Estimating Time-Varying Coefficients With the VC Program

Ekkehart Schlicht

University of Munich,
Schackstr. 4, 80539 Munich, Germany,
schlicht@lmu.de

version 4 , December 2003

Abstract

The estimation of models with time-varying coefficients is usually performed by Kalman-Bucy filtering. The two-sided filter proposed by Schlicht (1988) is statistically and computationally superior to the one-sided Kalman-Bucy filter. This paper describes the estimation procedure and the program package that implements the two-sided filter.

Keywords: Kalman filtering, Kalman-Bucy, random walk, time-varying coefficients, adaptive estimation, time-series

Journal of Economic Literature classification numbers: C22

CONTENTS

> The object of a model is to segregate
> the semipermanent or relatively constant factors
> from those which are transitory or fluctuating
> so as to develop a logical way of thinking
> about the latter, and of understanding
> the time sequences to which they give rise
> in particular cases.
>
> JOHN MAYNARD KEYNES (1978, 296)

1 INTRODUCTION

The estimation of models with time-varying coefficients is usually considered an arcane affair, and practical implementation requires much expertise and manual fine-tuning. The VC program described in this note is intended to improve the situation. The program offers a statistically superior alternative to Kalman-Bucy filtering, is easy to use, and freely available.

The following sections restate the underlying methodological considerations briefly, explain the statistical model and the estimation procedure, and give a description of the program and its features. A concluding note comments on some more advanced statistical details.

2 THE METHODOLODICAL FOUNDATION: THE LOGIC OF *Ceteris Paribus* ANALYSIS

Economic analysis proceeds *ceteris paribus* The data of any problem - those things we presuppose in our argument in order to explain the behavior of some economic variables - are assumed constant, yet constant only for the time being. The assumption makes it possible to analyze selected issues in isolation and derive results that remain valid even if the factors frozen under a *ceteris paribus* clause are changing over time.[1]

The main point in economic analysis seems to be that the data of any model must be *relatively* stable with regard to the processes under study, but it is neither necessary nor realistic to assume *absolute* stability here. Yet standard econometric models start from the premise of time-invariant coefficients. The more appropriate way is to estimate models with time-varying coefficients. The VC program is designed for the purpose of estimating time-varying parameter models. The classical regression model, which presupposes constant coefficients, is just a special case and may come about because the data say so. We need not assume stability from the outset, though.

---

[1] See Schlicht (1978; 1997) for a mathematical statement and Schlicht (1977; 1985) for a comprehensive discussion of *ceteris paribus* analysis and a justification of using models with time-varying coefficients.

The assumption made is that the behavioral parameters of the model move randomly over time, but move *relatively* slowly, as compared to the speed of the endogenous variables. This means that the coefficients must be auto-correlated. As we do not want to explain the wanderings of these data in our analysis, because they are data, it seems sensible to assume a random walk here. This assumption assures that all changes in the coefficients are data-driven, rather than induced by the structure of the model. The assumption which really enters and links the past with the future - and we need an assumption of this kind - is that the stability, or variance, of any coefficient be time-invariant, and we estimate this stability in order to draw conclusions about the parameters of interest.[2]

3   THE MODEL

1   *A Model With Coefficients Generated by a Random Walk*

The regression model considered in the following is

$$y_t = a_t' \cdot x_t + u_t, \quad t = 1, 2, \dots \ , T, \quad u_t \sim N(0, \sigma^2), \quad y_t, \in \Re, \quad x_t, \, a_t \in \Re^n \quad (1)$$

with y_t and x_t denoting the observations at time t, and T denoting the number of time periods covered by the observations. All vectors are taken as column vectors unless a transposition is indicated by an apostrophe "'". The vector $y_t = (y_1, y_2, \dots y_T)$ denotes the observations of the dependent variable. The vector $x_t = (x_{1t}, x_{2t}, \dots x_{nt})$ denotes the vector of the n independent variables observed at time t. The disturbance u_t is assumed to be normally *i.i.d.* with mean zero and variance σ^2.

