_1 x_{1,t} + a_2 x_{2,t} + \dots + a_n x_{n,t} + u_t, \quad t = 1, 2, \dots T. \tag{2.2}$$

It appears, however, improbable, that outside influences not captured in the theoretical model (and theoretically held constant under a *ceteris paribus* clause) affect only the disturbance term, and not the coefficients themselves – think of changes in technology, preferences, market structure, and the composition of aggregates over time. These outside influences may affect the coefficients themselves, and they might change over time.

The problem of possibly time-varying coefficients was the subject of the famous Keynes-Tinbergen controversy around 1940.[1] While Tinbergen (1940, p. 153) defended the use of regression analysis with the argument that in "many cases only small changes in structure will occur in the near future", Keynes (1973, p. 294) objected that "the method requires not too short a series whereas it is only in a short series, in most cases, that there is a reasonable expectation that the coefficients will be fairly constant."

It appears that both arguments are correct. The VC model takes care of both

[1] See Tinbergen (1940), Keynes (1939), Keynes (1973, pp. 285–321).

by assuming that the coefficients change only *slowly* over time: They are highly auto-correlated. This is formalized by a random walk (Cooley and Prescott 1973, Schlicht 1973, Athans 1974). If $a_{i,t}$ denotes the state of coefficient a_i at time t, it is assumed that

$$a_{i,t+1} = a_{i,t} + v_{i,t} \tag{2.3}$$

with the disturbance term $v_{i,t}$ of expectation zero and with variance σ_i^2. The assumption of expectation zero formalizes the idea that "the coefficients will be fairly constant" in the short run, while the variance σ_i^2 is a measure of the stability of coefficient i and is to be estimated. For $\sigma_i^2 = 0$ for some i, the case of a constant (time-invariant) coefficients is covered as well. As a consequence, the standard linear model is replaced by

$$\begin{aligned} y_t &= a_{1,t}x_{1,t} + a_{2,t}x_{2,t} + \ldots + a_{n,t}x_{n,t} + u_t \\ &\qquad E\left\{u_t\right\} = 0, \quad E\left\{u_t^2\right\} = \sigma^2 \qquad (2.4) \\ a_{i,t+1} &= a_{i,t} + v_{i,t}, \\ &\qquad E\left\{v_{i,t}\right\} = 0, \quad E\left\{v_{i,t}^2\right\} = \sigma_i^2 \qquad (2.5) \end{aligned}$$

The VC method estimates the expected time-paths of the coefficients. It can be viewed as a straightforward generalization of the method of least squares:

- While the method of ordinary least squares selects estimates that minimize the sum of squared disturbances $\sum_{t=1}^{T} u_t^2$ in the equation, VC selects estimates that minimize the sum of squared disturbances in the equation and a weighted sum of squared disturbances in the coefficients $\sum_{t=1}^{T} u_t^2 + \gamma_1 \sum_{t=2}^{T} v_{1,t}^2 + \gamma_2 \sum_{t=2}^{T} v_{2,t}^2 + \ldots + \gamma_n \sum_{t=2}^{T} v_{n,t}^2$, where the weights for the changes in the coefficients $\gamma_1, \gamma_2, \ldots, \gamma_n$ are determined by the inverse variance ratios, *i.e.* $\gamma_i = \sigma^2/\sigma_i^2$. In other words, it balances the desiderata of a good fit and parameter stability over time.
- Estimation can proceed by focusing on some selected coefficients and keeping the remaining coefficients constant over time. This is done by

keeping the corresponding variances σ_i^2 close to zero, rather than estimating them. (If all coefficients are frozen in this manner, the OLS result is obtained.)

- The time-averages of the regression coefficients are GLS estimates of the corresponding regression with fixed coefficients, *i.e.* $\frac{1}{T}\sum_t a_t = a_{GLS}$.
- The VC method does not require initial values for the initial state and the initial variances. Rather all states and variances are estimated in an integrated unified procedure. This is an advantage over Kalman filtering and of some importance for shorter time series.
- All estimates are moments estimates. It is not necessary to presuppose Gaussian disturbances.
- For increasing sample sizes T and under the assumption that all disturbances are normally distributed, the moments estimates approach the maximum likelihood estimates.

3 Notation and Basic Assumptions

All vectors are conceived as column vectors, and their transposes are indicated by an apostrophe. The observations at time t are $x_t' = (x_{1,t}, x_{2,t}, ..., x_{n,t})$ and y_t for $t = 1, 2, .. , T$. We write

$$y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_T \end{pmatrix}, \quad x = \begin{pmatrix} x_1' \\ x_2' \\ . \\ . \\ x_T' \end{pmatrix}, \quad X = \begin{pmatrix} x_1' & & & & 0 \\ & x_2' & & & \\ & & . & & \\ & & & . & \\ 0 & & & & x_T' \end{pmatrix}$$

$$\text{order} \quad T \qquad\qquad T \times n \qquad\qquad T \times Tn$$

$$a_t = \begin{pmatrix} a_{1,t} \\ a_{2,t} \\ . \\ . \\ a_{n,t} \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_T \end{pmatrix}, \quad v_t = \begin{pmatrix} v_{1,t} \\ v_{2,t} \\ . \\ . \\ v_{n,t} \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ . \\ . \\ v_{T-1} \end{pmatrix}$$

$$\text{order} \quad n \qquad\qquad Tn \qquad\qquad n \qquad\qquad (T-1)\,n$$

We write further

$$\Sigma = \operatorname{diag}\begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ . \\ . \\ \sigma_n^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 & & & 0 \\ 0 & \sigma_2^2 & & & \\ & & . & & \\ & & & . & 0 \\ 0 & & & 0 & \sigma_n^2 \end{pmatrix}$$

$$\text{order} \quad n \qquad\qquad\qquad n \times n$$

and define

$$p = \begin{pmatrix} -1 & 1 & 0 & & & 0 \\ 0 & -1 & 1 & 0 & & \\ & & . & . & & \\ & & & . & . & 0 \\ 0 & & & 0 & -1 & 1 \end{pmatrix}, \quad P = p \otimes I_n = \begin{pmatrix} -I_n & I_n & & & & 0 \\ & -I_n & I_n & & & \\ & & . & . & & \\ & & & . & . & \\ 0 & & & & -I_n & I_n \end{pmatrix}$$

$$\text{order} \quad (T-1) \times T \qquad\qquad\qquad (T-1)\,n \times Tn$$

with I_n denoting the identity matrix of order n.

The model is obtained by writing equations (2.4) and (2.5) in matrix form:

The model

$$y = Xa + u, \quad E\{u\} = 0, \quad E\{uu'\} = \sigma^2 I_T \tag{3.1}$$

$$Pa = v, \quad E\{v\} = 0, \quad E\{vv'\} = V = I_{T-1} \otimes \Sigma \tag{3.2}$$

Note that the explanatory variables X are taken as predetermined, rather than stochastic.

Regarding the observations X and y we assume that a perfect fit of the model to the data is not possible:

This assumption rules out the (trivial) case that the standard linear model (2.2) fits the empirical data perfectly, a case that cannot reasonably be expected to occur in practical applications. Further, the assumption implies that the number of observations exceeds the number of coefficients to be estimated:

$$T > n. \tag{3.3}$$

Assumption ("No Perfect Fit").

$$Pa = 0 \quad \text{implies} \quad y \neq Xa. \tag{3.4}$$

4 Least Squares

In a descriptive spirit, the time-paths of the coefficients can be determined by following the penalized least square approach, where some criteria are employed that formalize some descriptive desiderata.[2] In the case at hand, the desiderata are that the model fits the data well and that the coefficients change only slowly over

[2] For the penalized least squares approach, see Green and Silverman (2000). The approach was introduced by Whittaker (1923), Henderson (1924) and Leser (1961). It has been used also by Hodrick and Prescott (1997), and has been further developed by Leser (1963), Schlicht (1981), Schlicht and Pauly (1983), Schlicht (1984) and Schlicht (2005a).

time – u and v ought to be as small as possible. The sum of the squared errors $u'u$ is taken as a criterion for the goodness of fit of equation (3.1), the weighted sum of the squared changes of the coefficients $v_i'v_i$ over time give criteria for the stability of the coefficients over time. The combination of all these criteria gives an overall criterion that combines the desiderata of a good fit and stability of coefficients over time. The weights $(\gamma_1, \gamma_2, .., \gamma_n)$ give the relative importance of the stability of the coefficients over time, where weight γ_i relates to coefficient a_i. For the time being, these weights are taken as given but will later be estimated, too.

