Time-varying Coefficients Estimation

Ekkehart Schlicht

1 Purpose

The program estimates the time-varying coefficients $a_t' = (a_{1,t}, a_{2,t}, \ldots, a_{n,t})$ of the linear regression

$$y_t = a_t' x_t + u_t, \;\; t = 1, 2, \ldots, T \tag{1}$$

with regressors $x_t' = (x_{1,t}, x_{2,t}, \ldots, x_{n,t})$ and the dependent variable y_t. The time series x_t and y_t denote the observations at time $t = 1, 2, \ldots, T$. The disturbance u_t is a random variable with mean zero and variance σ^2. The coefficients $a_{i,t}$ are assumed to be generated by a random walk

$$a_{i,t+1} = a_{i,t} + v_{i,t}, \;\; i = 1, 2, \ldots, n \tag{2}$$

with disturbances $v_{i,t}$ of mean zero and variances σ_i^2. The idea is that the coefficients are moving only slowly - they are highly auto-correlated. Their variability is given by the variances σ_i^2. The case of one or more time-invariant coefficients is covered as the limiting case that the corresponding variances σ_i^2 are zero.

The program estimates the conditional expected values of the coefficients $a_{1,t}, a_{2,t}, \ldots, a_{n,t}$ and the variances $\sigma^2, \sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, given the observations $x_{i,t}$ and y_t for $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, n$. The method used is the VC method described in Schlicht (2021).

2 Installation

Install TVC by opening the Gretl window. Select *File → Function packages → On server,* search for “TVC”, right-click and select “Install”. A window pops up that asks to attach the function to Gretl’s menu. Confirm with “Yes”. For generating graphics, the program needs the program “TVCplots” that provides a configurable back-end and is installed in the same way as TVC.

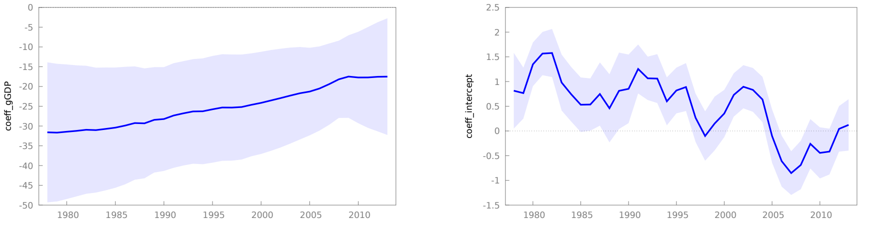
Version 2.0 of December 1, 2023. The author is at the Ludwig-Maximilians-Universität München, Germany and can be contacted at schlicht@lmu.de.

3 Try the GUI

In Gretl's main window, select the *fx* button icon (the fifth at the bottom bar), search for TVC, right click and select "sample script". Follow the instructions given there. This will load time-series for estimating Okun's law for Germany (dU, gGDP, intercept) and performs the estimation of the time-varying coefficients for Okun's law for Germany. To open the graphical user interface (GUI), follow the instructions given at the end of the output. Select "Apply" or "OK" to run the program and select "Help" for explanations and input requirements.[1]

Running the program adds the list of estimated time series for the time-varying coefficients and their standard errors, denoted by coeff_gGDP, stderr_gGDP, coeff_intercept, and stderr_intercept to the list of time-series in Gretl's main window.

If the option "plot results" is selected, the function TVCplots is called. It plots time series with confidence bands (Figure 1). As can be seen, the Okun slope is negative and decreases in absolute value over time while the intercept (the constant term of the Okun relationship at a given point in time) follows the business cycle with a downward trend.[2]



(a) slope (b) intercept

Fig. 1: Estimated time varying coefficients.