The coefficients a_t in equation (1) are assumed to be generated by a random walk:

$$a_t = a_{t-1} + v_t, \quad t = 2, 3, \dots, T, \quad v_t \sim N(0, \Sigma), \quad v_t \in \Re^n \quad (2)$$

with

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

The classical regression model is covered as the limiting case $\Sigma \to 0$ where the coefficients remain invariant over time. The case of interest is that the coefficients move slowly over time (that Σ is small), but in the end the volatility

---

[2] A detailed elaboration of the underlying "Isolation Principle" and its entailments for econometrics is provided in Schlicht (1985).

of the coefficients is to be determined empirically rather than assumed in advance.

The model (1), (2) can be reformulated in matrix notation as

$$y = X \cdot a + u, \qquad u \sim N(0, \sigma^2 I) \tag{3}$$

$$P \cdot a = v, \qquad v \sim N(0, S), \qquad S = I \otimes \Sigma \tag{4}$$

with

$$y, \; u \in \Re^T, \; v \in \Re^{(T-1)n}$$

$$X = \begin{pmatrix} x_1' & & & 0 \\ & x_2' & & \\ & & . & \\ & & & . \\ 0 & & & x_T' \end{pmatrix} \in \Re^T \times \Re^{Tn}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_T \end{pmatrix} \in \Re^{Tn}$$

and

$$P = \begin{pmatrix} -I & I & & & 0 \\ & -I & I & & \\ & & . & . & \\ & & & . & . \\ 0 & & & -I & I \end{pmatrix} \in \Re^{(T-1)n} \times \Re^{Tn}$$

2 Estimation

For given observations y and X and given variances σ^2 and $\sum$, the conditional distribution of the coefficients a is normal with expectation

$$\hat{a} = (X'X + \sigma^2 P' S^{-1} P)^{-1} X' y \tag{5}$$

(see Schlicht 1985; 1988).

The covariance matrix of the estimates of the coefficients can be obtained from Schlicht (1988, eq. 40) as

$$E\left((\hat{a} - E(a))(\hat{a} - E(a))'\right) = E\left((\hat{a} - a)(\hat{a} - a)'\right)$$

$$= \sigma^2 (X'X + \sigma^2 P' S^{-1} P)^{-1} \tag{6}$$

The square roots of the main diagonal elements of (6) are given along with the estimates $\hat{a}$. As an option, the entire matrix (6) is calculated.

The variances are estimated by a generalized method of moments procedure and plugged into (5). The algorithm finds variances σ^2 and Σ such that the variances of the estimated errors $\hat{u}_t$ and $\hat{v}_{it}$ are equalized with their expectations, see Schlicht (1988). These estimators for σ^2 and Σ are asymptotically equivalent to the corresponding maximum-likelihood estimators.

4   The Program

When the program is called, a window opens that displays a number of options and controls. On these later. The main buttons are "Get input!," "Info," and "Cancel." The "Cancel" button stops the program, the "Info" button opens the information file in PDF format, and the "Get input!" button starts the program. A window opens that allows for selecting an input file.

1   Input

The program requires input data in the .CSV format (comma separated values).[3] (An example input file (OkunGER.csv) is provided that contains data for estimating Okun's law for Germany.) The first row of the input file is reserved for data description - one label for each cell. No cell should be left empty. The number of cells in the first row tells the program how many cells to expect in the subsequent rows. The following rows give the data:

Each row contains

-cell 1: time index, such a year or month, ignored by the program, but must be non-empty and can't contain a comma,

-cell 2: dependent variable (typically $y(t)$),

-cell 3: first independent variable ($x_0(t)$, typically constant term, $= 1$),

-cell 4: second independent variable ($x_1(t)$),

-cell 5: third independent variable ($x_2(t)$),

etc...

