



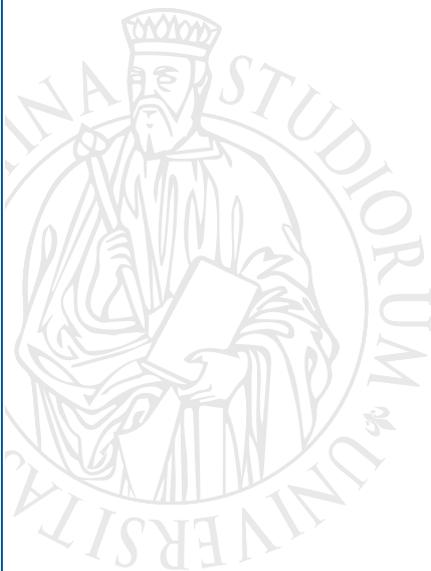
UNIVERSITÀ
DEGLI STUDI
FIRENZE

DISIA

DIPARTIMENTO DI STATISTICA,
INFORMATICA, APPLICAZIONI
"GIUSEPPE PARENTI"

**Indirect estimation
and econometrics exams:
how to live a round life**

Giorgio Calzolari



**DISIA WORKING PAPER
2015/01**

© Copyright is held by the author(s).

Indirect estimation and econometrics exams: how to live a round life*

Giorgio CALZOLARI[§]

January 2015

ABSTRACT

The use of Monte Carlo methods to generate exam data sets is nowadays a well-established practice among econometrics examiners all over the world. Its advantages are well known: providing each student a different data set ensures that estimates are actually computed individually, rather than copied from someone sitting nearby. The method however has a major fault: initial “random errors”, such as mistakes in downloading the assigned dataset, might generate downward bias in student evaluation. We propose a set of calibration algorithms, typical of indirect estimation methods, that solve the issue of initial “random errors” and reduce evaluation bias. Ensuring round initial estimates of the parameters for each individual data set, our calibration procedures allow the students to determine if they have started the exam correctly. When initial estimates are not round numbers, this random error in the initial stage of the exam can be corrected for immediately, thus reducing evaluation bias. The procedure offers the further advantage of rounding markers life by allowing them to check round numbers answers only, rather than lists of numbers with many decimal digits.

KEYWORDS: Indirect estimation, calibration procedure, round number estimates, econometrics exams.

*This work was presented at *Abbey Hegne* during the Conference on *Indirect Estimation Methods in Finance and Economics*, organized by the Universität Konstanz on 30-31 May 2014. I am grateful to the conference participants and to its organizers, Roxana Halbleib and David Veredas, for their comments and suggestions.

There is a long list of colleagues and friends who over the years verified the compatibility of the dataset generated by my algorithm with the most popular econometric packages: Alessandra Amendola, Michela Baccini, Matteo Barigozzi, Erich Battistin, Christiane Baumeister, Luigi Benfratello, Monica Billio, Silvia Bozza, Salvatore Bruno, Christian Brownlees, Massimiliano Caporin, Massimiliano Cecconi, Elena Cefis, Fabrizio Cipollini, Michele Costola, Gabriele Fiorentini, Giampiero M. Gallo, Monica Gentile, Andrea Giacomelli, Roberto Golinelli, Lara Lusa, Riccardo “Jack” Lucchetti, Laura Magazzini, Alessandra Mattei, Giovanni Millo, Laura Neri, Giulio Palomba, Giuseppe Storti, Umberto Triacca, Margherita Velucchi; I am really grateful to all of them. Malvina Marchese kindly tried to improve my *poor* use of the English language in this paper. Francesca Di Iorio played a crucial role in checking the *robustness* of the algorithms (i.e. the invariance of the results) to all the last releases of the most popular *commercial* software packages in the last 15 years. Finally, without the encouragement of Enrique Sentana these algorithms would never have been in a paper, but *only used* (and *suffered*) by the graduate and undergraduate students of Firenze, Pisa and Venezia Universities, and by more than a thousand summer school PhD students of the Italian Interuniversity Econometrics Center (CIDE). Needless to say the full responsibility for the contents of the paper is mine.