Write

$$\Gamma = \begin{pmatrix} \gamma_1 & 0 & . & & 0 \\ 0 & \gamma_2 & 0 & . & \\ . & 0 & . & . & \\ & & . & . & 0 \\ 0 & & & 0 & \gamma_n \end{pmatrix} \tag{4.1}$$

and

$$G = I_{T-1} \otimes \Gamma. \tag{4.2}$$

Adding the sum of squares $u'u$ and the weighted sum of squares $v'Gv$ gives the overall criterion

$$Q = u'u + v'Gv \tag{4.3}$$

This expression is to be minimized under the constraints given by the model (3.1), (3.2) with the observations X and y

$$u = y - Xa \tag{4.4}$$

$$v = Pa. \tag{4.5}$$

This determines the time-paths of the coefficients a that optimize this criterion.

Hence we can write

$$Q = (y - Xa)'(y - Xa) + a'P'GPa \tag{4.6}$$

The weighted sum of squares Q is the sum of two positive semi-definite quadratic forms. The "no perfect fit" assumption (3.4) rules out the case that Q can be zero. Hence Q is positive definite and of full rank. The first order condition for a minimizing a is

$$\frac{\partial Q}{\partial a} = -2Xy + 2(X'X + P'GP)a = 0 \tag{4.7}$$

and the second order condition is that the Jacobian

$$\frac{\partial^2 Q}{\partial a \partial a'} = 2(X'X + P'GP) \tag{4.8}$$

be positive definite, which is the case. Solving (4.7) for a and plugging this into (4.4) and (4.5) gives the estimates

$$a_{LS} = (X'X + P'GP)^{-1} X'y \tag{4.9}$$

$$u_{LS} = \left(I_T - X(X'X + P'GP)^{-1} X'\right) y \tag{4.10}$$

$$v_{LS} = P(X'X + P'GP)^{-1} X'y \tag{4.11}$$

where the subscript LS stands for "least squares".

5 Orthogonal Parametrization

For purposes of estimation we need a model that explains the observation y as a function of the observations X and the random variables u and v. This would permit calculating the probability distribution of the observations y contingent on the parameters of the distributions of u and v, *viz.* σ^2 and Σ. The true model does not permit such an inference, though, because the matrix P is of rank $(T-1)n$ rather than of rank Tn and cannot be inverted. Hence v does not determine a unique a but rather the set of solutions

$$A := \left\{ a = \tilde{P}v + Z\beta \middle| \beta \in \mathbb{R}^n \right\}. \tag{5.1}$$

with β as a shift parameter and

$$\tilde{P} \quad := \quad P' \left(PP'\right)^{-1} \tag{5.2}$$

of order $Tn \times (T-1)\,n$ as the right-hand pseudo-inverse of P. For any v we have $a \in A \Leftrightarrow Pa = v$. Hence equation (3.1) and the set (5.1) give equivalent descriptions of the relationship between a and v.

Define further the $Tn \times n$ matrix

$$Z := \begin{pmatrix} I_n \\ I_n \\ \cdot \\ I_n \end{pmatrix}. \tag{5.3}$$

It is orthogonal to P:

$$PZ \quad = \quad 0$$

and the square matrix (P', Z) is of full rank. Note further that

$$Z'Z = T \cdot I_n, \;\; P' \left(PP'\right)^{-1} P + ZZ' = I_{Tn}. \tag{5.4}$$

The last equality is implied by the identity

$$\begin{pmatrix} P' & Z \end{pmatrix} \left(\begin{pmatrix} P \\ Z' \end{pmatrix} \begin{pmatrix} P' & Z \end{pmatrix} \right)^{-1} \begin{pmatrix} P \\ Z' \end{pmatrix} = I_{Tn}.$$

Regarding the matrices P, $\tilde{P}$, and Z we have

$$\begin{array}{rcccc} P\tilde{P} & = & \tilde{P}'P' & = & I_{(T-1)n} \\ \tilde{P}P & = & P'\tilde{P}' & = & I_{Tn} - ZZ' \\ Z'\tilde{P} & = & \tilde{P}'Z & = & 0. \end{array} \tag{5.5}$$

In view of (5.1), any solution a to $Pa = v$ can be written as

$$a = \tilde{P}v + Z\beta \tag{5.6}$$

for some $\beta \in \mathbb{R}^n$. Equation (3.1) can be re-written as

$$y = u + X\tilde{P}v + XZ\beta. \tag{5.7}$$

The model (5.6), (5.7) will be referred to as the *equivalent orthogonally parameterized model*. It implies the *true model* (3.1), (3.2). It implies, in particular, that a_t is a random walk even though a_t depends, according to (5.6), on past *and* future realizations of v_t.

The formal parameter β has a straightforward interpretation. Pre-multiplying (5.6) by Z' gives

$$Z'a = Z'Z\beta = T\beta$$

and therefore

$$\beta = \frac{1}{T}\sum_{t=1}^{T} a_t.$$

Hence β gives the averages of the coefficients $a_{i,t}$ over time.

Equation (5.7) permits calculating the density of y dependent upon the parameters of the distributions of u and v and the formal parameters β. In a second step, all these parameters – σ^2, Σ, and β – can be determined by moments estimators that will be derived in Section 8.

The orthogonal parametrization, proposed in Schlicht (1985, Sec. 4.3.3), entails some advantages with respect to symmetry and mathematical transparency, as compared to more usual procedures, such as parametrization by initial values. It permits to derive our moments estimator that does not require normally distributed disturbances, and to write down an explicit likelihood function for the case of normally distributed disturbances that permits estimation of all relevant parameters in a unified one-shot procedure.

The formal parameter vector β relates directly to the coefficient estimates of a standard generalized least squares (GLS, Aitken) regression. Equation (5.7) can be interpreted as a standard regression for this parameter vector with the matrix $x = XZ$ giving the explanatory variables:

$$y = x\beta + w \tag{5.8}$$

and the disturbance

$$w = X\tilde{P}v + u. \tag{5.9}$$

It has expectation zero

$$E\{w\} = 0 \tag{5.10}$$

and covariance

$$E\{ww'\} = X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T = W. \tag{5.11}$$

The Aitken estimate β_A satisfies

$$x'W^{-1}(y - x\beta_A) = 0 \tag{5.12}$$

or

$$\beta_A = \left(x'W^{-1}x\right)^{-1} x'W^{-1}y. \tag{5.13}$$

where the subscript A stands for "Aitken". As $x = XZ$ and $W = X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T$, equations (5.12) and (5.13) can be written as

$$Z'X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) = 0 \tag{5.14}$$

and equation (5.13) gives rise to

$$\beta_A = \left(Z'X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)^{-1} XZ\right)^{-1} Z'X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)^{-1} y. \tag{5.15}$$

6 The Filter

This section derives the VC filter which gives the expectation of the coefficients a for given observations X and y, a given shift parameter β, and given variances σ^2 and Σ.

For given β and X, the vectors y and a can be viewed as realizations of random variables determined jointly by the system (5.6), (5.8) as brought about by the disturbances u and v:

$$\begin{pmatrix} a \\ y \end{pmatrix} = \begin{pmatrix} Z \\ XZ \end{pmatrix}\beta + \begin{pmatrix} \tilde{P} & 0 \\ X\tilde{P} & I_T \end{pmatrix}\begin{pmatrix} v \\ u \end{pmatrix}$$

The covariance is

$$\begin{aligned} E\left\{\begin{pmatrix} a \\ y \end{pmatrix}\begin{pmatrix} a' & y \end{pmatrix}\right\} &= \begin{pmatrix} \tilde{P} & 0 \\ X\tilde{P} & I_T \end{pmatrix}\begin{pmatrix} V & 0 \\ 0 & \sigma^2 I_T \end{pmatrix}\begin{pmatrix} \tilde{P}' & \tilde{P}'X' \\ 0 & I_T \end{pmatrix} \\ &= \begin{pmatrix} \tilde{P}V\tilde{P}' & \tilde{P}V\tilde{P}'X' \\ X\tilde{P}V\tilde{P}' & X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T \end{pmatrix}. \end{aligned}$$

The marginal distribution of y is as given by (5.8) and (5.11). On this basis, we take our estimate of a as

$$a_A = Z\beta_A + \tilde{P}V\tilde{P}'X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A). \tag{6.1}$$

which is the expectation of a for the case that u and v are Gaussian and y, β, σ^2, and

Σ are given. (It will turn out later on that a_A is the expectation of a for non-Gaussian disturbances as well, see equation (7.10) below.)