2   Output

If the input file is OK, the iteration starts. The number of iterations is displayed in a window, as well as the achieved precision (measured by the maximum relative deviation of expected from theoretical variances). If a very high precision is reached or if non-computable values occur, the program stops. If the program is interrupted (because precision is good enough, for instance), or stops on its own, it asks for the name of the output file. This file will be written in .CSV format as well

---

[3]Note that the English number format is presupposed, with a decimal dot denoting decimals, and a comma separating values. Other formats (the German one, for instance) use other conventions and need to be transformed accordingly.

The output file will contain the original data and, in additional colums, the associated calculated coefficients (a_var), with their standard errors (stderr_var), lower bound (lb_var = a_var - 2*stderr_var), and upper bound (ub_var = a_var + 2*stderr_var).

At the end a list of the initial variances and the computed variances are given.

-cell 1: variance of disturbance in the equation,

-cell 2: variance of the first coefficient,

-cell 3: variance of the second coefficient,

etc...

Further, information on initial variances, blocked variances and achieved precision is provided.

3 *Options*

The program allows for three options:

Estimate all variances. This is the standard case. If this is selected, all variances are estimated.

Estimate selected variances only. In this case a new window opens after the input data are read. Here you can force selected variances to the initial value given in the input file and adjust only the other variances. The output file will contain information on the enforced variances and on the corresponding empirical variances in this case. This option ought to be used in conjunction with the control 'obtain initial variances from input' described below.

Don't estimate variances. In this case, the coefficients are computed for the variances given in the input file according to eq. (5). The corresponding empirical variances are given in the output file.

4 *Controls*

Three checkboxes are provided:

Slow down adjustment. This slows down the adjustment during iterations and may improve convergence to cope with some stability issues.

Iterate until aborted. This option will disable automatic termination of the iteration process. You will need to close the info window in order to terminate the iteration process in this case.

Calculate complete covariance matrix. If this is selected, the program will ask for a further output file name for storing the covariance matrix (6).

Obtain initial variances from input. If this is selected it is assumed that the second rowof the input file contains in the first cell a label (dummy) , and in the following cells initial values for the variances $\sigma^2, \sigma_1^2, \sigma_2^2, \sigma_3^2$.... This permits checking for uniqueness of the solution. If this remains unchecked, all initial variances are set to one.

5   A NOTE ON KALMAN-BUCY
FILTERING

It is nowadays usual to employ Kalman-Bucy filtering in order to estimate models with time-varying coefficients like (1), (2). I have proposed this technique myself some time ago (Schlicht 1977), but have convinced myself that the estimation technique described in Schlicht (1988) and implemented in the VC program here is preferable because it is statistically more efficient and numerically more transparent and straightforward. The main issue in both Kalman-Bucy filtering and in my approach relates to the estimation of variances. But given the variances, the VC program is clearly statistically superior because it is a two-sided filter, whereas the Kalman-Bucy filter is, by construction, one-sided. As a consequence, the estimate of the state of coefficients a_t at time t will depend in the Kalman-Bucy case only on past and current, but not on future, observations - that is, on $(y_\tau, X_\tau)_{\tau=1,2,...,t}$. But the current state a_t is correlated with the future states, and future observations will provide information about the current state. Hence it is inefficient to use a one-sided filter. The method implemented in VC uses all observations for estimating each coefficient - that is, on $(y_\tau, X_\tau)_{\tau=1,2,...,T}$. Further, and as a consequence of one-sided filtering, the initial state a_1 must be estimated somehow, as it is needed for starting the updating process. In short, the Kalman-Bucy filter gives an estimate

$$\hat{a}_t = E\left(a_t \,|(y_\tau, X_\tau)_{\tau=1,2,...,t}, \; a_1, \; \sigma^2, \; \Sigma\right) \quad t = 2, 3, ..., T$$

that throws away the information contained in $(y_\tau, X_\tau)_{\tau=t+1,...,T}$ and requires the information a_1 we don't have, wheras the VC filter gives just what we want:

$$\hat{a}_t = E\left(a_t \,|(y_\tau, X_\tau)_{\tau=1,2,...,T}, \; \sigma^2, \; \Sigma\right) \quad t = 1, 2, ..., T$$