4 Initial Variance Ratios

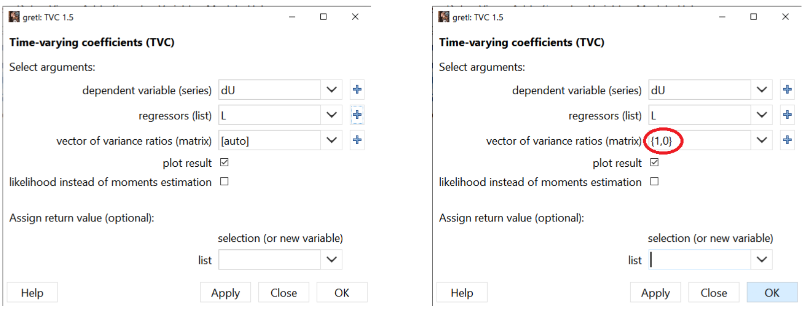
The variance ratios are the inverses of the smoothing weights, see equation (4). A high variance ratio indicates a volatile coefficient, whereas a low variance ratio stands for a more time-invariant coefficient. The solution is obtained by Gretl's BFGScmax-routine that starts with the initial variance ratios. If TVC is called without supplying a vector of initial variance ratios for the regressors, the default values of 1 for each initial variance ratio are assumed (Figure 2). You can supply other initial variance ratios by simply typing in the initial variance ratios as a vector in the initial variance ratio box (Figure 2b).

The initial variance ratios induce the following:

- A positive initial variance ratio simply sets the starting point for the iterations for the respective coefficients. Default is $\{1, 1, \ldots, 1\}$, that is 1 for all coefficients. Trying other variance ratios may be used to check uniqueness of the result.

[1] The data are taken from a study by Jalles (2018) on the topic of Okun's Law that covers many countries.

[2] The time-series intercept in the data set OkunGer.gdt equals Gretl's const which is (1,1,... ,1). As "const" appears inappropriate as a name for a term that is time-varying, the term "intercept" has been chosen.



(a) Default initial variance ratios. This gives the result depicted in Fig. 1.

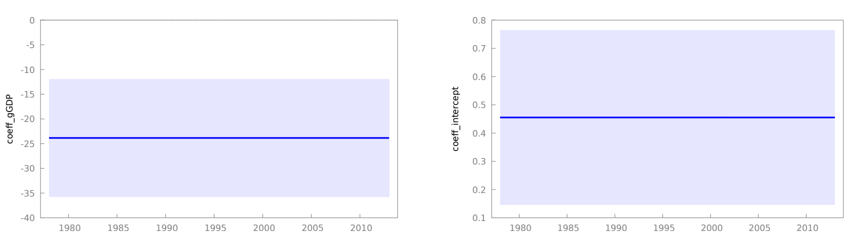
(b) Assigned initial variance ratios of {1,0}. This gives the result depicted in Fig. 3.

Fig. 2: Select initial variance ratios

- An initial variance ratio of zero tells the program to treat the respective coefficient as time-invariant.
- A negative initial variance ratio tells the program to keep the variance ratio for the corresponding coefficient constant at the absolute value of the variance ratio given. If all initial variance ratios are negative, the data are just filtered with the absolute values of these ratios.[3]

The purpose of selecting initial variance ratios is to determine which coefficients are to be taken as time-variant and which are to be considered as time-invariant, *i.e.* choosing for each regressor either the first or the second of the above three cases. If initial variance ratios are set negative, this allows filtering with the absolute values of these variance ratios.

The output obtained, for example, by setting the initial variance ratio for the coefficient of gGDP to *1* and for the intercept to *0* – that is, allowing the slope to vary over time and to take the intercept as time-invariant – gives the result depicted in Fig. 3. It can be seen that the suppression of the influence of the business cycle by forcing the intercept to remain