Note that the variance-covariance matrix of w, as given in equation (5.11), tends to $\sigma^2 I_T$ if the the variances σ_i^2 go to zero, and equation (5.7) approaches the standard unweighted linear regression. In this sense, the OLS regression model is covered as a special limiting case by the model discussed here.

7 Least Squares and Aitken

The following theorem states that the least squares estimator a_{LS} and the Aitken estimator a_A coincide if the weights are given by the variance ratios.

Claim 1. $G = \sigma^2 V^{-1}$ implies $a_{LS} = a_A$.

Proof. Consider first the necessary conditions for a minimum of (4.3). The first-order condition (4.7) defines a_{LS} with weights $G = \sigma^2 V^{-1}$ uniquely and can be written as

$$\left(X'X + \sigma^2 P'V^{-1}P\right) a_{LS} = X'y \tag{7.1}$$

It will be shown that (6.1) implies

$$\left(X'X + \sigma^2 P'V^{-1}P\right) a_A = X'y \tag{7.2}$$

which will establish the proposition.

Pre-multiplication of (6.1) by $\left(X'X + \sigma^2 P'V^{-1}P\right)$ gives

$$\begin{aligned}\left(X'X + \sigma^2 P'V^{-1}P\right) a_A \;=\; & \left(X'X + \sigma^2 P'V^{-1}P\right) Z\beta_A + \\ & + \left(X'X + \sigma^2 P'V^{-1}P\right) \tilde{P}V\tilde{P}'X' \left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1} \cdot \\ & \cdot \left(y - XZ\beta_A\right).\end{aligned}$$

Because of $PZ = 0$ this can be written as

$$\begin{aligned}\left(X'X + \sigma^2 P'V^{-1}P\right) a_A &= X'XZ\beta_A + \\ &+X'X\tilde{P}V\tilde{P}'X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) \\ &+\sigma^2 P'\tilde{P}'X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A).\end{aligned}$$

Adding and subtracting $\sigma^2 X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A)$ and using $P'\tilde{P}' = (I_{Tn} - ZZ')$ results in

$$\begin{aligned}\left(X'X + \sigma^2 P'V^{-1}P\right) a_A &= X'XZ\beta_A + \\ &+X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) \\ &-\sigma^2 X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) \\ &+\sigma^2 (I_{Tn} - ZZ') X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A)\end{aligned}$$

which reduces to

$$\begin{aligned}\left(X'X + \sigma^2 P'V^{-1}P\right) a_A &= X'XZ\beta_A + \\ &+X'(y - XZ\beta_A) \\ &-\sigma^2 X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) \\ &+\sigma^2 X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A) \\ &-\sigma^2 ZZ'X'\left(X\tilde{P}V\tilde{P}'X + \sigma^2 I_T\right)^{-1}(y - XZ\beta_A).\end{aligned}$$

According to (5.14), the last term is zero and we obtain

$$\left(X'X + \sigma^2 P'V^{-1}P\right) a_A = X'y.$$

This shows that the least squares estimator a_{LS} and the Aitken estimator a_A coincide. □

As a consequence of Claim 1, the least-squares estimates for u, v, and w and their Aitken counterparts coincide for $G = \sigma^2 V^{-1}$. We need not distinguish them and denote all our estimates by circumflex:

$$a_A = a_{LS} = \hat{a} = Z\hat{\beta} + \tilde{P}V\tilde{P}'X'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)^{-1}\left(y - XZ\hat{\beta}\right) \quad (7.3)$$

$$u_A = u_{LS} = \hat{u} = \left(I_T - X\left(X'X + \sigma^2 P'V^{-1}P\right)^{-1}X'\right)y \quad (7.4)$$

$$v_A = v_{LS} = \hat{v} = P\left(X'X + P'\sigma^2 P'V^{-1}P\right)^{-1}X'y \quad (7.5)$$

$$w_A = w_{LS} = \hat{v} = X\tilde{P}\hat{v} + \hat{u}. \quad (7.6)$$

For the sake of completeness and later use, the following observation is added:

Claim 2. $G = \sigma^2 V^{-1}$ implies $\hat{Q} = \sigma^2 \hat{w}' W^{-1} \hat{w}$. In other words: the sum of squared deviations weighted by the variance ratios $\frac{\sigma^2}{\sigma_1^2}, \frac{\sigma^2}{\sigma_2^2}, \ldots, \frac{\sigma^2}{\sigma_n^2}$ equals the weighted sum of squares (the squared Mahalanobis distance) in the Aitken regression.

Proof. As $\hat{v} = X\tilde{P}\hat{v} + \hat{u}$, we have

$$\begin{aligned}
\hat{Q} &= \hat{u}'\hat{u} + \sigma^2\hat{\hat{v}}'V^{-1}\hat{\hat{v}} \\
&= \hat{u}'\left(\hat{w} - X\tilde{P}\hat{v}\right) + \sigma^2\hat{v}'V^{-1}\hat{v} \\
&= \hat{u}'\hat{w} - \hat{u}'X\tilde{P}\hat{v} + \sigma^2\hat{v}'V^{-1}\hat{v} \\
&= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P} - \sigma^2\hat{v}'V^{-1}\right)\hat{v} \\
&= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P} - \sigma^2\hat{v}'V^{-1}\right)P\hat{a}
\end{aligned}$$

With (7.3) and (5.8) this gives

$$
\begin{aligned}
\hat{Q} &= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P} - \sigma^2\hat{v}'V^{-1}\right)P\left(Z\hat{\beta} + \tilde{P}V\tilde{P}'X'W^{-1}\hat{w}\right) \\
&= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P} - \sigma^2\hat{v}'V^{-1}\right)P\tilde{P}V\tilde{P}'X'W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P} - \sigma^2\hat{v}'V^{-1}\right)V\tilde{P}'X'W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \left(\hat{u}'X\tilde{P}V\tilde{P}'X' - \sigma^2\hat{v}'\tilde{P}'X'\right)W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \hat{u}'X\tilde{P}V\tilde{P}'X'W^{-1}\hat{w} + \sigma^2\hat{v}'\tilde{P}'X'W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \hat{u}'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T - \sigma^2 I_T\right)W^{-1}\hat{w} + \sigma^2\hat{v}'\tilde{P}'X'W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \hat{u}'\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right)W^{-1}\hat{w} + \sigma^2\hat{u}'W^{-1}\hat{w} + \sigma^2\hat{v}'\tilde{P}'X'W^{-1}\hat{w} \\
&= \hat{u}'\hat{w} - \hat{u}'\hat{w} + \sigma^2\hat{u}'W^{-1}\hat{w} + \sigma^2\hat{v}'\tilde{P}'X'W^{-1}\hat{w} \\
&= \sigma^2\left(\hat{u}' + \hat{v}'\tilde{P}'X'\right)W^{-1}\hat{w}
\end{aligned}
$$

and finally

$$\hat{Q} = \sigma^2\hat{w}'W^{-1}\hat{w}.$$

Hence the weighted sum of squares Q equals the squared Mahalanobis distance. □

Consider now the distribution of $\hat{a}$. The matrix $(X'X + \sigma^2 P'V^{-1}P)$, henceforth referred to as the "system matrix", will be denoted by M:

$$M = \left(X'X + \sigma^2 P'V^{-1}P\right). \tag{7.7}$$

With this, the normal equation (7.2), which defines the solution for the vector of the coefficients $\hat{a}$ can be written as

$$M\,\hat{a} = X'y. \tag{7.8}$$

With (3.1) and (7.7) we obtain

$$\begin{aligned}\hat{a} &= M^{-1}X'(Xa+u) \\ &= M^{-1}\left(X'Xa + X'u + \sigma^2 P'V^{-1}Pa - \sigma^2 P'V^{-1}Pa\right) \\ &= a + M^{-1}\left(X'u - \sigma^2 P'V^{-1}v\right). \end{aligned} \tag{7.9}$$

Given a realization of the time-path of the coefficients a, the estimator $\hat{a}$ is distributed with mean

$$E\{\hat{a}\,|a\} = a \tag{7.10}$$

and covariance

$$E\left\{(a-\hat{a})(a-\hat{a})'\right\} = M^{-1}\begin{pmatrix} X' & -\sigma^2 P'V^{-1} \end{pmatrix}\begin{pmatrix} \sigma^2 I_T & 0 \\ 0 & V \end{pmatrix}\begin{pmatrix} X \\ -\sigma^2 V^{-1}P \end{pmatrix} M^{-1}$$

which reduces to

$$E\left\{(a-\hat{a})(a-\hat{a})'\right\} = M^{-1}\begin{pmatrix} \sigma^2 X'X & +\sigma^4 P'V^{-1}P \end{pmatrix} M^{-1}$$

and finally to

$$E\left\{(a-\hat{a})(a-\hat{a})'\right\} = \sigma^2 M^{-1}. \tag{7.11}$$

The system matrix (7.7) is determined by the observations X, the variance σ^2 and the variances Σ. Equation (7.11) gives the precision of our estimate which is directly related to the system matrix M. The next step is to determine the variance σ^2 and the variances Σ.