The second issue relates to the estimation of variances. This is usually done by maximum-likelihood methods in the Kalman-Bucy framework. This can be done in the VC context as well; it is, in fact, much more easy to implement here. My experience with maximum-likelihood estimation for smaller samples was, however, entirely disappointing.[4] So I developed the GMM estimator, which is part of the VC program and which seems to work much better in smaller samples, has a nice intuitive interpretation and is asymptotically equivalent to the maximum-likelihood estimator. A comparably transparent and computationally manageable method seems not available for the Kalman-Bucy method. It is, however possible, to use the VC program just as a decomposition algorithm for given variances and to calculate the variances by other means, *viz.* by constructing a target function (like maximum likelihood) from the estimates, their variances and their co-variances, as all these numbers are can be calculated by VC for any given choice of σ^2 and Σ.

---

[4]See the example in Schlicht (1988).

6   TERMS AND CONDITIONS

All rights reserved. This program is free for non-commercial use. It may be freely used and distributed provided that all the original files are included - in particular this info file. The program is provided as-is, without any warranty or obligation on my part. The user is invited to improve the program. If he distributes his own version, he is required to include a copyright notice which states how the derived program is related to the original one. (The source code in C++ for the computations is included in the earlier console version (ver. 1▷0), available on my my web site.)

I would appreciate comments, suggestions, bug reports, and observations on overall performance at schlicht@lmu.de

The most recent version is available on my website www.lrz.de/~ekkehart→downloads.

7   HISTORY OF PROGRAM AND DOCUMENTATION

December 19, 2003. version 4: confidence bands (cefficient ± 2 * stderr) added, label format changed to comply with Gretl and similar econometrics packages. Hopefully the last change!

December 18, 2003. version 4 beta: data format changed slightly. As a standard, no input of initial variances required. To process files prepared for previous versions, check 'obtain initial variances from input'.

December 15, 2003, example file changed to data for Okun's Law for Germany, data source: Statistische Bundesamt.

July 11, 2003, revision of this file

June 27, 2003ᶜ revision of this file

March 7, 2003: corrections of this info file.

March 6, 2003: version 3.1 with covariance matrix output.

February 26, 2003: version 3.0 with new GUI and extended information.

February 12, 2003: version 2.0, improved by Robert Schlicht, output variances, additions to Readme.txt.

February 7, 2003: addition to Readme.txt

February 1, 2003: fist public release, version 1.0

December 2002: translation of the original code into C by Robert Schlicht

September 25, 1987: original program written in an obsolete dialect by Ekkehart Schlicht

8   ACKNOWLEDGEMENTS

Robert Schlicht has ported the original program into C and wrote the input-output modules.

Johannes Ludsteck made several helpful suggestions.

9

Tobias Lampe provided the data.

9   REFERENCES

Keynes, John Maynard, 1973, *The General Theory and After, Part II: Defense and Development, The Collected Works of John Maynard Keynes*, vol. XIX, London: Macmillan.

Schlicht, Ekkehart, 1978, "Die Methode der Gleichgewichtsbewegung als Approximationsverfahren" ("Approximation by Moving Equilibrium"), in: E. Helmstädter (ed.): *Neuere Entwicklungen in den Wirtschaftswissenschaften*, Berlin: Duncker und Humblot, 293-305.

—, 1977, *Grundlagen der ökonomischen Analyse,* Reinbek: Rowohlt.

—, 1985, *Isolation and Aggregation in Economics,* Berlin-Heidelberg-New York-Tokyo: Springer-Verlag.

—, 1988, "Variance Estimation in a Random Coefficients Model," paper presented at the Econometric Society European Meeting Munich 1989, available at http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/paper/schlicht_variance-estimation-in-random-coeff-model.pdf.

—, 1997, "The Moving Equilibrium Theorem Again," *Economic Modelling*, 14(2), 1997, 271-278.

—, 2003, Schlicht's VC Program Package, available at http://www.semverteilung.vwl.uni-muenchen.de/mitarbeiter/es/linkpages/linkpage-vc-package.htm.