(a) slope

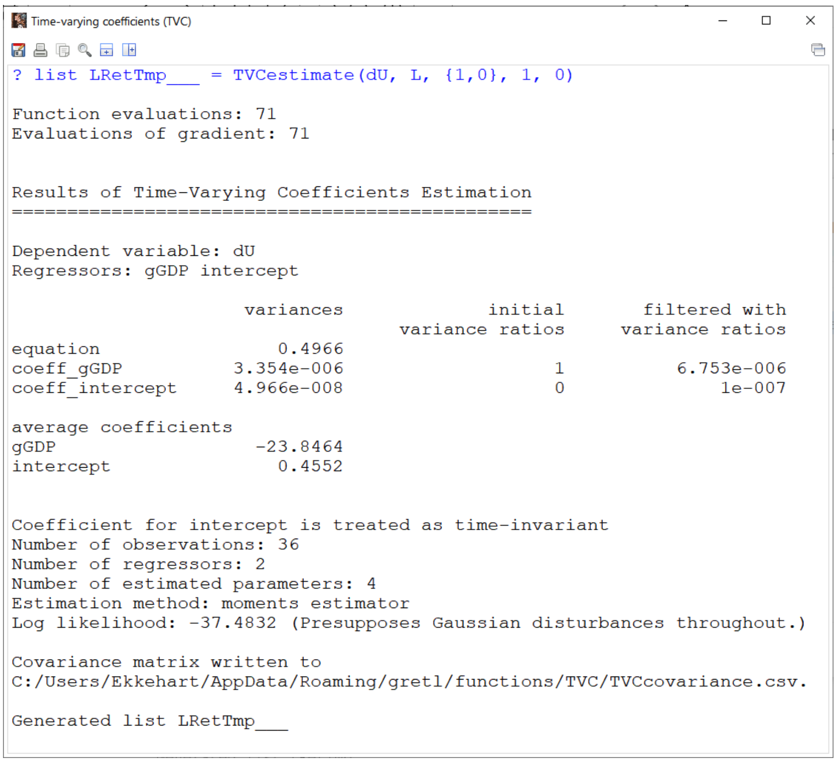
(b) intercept

Fig. 3: Estimated time varying slope with time-invariant intercept.

[3] The program uses the *inverse* variance ratios $\gamma_i = \sigma^2/\sigma_i^2$, see (3) and (4) below. For a variance ratio of zero, γ_i would be infinite. Instead a variance ratio of *1.e-7* is used which implies an inverse variance ratio of *1.e7* and puts a very high, albeit not infinte, penalty on any variability of the corresponding coefficient. Other values, like *1.e-10*, can be tried by entering corresponding negative values, *e.g. -1.e-10,* as variance ratios.

time-invariant reduces the average size of the slope and practically removes its upward trend.

Further, the output screen (Fig. 4) provides, as far as applicable, information about which coefficients are taken as time-invariant and which variance ratios are fixed, about the number of observations and regressors, and the logarithmic likelihood in case the user wishes to calculate an information criterion like AIC or BIC. [4]



```
Time-varying coefficients (TVC)
? list LRetTmp___ = TVCestimate(dU, L, {1,0}, 1, 0)

Function evaluations: 71
Evaluations of gradient: 71


Results of Time-Varying Coefficients Estimation
===============================================

Dependent variable: dU
Regressors: gGDP intercept

                     variances            initial      filtered with
                                   variance ratios   variance ratios
equation                0.4966
coeff_gGDP           3.354e-006                  1        6.753e-006
coeff_intercept      4.966e-008                  0            1e-007

average coefficients
gGDP                  -23.8464
intercept               0.4552


Coefficient for intercept is treated as time-invariant
Number of observations: 36
Number of regressors: 2
Number of estimated parameters: 4
Estimation method: moments estimator
Log likelihood: -37.4832 (Presupposes Gaussian disturbances throughout.)

Covariance matrix written to
C:/Users/Ekkehart/AppData/Roaming/gretl/functions/TVC/TVCcovariance.csv.

Generated list LRetTmp___
```

Fig. 4: Output screen

For completeness, the address of the covariance matrix is given in case the user wants to have a look. This matrix gives the expectation of $(a-\hat{a})\cdot(a-\hat{a})'$ where $a = \operatorname{diag}(A)$ and $\hat{a} = \operatorname{diag}(\hat{A})$ with A as the matrix of true coefficients and $\hat{A}$ as the matrix of estimated coefficients as in equation (5) below; see equation (2.28) in Schlicht (2021).