8 Moments Estimation of the Variances

The moments estimator that will be developed in this section has, for any sample size, a straightforward interpretation: It is defined by the property that the variances of the disturbances in the estimated coefficients equal their expectations. It has, thus, a straightforward connotation even in shorter time series and does not presuppose that the perturbations u and v are normally distributed. It will be shown later that the moments estimators approach the respective maximum likelihood estimators in large samples if the disturbances are normally distributed. Hence the intuitive appeal of the moments estimator carries over to the likelihood estimator, and the attractive large-sample properties of the likelihood estimator carry over to the moments estimator.

In the following we denote the estimated coefficients by $\hat{a}$ and the estimated perturbations by $\hat{u}$ and $\hat{v}$. For some variances σ^2 and $\sum = \text{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\right)$, the estimated coefficients $\hat{a}$ along with the estimated disturbances $\hat{u}$ and $\hat{v}$ are random variables brought about by realizations of the random variables u and v. Consider $\hat{u} = y - X\hat{a} = X(a - \hat{a}) + u$ first. With (7.9) we obtain

$$\begin{aligned}\hat{u} &= -X\left(M^{-1}\left(X'u - \sigma^2 P'V^{-1}v\right)\right) + u \\ &= \left(I_T - XM^{-1}X'\right)u + \sigma^2 XM^{-1}P'V^{-1}v.\end{aligned}$$

Regarding $\hat{v}$, consider the vectors $\hat{v}_i' = \left(\hat{v}_{i,1}^2, \hat{v}_{i,3}^2, \ldots, \hat{v}_{i,T-1}^2\right)$ for $i = 1, 2, .., n$, that is, the disturbances in the coefficients $\hat{a}_i$ for each coefficient separately. These are obtained as follows.

Denote by $e_i \in \mathbb{R}^n$ the n-th column of an $n \times n$ identity matrix and define the $(T-1) \times (T-1)\, n$-matrix

$$E_i := I_{T-1} \otimes e_i' \tag{8.1}$$

that picks the time-path of the i–th disturbance $v_i = (v_{i,1}, v_{i,3}, \ldots v_{i,T-1})'$ from the disturbance vector v:

$$v_i := E_i v.$$

Note that

$$\sum_{i=1}^{n} \sigma_i^2 E_i' E_i \quad = \quad V. \tag{8.2}$$

Pre-multiplying (7.9) with the matrices E_i yields

$$\hat{v}_i = E_i \left(I_{(T-1)n} - \sigma^2 P M^{-1} P' V^{-1} \right) v + E_i P M^{-1} X' u.$$

Thus $\hat{u}$ and $\hat{v}_i$ are linear functions of the random variables u and v, and their expected squared errors can be calculated.

Claim 3. For given observations X and y and given variances $\sigma^2$and Σ, the expected squared deviations of $\hat{u}$ and $\hat{v}_i$, $i = 1, 2, \dots, n$ are

$$E\left\{\hat{u}'\hat{u}\right\} \quad = \quad \sigma^2 \left(T - \mathrm{tr} X M^{-1} X' \right) \tag{8.3}$$

$$E\left\{\hat{v}_i'\hat{v}_i\right\} \quad = \quad (T-1)\,\sigma_i^2 - \sigma^2 \mathrm{tr} E_i P M^{-1} P' E_i'. \tag{8.4}$$

This implies that the expected sum of squares is

$$E\left\{\hat{Q}\right\} \quad = \quad \sigma^2 (T-n). \tag{8.5}$$

Proof. The expectation of the squared estimated error $\hat{u}$ is

$$
\begin{aligned}
E\left\{\hat{u}'\hat{u}\right\} &= E\left\{\left(u'\left(I_T + XM^{-1}X'\right) + \sigma^2 v'V^{-1}PM^{-1}X'\right)\cdot\right. \\
&\quad \left.\left(\left(I_T - XM^{-1}X'\right)u + \sigma^2 XM^{-1}P'V^{-1}v\right)\right\} \\
&= E\left\{u'\left(I_T - XM^{-1}X'\right)\left(I_T - XM^{-1}X'\right)u\right\} + \\
&\quad +\sigma^4 E\left\{v'V^{-1}PM^{-1}X'XM^{-1}P'V^{-1}v\right\} \\
&= \mathrm{tr}E\left\{u'\left(I_T - XM^{-1}X'\right)\left(I_T - XM^{-1}X'\right)u\right\} + \\
&\quad +\sigma^4\mathrm{tr}E\left\{v'V^{-1}PM^{-1}X'XM^{-1}P'V^{-1}v\right\} S \\
&= \mathrm{tr}E\left\{\left(I_T - XM^{-1}X'\right)uu'\left(I_T - XM^{-1}X'\right)\right\} + \\
&\quad +\sigma^4\mathrm{tr}E\left\{XM^{-1}P'V^{-1}vv'V^{-1}PM^{-1}X'\right\} \\
&= \mathrm{tr}\sigma^2\left(I_T - XM^{-1}X'\right)\left(I_T - XM^{-1}X'\right) + \mathrm{tr}\sigma^4 XM^{-1}P'V^{-1}PM^{-1}X' \\
&= \sigma^2\mathrm{tr}\left(\left(I_T - XM^{-1}X'\right)\left(I_T - XM^{-1}X'\right) + \sigma^2 XM^{-1}P'V^{-1}PM^{-1}X'\right) \\
&= \sigma^2\mathrm{tr}\left(I - 2XM^{-1}X' + XM^{-1}X'XM^{-1}X' + \sigma^2 XM^{-1}P'V^{-1}PM^{-1}X'\right) \\
&= \sigma^2\mathrm{tr}\left(I_T - 2XM^{-1}X' + XM^{-1}\left(X'X + \sigma^2 P'V^{-1}P\right)M^{-1}X'\right) \\
&= \sigma^2\mathrm{tr}\left(I_T - XM^{-1}X'\right) \\
&= \sigma^2\left(T - \mathrm{tr}XM^{-1}X'\right).
\end{aligned}
$$

In a similar way, the expectation of the squared estimated disturbance in the i-th coefficient $\hat{v}_i$ is evaluated as

$$
\begin{aligned}
E\left\{\hat{v}_i'\hat{v}_i\right\} &= E\left\{\left(u'XM^{-1}P'E_i' + \nu'\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'\right)\right. \\
&\cdot \quad \left.\left(E_iPM^{-1}X'u + E_i\left(I_{(T-1)n} - \sigma^2 PM^{-1}P'V^{-1}\right)\nu\right)\right\} \\
&= E\left\{u'XM^{-1}P'E_i'E_iPM^{-1}X'u \, + \right. \\
&\quad \left.\nu'\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'E_i\left(I_{(T-1)n} - \sigma^2 PM^{-1}P'V^{-1}\right)\nu\right\} \\
&= E\left\{\mathrm{tr}\left(u'XM^{-1}P'E_i'E_iPM^{-1}X'u \; + \right.\right. \\
&\quad \left.\left.\nu'\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'E_i\left(I_{(T-1)n} - \sigma^2 PM^{-1}P'V^{-1}\right)\nu\right)\right\} \\
&= E\left\{\mathrm{tr}\left(E_iPM^{-1}X'uu'XM^{-1}P'E_i' \; + \right.\right. \\
&\quad \left.\left.E_i\left(I_{(T-1)n} - \sigma^2 PM^{-1}P'V^{-1}\right)\nu\nu'\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'\right)\right\} \\
&= \mathrm{tr}\left(\sigma^2 E_iPM^{-1}X'XM^{-1}P'E_i'\right) + \\
&\quad \mathrm{tr}\left(E_i\left(I_{(T-1)n} - \sigma^2 PM^{-1}P'V^{-1}\right)V\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(\sigma^2 E_iPM^{-1}X'XM^{-1}P'E_i'\right) + \\
&\quad \mathrm{tr}\left(E_i\left(V - \sigma^2 PM^{-1}P'\right)\left(I_{(T-1)n} - \sigma^2 V^{-1}PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(\sigma^2 E_iPM^{-1}X'XM^{-1}P'E_i'\right) + \\
&\quad \mathrm{tr}\left(E_i\left(V - \sigma^2 PM^{-1}P'\right)E_i' - \sigma^2 E_i\left(V - \sigma^2 PM^{-1}P'\right)V^{-1}PM^{-1}P'E_i'\right) \\
&= \mathrm{tr}\left(\sigma^2 E_iPM^{-1}X'XM^{-1}P'E_i'\right) + \\
&\quad \mathrm{tr}\left(E_i\left(V - \sigma^2 PM^{-1}P' - \sigma^2 PM^{-1}P' + \sigma^4 PM^{-1}P'V^{-1}PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(\sigma^2 E_iPM^{-1}X'XM^{-1}P'E_i' \; + \right. \\
&\quad \left.E_i\left(V - \sigma^2 PM^{-1}P' - \sigma^2 PM^{-1}P' + \sigma^4 PM^{-1}P'V^{-1}PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(E_i\left(\left(\sigma^2 PM^{-1}\left(X'X + \sigma^2 P'V^{-1}P\right)M^{-1}P'\right) + V - 2\sigma^2 PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(E_i\left(V - \sigma^2 PM^{-1}P'\right)E_i'\right) \\
&= \mathrm{tr}\left(E_iVE_i' - \sigma^2 E_iPM^{-1}P'E_i'\right) \\
&= \mathrm{tr}\left(\left(I_{T-1}\otimes e_i'\right)\left(I_{T-1}\otimes\Sigma\right)\left(I_{T-1}\otimes e_i\right) - \sigma^2 E_iPM^{-1}P'E_i'\right) \\
&= \mathrm{tr}\left(I_{T-1}\otimes e_i'\Sigma e_i\right) - \sigma^2\mathrm{tr}\left(E_iPM^{-1}P'E_i'\right) \\
&= \left(T-1\right)\sigma_i^2 - \sigma^2\mathrm{tr}\left(E_iPM^{-1}P'E_i'\right).
\end{aligned}
$$

Regarding $\hat{Q}$ we note that

$$X'X + \sigma^2 P'V^{-1}P = X'X + \sigma^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2} P_i' P_i = M$$

and obtain

$$\begin{aligned}
E\left\{\hat{Q}\right\} &= \sigma^2\left(T - \mathrm{tr}XM^{-1}X'\right) + \sum_{i=1}^{n} \frac{\sigma^2}{\sigma_i^2}\left(\left(T-1\right)\sigma_i^2 - \sigma^2 \mathrm{tr}E_i P M^{-1} P' E_i'\right) \\
&= \sigma^2\left(T - \mathrm{tr}XM^{-1}X' + \sum_{i=1}^{n}\left(T-1\right) - \sum_{i=1}^{n}\frac{\sigma^2}{\sigma_i^2}\mathrm{tr}E_i P M^{-1} P' E_i'\right) \\
&= \sigma^2\left(T + n\left(T-1\right) - \mathrm{tr}XM^{-1}X' - \mathrm{tr}\left(\sum_{i=1}^{n}\frac{\sigma^2}{\sigma_i^2}E_i P M^{-1} P' E_i'\right)\right) \\
&= \sigma^2\left(Tn - T - n - \mathrm{tr}M^{-1}X'X - \mathrm{tr}\left(M^{-1}\sum_{i=1}^{n}\frac{\sigma^2}{\sigma_i^2}P'E_i'E_iP\right)\right) \\
&= \sigma^2\left(Tn - T - n - \mathrm{tr}M^{-1}X'X - \mathrm{tr}\left(M^{-1}\sum_{i=1}^{n}\sigma^2 P'V^{-1}P\right)\right) \\
&= \sigma^2\left(Tn - T - n - \mathrm{tr}M^{-1}\left(X'X - \sigma^2 P'V^{-1}P\right)\right) \\
&= \sigma^2\left(Tn - T - n - \mathrm{tr}I_{nT}\right) \\
&= \sigma^2\left(T - n\right).
\end{aligned}$$

□

The moments estimators are obtained by selecting variances σ^2 and $\sigma_i^2, i = 1, 2, ..., n$ such that the expected moments $E\{\hat{u}'\hat{u}\}$ and $E\{\hat{v}_i'\hat{v}_i\}, i = 1, 2, ..., n$ are equalized to the estimated moments $\hat{u}'\hat{u}$ and $\hat{v}_i'\hat{v}_i,\ i = 1, 2, ..., n$. As both the expected moments and the estimated moments are functions of the variances, the moments estimators, denoted by $\hat{\sigma}^2$ and $\hat{\sigma}_i^2,\ i = 1, 2, ..., n$, respectively, are defined as a fix point of the system

$$\begin{aligned}
E\{\hat{u}'\hat{u}\} &= \hat{u}'\hat{u} \\
E\{\hat{v}_i'\hat{v}_i\} &= \hat{v}_i'\hat{v}_i.
\end{aligned}$$

Alternatively, the moments estimators can be equivalently defined as a fix point of the system:

$$
\begin{aligned}
E\left\{\hat{v}_i'\hat{v}_i\right\} &= \hat{v}_i'\hat{v}_i \\
E\left\{\hat{Q}\right\} &= \hat{Q}.
\end{aligned}
$$

The implementations Schlicht (2005a, 2005b) use the latter alternative and employ a gradient process to find the solution of the equation system

$$
\begin{aligned}
\hat{v}_i'\hat{v}_i &= (T-1)\,\hat{\sigma}_i^2 - \hat{\sigma}^2 \mathrm{tr} E_i P \hat{M}^{-1} P' E_i' \\
\frac{1}{T-n}\hat{Q} &= \hat{\sigma}^2.
\end{aligned}
$$

This can be written as

$$
\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} = \left(\frac{\hat{v}_i'\hat{v}_i}{\hat{Q}}(T-n) - \mathrm{tr} E_i P \hat{M}^{-1} P' E_i'\right)\frac{1}{T-1} \tag{8.6}
$$

$$
\hat{\sigma}^2 = \frac{1}{T-n}\hat{Q}. \tag{8.7}
$$

Iteration starts with some variance ratios $\gamma_i = \frac{\sigma^2}{\sigma_i^2}$. This permits to determine the right-hand sides of equations (8.6) and (8.7). The variance ratios at the left-hand side of (8.6) and the variance at the left hand side of (8.7) are used for a new iteration, and this continues until convergence is reached, delivering the fix-point values $\hat{\gamma}_i = \frac{\hat{\sigma}^2}{\hat{\sigma}_i^2}$ and $\hat{\sigma}^2$ and the corresponding variances $\hat{\sigma}_i^2 = \frac{\hat{\sigma}^2}{\hat{\gamma}_i}$. (If this process does not converge, another solution procedure is available that will be discussed in Section 10 below.)

9 Maximum Likelihood Estimation of the Variances

This section derives a maximum-likelihood estimator for the variances under the additional assumption that the disturbances u and v are normally distributed.

Using equations (3.2) and (5.9) – (5.13) together with the identity $x = XZ$, the concentrated log-likelihood function for the Aitken regression (5.8) can be written as

$$\mathscr{L}\left(\ \sigma^2, \Sigma\ \right) = -\frac{1}{2}\left(T\left(\log 2 + \log \pi\right) + \log\det W\right) - \frac{1}{2}\left(y - XZ\beta\right)' W^{-1}\left(y - XZ\beta\right) \tag{9.1}$$

with

$$W \quad = \quad X\tilde{P}\left(I_{T-1} \otimes \Sigma\right)\tilde{P}'X' + \sigma^2 I_T.$$

By maximizing (9.1) with respect to β,σ^2 and Σ, the maximum likelihood estimates for the variances are obtained and the corresponding expectation for the parameter a is given in analogy to (7.3) as

$$\check{a} \quad = \quad Z\check{\beta} + \tilde{P}\check{V}\tilde{P}'X'\left(X\tilde{P}\check{V}\tilde{P}'X' + \check{\sigma}^2 I_T\right)^{-1}\left(y - XZ\check{\beta}\right)$$

with a caron denoting the maximum likelihood estimates and $\check{V} = \left(I_{T-1} \otimes \check{\Sigma}\right)$.