The option "likelihood estimation rather than moments es timation" is provided for completeness but is not recommended, as it tends to produce corner solutions - extreme values for the variance ratios and the variance - in shorter time series. If selected, a warning is issued in this regard.

[4] These criteria require to specify the sample size and the number of estimated parameters. The sample size is given by the number of observations, As to the number of estimated parameters, this varies with the choice of initial variance ratios, see Section 2. In case that all parameters are declared as time-invariant, the parameters include the coefficients of the n regressors and the variance in the equation, that is, the number of regressors plus 1. If $m \leq n$ regressors are taken as varying over time, one additional variance or variance ratio for each of them must be estimated. This amounts to $n+m+1$ estimated parameters (4 in the example depicted in Figs. 2b and 4). See Section 3.1 in Schlicht (2021) for details.

Calculations involving long time series and many regressors may take a while. In such cases, the current state of estimation is reported by a self-actualizing step monitor (Fig. 5).

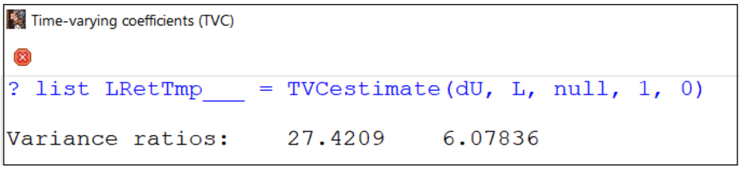


Fig. 5: For long calculations, the current state of calculations is displayed.

5 Running TVC in a script

The TVC function declaration is

```
function list TVCestimate(series S             "dependent variable",
                          list   L             "regressors",
                          matrix ratios[auto]  "vector of variance ratios",
                          bool   plt[1]        "plot result",
                          bool   like[0]       "likelihood instead of moments estimation")
```

with the arguments

argument	*example*
S	*S = dU*
L	*L=gGDP intercept*
ratios	*{1,0}* *(coefficient of gGDP time-varying, coefficient of intercept constant)*
plt	*1* *(1: plot, 0: don't plot)*
like	*0* *(0: moments estimation, 1: maximum likelihood estimation)*

The first two arguments are mandatory, the remaining arguments are optional. If the function is called with the mandatory arguments only, *i.e.* $TVC(S,L)$, the remaining arguments are filled in with their defaults. If some arguments are inserted and others are left blank, as in $TVC(S,L,,0))$, the selected arguments are set to the chosen values ($plt = auto = \{1,1\}$ for the variance ratios and 0 for *don't plot*). The remaining arguments are set to their defaults (*like=0* for *moments estimation*).

6 Files in the TVC Package

The package “TVC.zip” contains the following files

TVC.gfn	TVC functions package for Gretl
TVC.pdf	TVC documentation

and a folder “examples” with

OkunGer.gdt	example data file

7 License

8 Changelog

Version 1.2: Likelihood calculation corrected, step monitor added.
Version 2: Plot routines made configurable and branched out to the program TVCplots. The sample script focuses now on the GUI. Help in the GUI window is provided with Markdown formatting, and documentation is adapted.

Acknowledgements Riccardo Lucetti has suggested to make the VC method available for Gretl. Sven Schreiber of the Gretl team kindly provided the foundation of the TVC.gfn function package by writing a Hansl script for a previous version and gave valuable advice on coding problems I ran into as a novice Hansl coder. I thank also João Tovar Jalles for permission to use the data for Germany from his study (Jalles 2018) as an example.

References

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.

Schlicht, Ekkehart (2021). "VC - A method for estimating time-varying coefficients in linear models". In: *Journal of the Korean Statistical Society* 50, 1164–1196. URL: https://doi.org/10.1007/s42952-021-00110-y.