The maximum likelihood estimator can be characterized in another way. This will be explained in the following. In order to do so, the following lemma is needed.

Claim 4.

$$\begin{aligned} \log\det W \quad = \quad & \log\det\left(PMP'\right) + \left(T-1\right)\sum_{i=1}^{n}\log\sigma_i^2 - \\ & \left(\left(T-1\right)n - T\right)\log\sigma^2 - 2\log\det\left(PP'\right). \end{aligned} \tag{9.2}$$

Proof.

$$
\begin{aligned}
\det W &= \det\left(X\tilde{P}V\tilde{P}'X' + \sigma^2 I_T\right) \\
&= \left(\sigma^2\right)^T \det\left(\frac{1}{\sigma^2}X\tilde{P}V^{\frac{1}{2}}V^{\frac{1}{2}}\tilde{P}'X' + I_T\right) \\
&= \left(\sigma^2\right)^T \det\left(\frac{1}{\sigma^2}V^{\frac{1}{2}}\tilde{P}'X'X\tilde{P}V^{\frac{1}{2}} + I_{(T-1)n}\right) \\
&= \left(\sigma^2\right)^T \det\left(V^{\frac{1}{2}}\left(\frac{1}{\sigma^2}\tilde{P}'X'X\tilde{P} + V^{-1}\right)V^{\frac{1}{2}}\right) \\
&= \left(\sigma^2\right)^T \det\left(V\left(\frac{1}{\sigma^2}\left(PP'\right)^{-1}PX'XP'\left(PP'\right)^{-1} + V^{-1}\right)\right) \\
&= \left(\sigma^2\right)^T \det\left(\frac{1}{\sigma^2}V\left(PP'\right)^{-1}P\left(X'X + \sigma^2 P'V^{-1}P\right)P'\left(PP'\right)^{-1}\right) \\
&= \left(\sigma^2\right)^T \det\left(\frac{1}{\sigma^2}V\right)\det\left(PP'\right)^{-1}\det\left(PMP'\right)\det\left(PP'\right)^{-1} \\
&= \left(\sigma^2\right)^T \left(\prod_{i=1}^{n}\frac{\sigma_i^2}{\sigma^2}\right)^{(T-1)}\det\left(PP'\right)^{-2}\det\left(PMP'\right).
\end{aligned}
$$

Hence the result

$$
\begin{aligned}
\log\det W = {} & \log\det\left(PMP'\right) + (T-1)\sum_{i=1}^{n}\log\sigma_i^2 - \\
& \left((T-1)n - T\right)\log\sigma^2 - 2\log\det\left(PP'\right)
\end{aligned}
$$

is obtained. □

Claim 5. Minimizing the criterion

$$
\begin{aligned}
\mathcal{C}_L = \log\det\left(PMP'\right) + (T-1)\sum_{i=1}^{n}\log\sigma_i^2 - \left((T-1)n - T\right)\log\sigma^2 + \\
+ \frac{1}{\sigma^2}u'u + v'V^{-1}v \qquad (9.3)
\end{aligned}
$$

is equivalent to maximizing the likelihood function (9.1).

Proof. With (9.2) we have

$$\mathcal{C}_L + 2\mathscr{L}\left(\sigma^2, \Sigma\right) = \frac{1}{\sigma^2}u'u + v'V^{-1}v - w'W^{-1}w + 2\log\det\left(PP'\right) - T\left(\log 2 + \log\pi\right).$$

As, according to Claim 2, $w'W^{-1}w = \left(y - XZ\beta\right)' W^{-1}\left(y - XZ\beta\right)$ equals $\frac{1}{\sigma^2}u'u + v'V^{-1}v$ and $\log\det\left(PP'\right)$ and $T\left(\log 2 + \log\pi\right)$ are independent of the variances, we can write

$$\mathcal{C}_L = -2\mathscr{L}\left(\sigma^2, \Sigma\right) + \text{constant}$$

where "constant" is independent of the variances and maximization of $\mathscr{L}$ with regard to the variances is equivalent to minimization of $\mathcal{C}_L$. □

10 Another Representation of the Moments Estimator

The relationship between the likelihood estimator and the moments estimator can be elucidated with the aid of a criterion that is very similar to the likelihood criterion (9.3). This criterion function is

$$\mathcal{C}_M\left(\sigma^2, \Sigma\right) = \log\det M + \left(T-1\right)\sum_{i=1}^{n}\log\sigma_i^2 - T\left(n-1\right)\log\sigma^2 + \\ + \frac{1}{\sigma^2}u'u + v'V^{-1}v. \tag{10.1}$$

Claim 6. Minimization of the criterion function (10.1) with respect to the disturbances u and v and the variances σ^2 and Σ yields the moments estimators as defined in (8.3) and (8.4).

Proof. Note that the envelope theorem together with (8.2) implies

$$\frac{\partial}{\partial\sigma^2}\left(\frac{1}{\sigma^2}\hat{u}'\hat{u}+\hat{v}'V^{-1}\hat{v}\right) = -\frac{1}{\sigma^4}\hat{u}'\hat{u} \tag{10.2}$$

$$\frac{\partial}{\partial\sigma_i^2}\left(\frac{1}{\sigma^2}\hat{u}'\hat{u}+\hat{v}'V^{-1}\hat{v}\right) = -\frac{\sigma^2}{\sigma_i^4}\hat{v}_i{}'\hat{v}_i. \tag{10.3}$$

In view of (8.2) we obtain further

$$\frac{\partial \log\det M}{\partial\sigma^2} = \operatorname{tr}\left(M^{-1}P'V^{-1}P\right). \tag{10.4}$$

By definition (7.7) we have

$$M^{-1}\left(X'X+\sigma^2P'V^{-1}P\right) = I$$

and hence

$$M^{-1}P'V^{-1}P = \frac{1}{\sigma^2}\left(I-M^{-1}X'X\right).$$

With this, equation (10.4) can be written as

$$\begin{aligned}\frac{\partial \log\det M}{\partial\sigma^2} &= \operatorname{tr}\left(\frac{1}{\sigma^2}\left(I_{Tn}-M^{-1}X'X\right)\right)\\ &= \frac{1}{\sigma^2}\left(\operatorname{tr}I_{Tn}-\operatorname{tr}M^{-1}X'X\right)\\ &= \frac{Tn}{\sigma^2}-\frac{1}{\sigma^2}\operatorname{tr}XM^{-1}X'.\end{aligned}$$

$$\frac{\partial \log\det M}{\partial\sigma_i^2} = -\frac{\sigma^2}{\sigma_i^4}\operatorname{tr}\left(M^{-1}P'E_i'E_iP\right)$$

and we find

$$\frac{\partial \mathcal{C}_M}{\partial \sigma^2} = \frac{Tn}{\sigma^2} - \frac{1}{\sigma^2}\mathrm{tr}XM^{-1}X'. - \frac{T(n-1)}{\sigma^2} - \frac{1}{\sigma^4}\hat{u}'\hat{u} = 0 \tag{10.5}$$

$$\frac{\partial \mathcal{C}_M}{\partial \sigma_i^2} = -\frac{\sigma^2}{\sigma_i^4}\mathrm{tr}P'F_i'F_iPM^{-1} + (T-1)\frac{1}{\sigma_i^2} - \frac{\sigma^2}{\sigma_i^4}\hat{v}_i{}'\hat{v}_i = 0 \tag{10.6}$$

which gives

$$\begin{aligned} \hat{u}'\hat{u} &= \sigma^2\left(T - \sigma^2\mathrm{tr}XM^{-1}X'\right) \\ \hat{v}_i{}'\hat{v}_i &= (T-1)\,\sigma_i^2 - \sigma^2\mathrm{tr}P'F_i'F_iPM^{-1}. \end{aligned}$$

These first-order conditions are equivalent to equations (8.3), (8.4) that define the moments estimator. □

Johannes Ludsteck's (2004, 2018) Mathematica packages for VC proceed by minimizing the criterion function (10.1). This permits very clean and transparent programming. As Claim 6 is confined to moments and does not require any assumption about the normality of the disturbances, Ludsteck's estimators are moments estimators as well.

11 The Relationship Between the Likelihood and the Moments Estimator

The likelihood estimates minimize, according to Claim 5, the criterion $\mathcal{C}_L$ and the moments estimates minimize, according to Claim 6, the criterion $\mathcal{C}_M$. It is claimed in the following that, for increasing T and bounded X, both estimates tend to coincide. To show that, the following lemma is needed.

Claim 7. For sufficiently large T and bounded explanatory variables X, the following holds true approximately:

$$\det PMP' \approx \det M \det(PP').$$

Proof. Define the $Tn \times Tn$ matrix

$$\mathbb{P} = \begin{pmatrix} P \\ T^{-\frac{1}{2}} Z' \end{pmatrix}$$

and consider the matrix $\mathbb{P}M\mathbb{P}'$. One way to calculate it is as follows:

$$\begin{aligned}
\mathbb{P}M\mathbb{P}' &= \begin{pmatrix} P \\ T^{-\frac{1}{2}} Z' \end{pmatrix} M \begin{pmatrix} P' & T^{-\frac{1}{2}} Z \end{pmatrix} \\
&= \begin{pmatrix} PMP' & T^{-\frac{1}{2}} PMZ \\ T^{-\frac{1}{2}} Z'MP' & T^{-1} Z'Z \end{pmatrix} \\
&= \begin{pmatrix} PMP' & T^{-\frac{1}{2}} PX'XZ \\ T^{-\frac{1}{2}} Z'X'XP' & I_n \end{pmatrix}.
\end{aligned}$$

This implies

$$\begin{aligned}
\det \mathbb{P}M\mathbb{P}' &= \det I_n \det \left(PMP' - \frac{1}{T} PX'XZZ'X'XP' \right) \\
&= \det \left(PMP' - \frac{1}{T} PX'xx'XP' \right) \\
&= \det \left(P \left(M - \frac{1}{T} X'xx'X \right) P' \right) \\
&= \det \left(P \left(X' \left(I_T - \frac{1}{T} xx' \right) X + \sigma^2 P'V^{-1}P \right) P' \right).
\end{aligned}$$

For increasing T and bounded x, $\frac{1}{T}xx'$ tends to zero and $\left(I_T - \frac{1}{T}xx' \right)$ tends to I_T. Hence $\det \mathbb{P}M\mathbb{P}'$ tends to $\det PMP'$ and we can write

$$\det \mathbb{P}M\mathbb{P}' \approx \det PMP' \tag{11.1}$$

for large T. Another way to evaluate $\det(\mathbb{P}M\mathbb{P})$ is the following:

$$\begin{aligned}\det \mathbb{P}M\mathbb{P}' &= \det(M\mathbb{P}'\mathbb{P}) \\ &= \det M \det(\mathbb{P}'\mathbb{P}) \\ &= \det M \det(\mathbb{P}\mathbb{P}')\end{aligned}$$

As

$$\det(\mathbb{P}\mathbb{P}') = \det\begin{pmatrix} PP' & 0 \\ 0 & I_n \end{pmatrix} = \det(PP'),$$

$$\det \mathbb{P}M\mathbb{P}' = \det M \det(PP') \tag{11.2}$$

is obtained. Combining (11.1) and (11.2) gives the result. □

Claim 8. For increasing T and with bounded explanatory variables X, the moments criterion and the likelihood criterion coincide.

Proof. For large T and in view Claim (7), $\mathcal{C}_M$ and $\mathcal{C}_L$ differ by the constant $\log\det(PP') + n$. Hence the minimization of both criteria with respect to the variances will generate the same result. □

In consequence, the descriptive appeal of the moments estimator carries over to the likelihood estimator, and the theoretical appeal of the likelihood estimator carries over to the moments estimator.

Acknowledgements This paper relies heavily on two earlier discussion papers (Schlicht 1989 and Schlicht and Ludsteck 2006). I thank Daniela Diekmann, Theo Dijkstra. Walter Krämer, José Alberto Bravo Lopez, Johannes Ludsteck, and Ralf Pauly for comments on earlier attempts. Not the least I wish to thank the authors of the works that have used VC and are mentioned at the end of the paper.. Their use of VC eventually convinced me that it may be worthwhile to provide this documentation.

References

Athans, Michael (1974). “The Importance of Kalman Filtering Methods for Economic Systems”. In: *Annals of Economic and Social Measurement, Volume 3, number 1*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 49–64. URL: https://ideas.repec.org/h/nbr/nberch/9994.html.

Cooley, Thomas F and Edward C Prescott (1973). “An Adaptive Regression Model”. In: *International Economic Review* 14.2, pp. 364–371. URL: https://ideas.repec.org/a/ier/iecrev/v14y1973i2p364-71.html.

Green, P.J. and B.W. Silverman (2000). *Nonparametric Regression and Generalized Linear Models. a Roughness Penalty Approach*. Boca Raton, Fla.: Chapman & Hall. ISBN: 0412300400.

Henderson, R. (1924). “A New Method of Graduation”. In: *Transactions of the Actuarial Society of America* 25, pp. 29–40.

Hodrick, Robert J and Edward C Prescott (1997). “Postwar U.S. Business Cycles: An Empirical Investigation”. In: *Journal of Money, Credit and Banking* 29.1, pp. 1–16. URL: https://ideas.repec.org/a/mcb/jmoncb/v29y1997i1p1-16.html.

Keynes, John Maynard (1939). “Professor Tinbergen’s Method”. In: *Economic Journal* 49.195, pp. 558–68. URL: https://www.jstor.org/stable/2224838.

— (1973). *The General Theory and After, Part II: Defense and Development*. Ed. by Moggridge. Vol. xiv. The Collected Works of John Maynard Keynes. London: Macmillan.

Leser, C. E. V. (1961). “A Simple Method of Trend Construction”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 23, pp. 91–107. URL: http://www.jstor.org/stable/2983845.

— (1963). “Estimation of Quasi-Linear Trend and Seasonal Variation”. In: *Journal of the American Statistical Association* 58, pp. 1033–43. URL: http://www.jstor.org/stable/2283329.

Ludsteck, Johannes (2004). *VC Package for Mathematica*. URL: http://library.wolfram.com/infocenter/MathSource/5195/.

— (2018). *VC Packages for Estimating Time-Varying Coefficients with Mathematica 8 - 11*. URL: https://epub.ub.uni-muenchen.de/59479/.

Schlicht, Ekkehart (1973). *Forcasting Markov Chains. A Theoretical Foundation for Exponential Smoothing*. Working paper B 13, Department of Economics, University of Regensburg. URL: http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht-exponential_smoothing.pdf.

— (1977). *Grundlagen der ökonomischen Analyse*. Reinbek: Rowohlt. URL: https://epub.ub.uni-muenchen.de/25821/.

— (1981). "A Seasonal Adjustment Principle and a Seasonal Adjustment Method Derived from This Principle". In: *Journal of the American Statistical Association* 76(374). Paper presented at the Econometric Society European Meeting Helsinki 1976, pp. 374–78. URL: http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht-seasonal_adjustment.pdf.

— (1984). "Seasonal Adjustment in a Stochastic Model". In: *Statistische Hefte* 25, pp. 1–12. URL: http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht_seasonal-adjustment-in-a-stochastic-model.pdf.

— (1985). *Isolation and Aggregation in Economics*. Annotated electronic reprint. 38821. Springer-Verlag. URL: https://ideas.repec.org/p/lmu/muenec/38821.html.

— (1989). "Variance Estimation in a Random Coefficients Model". In: *Munich Discussion Paper*. mwb 2006-12. Paper presented at the Econometric Society European Meeting Munich 1989. URL: https://epub.ub.uni-muenchen.de/59143/.

— (2005a). "Estimating the Smoothing Parameter in the So-Called Hodrick-Prescott Filter". In: *Journal of the Japan Statistical Society* 35, pp. 99–119. URL: https://doi.org/10.14490/jjss.35.99.

— (2005b). *VC - A Program for Estimating Time-Varying Coefficients*. URL: http://epub.ub.uni-muenchen.de/archive/00000684/.

— (2005c). *VCC - A Program for Estimating Time-Varying Coefficients. Console Version With Source Code in C*. URL: http://epub.ub.uni-muenchen.de/archive/00000719/.

Schlicht, Ekkehart and Johannes Ludsteck (2006). "Variance Estimation in a Random Coefficients Model". In: *Munich Discussion Paper* 2006-12. URL: https://ideas.repec.org/p/lmu/muenec/904.html.