Appendix

Notes on Computation

A Estimators

The estimators for the variances are moments estimators. With given variances, the time-paths of the coefficients are determined by minimizing the weighted sum of squares

$$Q = \sum_{t=1}^{T} u_t^2 + \sum_{t=2}^{T} \sum_{i=1}^{n} \gamma_i \left(a_{i,t} - a_{i,t-1}\right)^2 \tag{3}$$

with smoothing weights γ_i given by the inverse variance ratios

$$\gamma_i = \frac{\sigma^2}{\sigma_i^2} \ , \quad i = 1,2,\ldots n. \tag{4}$$

The matrix of coefficients

$$A = \begin{pmatrix} a_{1,1} & a_{2,1} & .. & a_{n,1} \\ a_{1,2} & a_{2,2} & .. & a_{n,2} \\ . & . & & \\ & & & \\ a_{1,T} & a_{2,T} & .. & a_{n,T} \end{pmatrix} \tag{5}$$

that minimizes the sum of squares Q gives the expectation of the coefficients $a_{i,t}$ for given observations and given variance ratios. The VC method calculates estimates for the variances and variance ratios by a moments estimator which coincides asymptotically with a maximum-likelihood estimator in case the disturbances are Gaussian.

B The Criterion Functions

The criterion functions used for finding the solution are

$$\mathscr{C}_M = \log\det M + (T-1)\sum_{i=1}^{n}\log\sigma_i^2 - T(n-1)\log\sigma^2 + \frac{1}{\sigma^2}Q \tag{6}$$

$$\mathscr{C}_L = \log\det\left(PMP'\right) + (T-1)\sum_{i=1}^{n}\log\sigma_i^2 - ((T-1)n-T)\log\sigma^2 + \frac{1}{\sigma^2}Q \tag{7}$$

where (6) refers to the moments estimator (with the constant $\log\det\left(PP'\right)$ dropped) and (7) gives the maximum likelihood criterion, as given in equations (3.12) and (3.13) in Schlicht (2021). The system matrix M is, according to equation (2.24) *ibid.*, a function of the variance ratios (4) alone and independent of the variance σ^2. Likewise the sum of squares Q is, according to (3), a function of the variance ratios alone. The variances of the coefficients can be expressed as $\sigma_i^2 = \gamma_i\sigma^2$. In consequence, the criteria (6) and (7) can be written as functions of the variance ratios γ and the variance σ^2:

$$\mathscr{C}_M = \log\det M + (T-1)\sum_{i=1}^{n}\log\gamma_i + (T-n)\log\sigma^2 + \frac{1}{\sigma^2}Q \tag{8}$$

$$\mathscr{C}_L = \log\det\left(PMP'\right) + (T-1)\sum_{i=1}^{n}\log\gamma_i^2 + T\log\sigma^2 + \frac{1}{\sigma^2}Q \tag{9}$$

Taking derivatives with respect to σ^2 gives

$$\frac{\partial\mathscr{C}_M}{\partial\sigma^2} = \frac{1}{\sigma^2}(T-n) - \frac{1}{\sigma^4}Q \tag{10}$$

$$\frac{\partial\mathscr{C}_L}{\partial\sigma^2} = \frac{1}{\sigma^2}T - \frac{1}{\sigma^4}Q. \tag{11}$$

A necessary condition for a minimum with regard to σ^2 is that these expressions are zero. This implies that $\sigma^2 = \frac{Q}{T-n}$ for the moments estimator and $\sigma^2 = \frac{Q}{T}$ for the likelihood estimator, respectively. Plugging these into (8) and (9) and dropping constants yields the concentrated criteria

$$\mathscr{C}_M^c = \log\det M + (T-1)\sum_{i=1}^{n}\log\gamma_i + (T-n)\log Q$$

$$\mathscr{C}_L^c = \log\det\left(PMP'\right) + (T-1)\sum_{i=1}^{n}\log\gamma_i + T\log Q$$

that involve only the variance ratios γ. These are used in the package to determine the variance ratios. Because a maximization routine is used, the criteria that are to be minimized are applied with a negative sign. All further values are derived from the variance ratios so obtained.

C LogLikelihood

For log-likelihood given in the output is defined by equation (3.10) in Schlicht (2021).