Schlicht, Ekkehart and Ralf Pauly (1983). "Descriptive Seasonal Adjustment by Minimizing Perturbations". In: *Empirica*. URL: http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht-pauly-perturbations.pdf.

Tinbergen, Jan (1940). "On a Method of Statistical Business-Cycle Research. A Reply". In: *Economic Journal* 50.197, pp. 141–54. URL: http://www.jstor.org/stable/2225763.

Whittaker, E. T. (1923). "On a New Method of Graduation". In: *Proceedings of the Edinburgh Mathematical Society* 41, pp. 63–75.

Contributions That Have Employed VC

Afonso, Antonio, Pedro Gomes, and Davide Furceri (2011). "Sovereign credit ratings and financial markets linkages: application to European data". In: *European Central Bank Working Paper Series* 1347. URL: https://ideas.repec.org/p/ecb/ecbwps/20111347.html.

Afonso, Antonio and Joao Tovar Jalles (2015). "Economic Volatility and Sovereign Yields' Determinants: A Time-Varying Approach". In: *Working Papers Department of Economics*. URL: https://ssrn.com/abstract=2712221.

— (2017a). "Euro area time-varying fiscal sustainability". In: *International Journal of Finance & Economics* 22.3, pp. 244–254. URL: https://ideas.repec.org/a/wly/ijfiec/v22y2017i3p244-254.html.

— (2017b). "Sovereign debt composition and time-varying public finance sustainability". In: *The North American Journal of Economics and Finance* 42.C, pp. 144–155. DOI: 10.1016/j.najef.2017.08.0. URL: https://ideas.repec.org/a/eee/ecofin/v42y2017icp144-155.html.

— (2019). "Decomposing and Analysing the Determinants of Current Accounts' Cyclicality: Evidence from the Euro Area". In: *Open Economies Review*. DOI: 10.1007/s11079-018-9503-2. URL: https://ideas.repec.org/a/kap/openec/v30y2019i1d10.1007_s11079-018-9503-2.html.

Afonso, António and João Tovar Jalles (Jan. 2017). *Sovereign Debt Effects and Composition: Evidence from Time-Varying Estimates*. Working Papers Department of Economics 2017/03. ISEG - Lisbon School of Economics and Management, Department of Economics, Universidade de Lisboa. URL: https://ideas.repec.org/p/ise/isegwp/wp032017.html.

Baxa, Jaromir, Roman Horvath, and Borek Vasicek (2011). "Time Varying Monetary Policy Rules and Financial Stress". In: ed. by S. Eijffinger and Donato Masciandaro. Edward Elgar Publishing. Chap. Time Varying Monetary Policy Rules and Financial Stress, pp. 268–312. URL: http://www.e-elgar.com/shop/handbook-of-central-banking-financial-regulation-and-supervision.

— (2014). "How Does Monetary Policy Change? Evidence on Inflation-Targeting Countries". In: *Macroeconomic Dynamics* 18.3, pp. 593–630. DOI: 10.1017/S1365100512000545.

Cohen, Gail et al. (2017). "Emissions and Growth: Trends and Cycles in a Globalized World". In: *IMF Working Paper* WP17/191. URL: https://www.imf.org/~/media/Files/Publications/WP/2017/wp17191.ashx.

Conway, P., R. Herd, and T. Chalaux (2010). "Reforming China's Monetary Policy Framework to Meet Domestic Objectives". In: *6th NIPFP DEA Research Program Conference*. DOI: doi.org/10.1787/5km32vmsq6f2-en.

Cournede, Boris, Rudiger Ahrend, and Robert Price (2008). "Have Long-Term Financial Trends Changed the Transmission of Monetary Policy?" In: *OECD Economic Department Working Papers* 634. URL: http://dx.doi.org/10.1787/238203348082.

Cournede, Boris and Frederic Gonand (2006). "Restoring Fiscal Sustainability in the Euro Area: Raise Taxes or Curb Spending?" In: *Working Paper*. URL: http://pcsi.pa.go.kr/files/JT03216787.pdf.

Flaschel, P. et al. (2008). *Topics in Applied Macrodynamic Theory*. Berlin-Heidelberg: Springer-Verlag. DOI: 10.1007/978-3-540-72542-8.

Flaschel, Peter (2008). "Linking Goods with Labor Markets: Okun's Law and Beyond". In: *Topics in Applied Macrodynamic Theory*. Berlin, Heidelberg:

Springer Berlin Heidelberg, pp. 213–283. ISBN: 978-3-540-72542-8. DOI: 10.1007/978-3-540-72542-8_5. URL: https://doi.org/10.1007/978-3-540-72542-8_5.

Flaschel, Peter and Ekkehart Schlicht (2006). "The New Keynesian Theory and the New Phillips Curves: A Competing Approach". In: *Quantitative and Empirical Analysis of Nonlinear Dynamic Macromodels*. Ed. by Carl Chiarella et al. Elsevier. Chap. The New Keynesian Theory and the New Phillips Curves: A Competing Approach, pp. 113–145. URL: https://books.google.de/books?id=5MmwM5vyhUQC.

Franke, Reiner (2006). *Themes on Okun's Law and Beyond*. Tech. rep. SCEPA. DOI: 10.1007/978-3-540-72542-8_5.

Furceri, Davide and Joao Tovar Jalles (2017). "Fiscal Stabilization and Productive Investment: Evidence from Advanced Economies". In: *Working Paper*. URL: https://www.bancaditalia.it/pubblicazioni/altri-atti-convegni/2018-workshop-public-finance/7-Furceri-Jalles.pdf?language_id=1.

— (2018). "Fiscal counter-cyclicality and productive investment: evidence from advanced economies". In: *The B.E. Journal of Macroeconomics*. DOI: 10.1515/bejm-2017-0222.

Furceri, Davide, Joao Tovar Jalles, and Aleksandra Zdzienicka (2017). "China Spillovers: New Evidence from Time-Varying Estimates". In: *Open Economies Review* 28.3, pp. 413–429. URL: https://link.springer.com/article/10.1007/s11079-016-9430-z.

He, Qizhi (2012). "Research on the Dynamic Relationship between Prices of Agricultural Futures in China and Japan". In: *Journal of Computers* 7(9). DOI: 10.4304/jcp.7.9.2342-2350.

Jalles, João Tovar (2018). "On the Time-Varying Relationship Between Unemployment and Output: What shapes it?" In: *Scottish Journal of Political Economy* 66, pp. 605–630. DOI: 10.1111/sjpe.12200.

— (2019a). *Explaining Africa's Public Consumption Procyclicality: Revisiting Old Evidence*. Working Papers REM 2019/0100. ISEG - Lisbon School of Economics and Management, REM, Universidade de Lisboa. URL: https://ideas.repec.org/p/ise/remwps/wp01002019.html.

Jalles, João Tovar (2019b). *On the Cyclicality of Social Expenditure: New Time-Varying evidence from Developing Economies*. Working Papers REM 2019/82. ISEG - Lisbon School of Economics and Management, REM, Universidade de Lisboa. URL: https://ideas.repec.org/p/ise/remwps/wp0822019.html.

— (2019c). *Polluting Emissions and GDP: Decoupling Evidence from Brazilian States*. Working Papers REM 2019/0104. ISEG - Lisbon School of Economics and Management, REM, Universidade de Lisboa. URL: https://ideas.repec.org/p/ise/remwps/wp01042019.html.

López-Herrera, Francisco, Humberto Valencia-Herrera, et al. (2016). "Hacia un Modelo de Valuación de Activos de Capital para México: Análisis de Activos Individuales con Coeficientes Variantes en el Tiempo". In: *Panorama Económico* 11.22, pp. 75–103. URL: http://yuss.me/revistas/panorama/pano2016v11n22a04p074_103.pdf.

Madsen, Michael Walter (2012). "Does Financial Stress Have an Impact on Monetary Policy?: An Econometric Analysis Using Norwegian Data". PhD thesis. Department of Economics, University of Oslo. URL: https://www.duo.uio.no/bitstream/handle/10852/17126/Madsen_Michael_Masteroppgave.pdf.

Nogueira, Reginaldo Pinto (2009). "Testing credibility with time-varying coefficients". In: *Applied Economics Letters* 16.18, pp. 1813–1817. DOI: 10.1080/13504850701719611.

Rodriguez, Cesar M. (2014). "Financial development, fiscal policy and volatility: Their effects on growth". In: *The Journal of International Trade & Economic Development* 23.2, pp. 223–266. DOI: 10.1080/09638199.2012.711014.