Financial support of the project MIUR (PRIN): MISURA - Multivariate Models for Risk Assessment is gratefully acknowledged.

[§]University of Firenze, Department of Statistics, Informatics, Applications “G.Parenti”, Viale Morgagni 59, I-50134 Firenze, Italy. *Fax:* (+39) 055 4223560. *E-mail:* calzolar@disia.unifi.it.

1 Introduction

The aim of this paper is to improve the quality of life of econometrics examiners through an *unusual* application of indirect estimation calibration algorithms.

It is a common practice for many examiners to include some practical applications of econometric techniques in their exam papers. This might entail OLS, 2SLS, 3SLS, NLLS, or MLE estimation as a preliminary step followed by some hypothesis testing, goodness of fit analysis, forecasting and maybe scenarios simulation. As a result of the strong and natural *cheating* attitude of students all over the world, Monte Carlo methods are now widespread. Thanks to their ability to generate quickly a large number of different datasets, they allow the examiners to provide each student with a different data set, and different expected estimation results. Should a student copy another student's estimation results, she/he would be immediately spotted and severely punished. The weakness of this evaluation method lies in its sensitivity to random mistakes. For example, the overemotional attitude of students during exams implies that data import mistakes are quite common even among first class students. Furthermore students display a natural tendency to fall into local maxima/minima instead of achieving the global one while programming iterative numerical algorithms for M estimators. The occurrence of such mistakes can induce a serious bias in the evaluation of students performance. Wrong estimation results and conclusions might arise not only as a consequence of inadequate exam preparation, but also as the outcome of some initial mistakes. This might generate either upward or downward bias in students evaluation. Bias reduction can only be achieved via exact replication of the computational steps which lead to the wrong result. This technique is extremely time consuming and it relies on an examiner ability to replicate an overemotional student computational mistakes. As the students sample size increases to infinity, the effectiveness of such bias reduction method converges to zero with rate of convergence depending on the examiner's good mood.

It has been suggested in the literature that providing the students with some of their expected estimation results could solve this issue, however it might also generate upward bias in the evaluation of poorly-prepared students.

This paper introduces a bias elimination method based on indirect estimation calibration algorithms. It is prompted by the consideration that if each individual data set is generated with indirect estimation calibration algorithms, it *can produce round numbers estimates*. Thus, while retaining the advantage of individuality of the data sets, this method enables students themselves to be aware of any computational mistake in their exams and to get rid of it. Obtaining round numbers estimates (that is numbers with a convenient amount of zeros, say for example 5.00000, or 10000.00000, or 0.20000), the student knows that no computational mistake occurred, otherwise she/he is aware of a mistake and of the necessity of computing the initial estimates again correctly. At the same time the student cannot get any *benefit* from watching at other students' estimation results, that could be numbers like 4.00000, or 2500.00000, or 0.10000. This method offers the further advantage of speeding up the marking process for examiners: checking scripts with round number estimates only is faster and by far less stressful than checking long lists of decimal digits numbers.

The method can be applied to a large variety of econometric models. The following sections provide some examples for the most commonly used models and describe in some details the calibration algorithms for each of them. For each of the models, the simple rule is always the following one: if parameters have been estimated as requested by the examiner, they *must be* round numbers (different

for each student, drawn at random from a prefixed *menu* of round numbers, and of course meaningful from the economic viewpoint in terms of sign and magnitude).

2 Static linear regression model and OLS

A dataset is given to each student, together with the equation

$$y_i = \delta_1 + \delta_2 x_i + \delta_3 d_i + \varepsilon_i \quad \varepsilon_i \text{ i.i.d. } (0, \sigma^2) \quad i = 1, 2, \dots, n$$

that must be estimated by OLS. The dataset may have a small, medium or large number of observations. The example here is with a large number of observations, so that a similar dataset can be used also in the following section, where a moderately large dataset is necessary.

i	y	x	d
1	588.3566145963907	1.100	0.000
2	466.2769870928506	0.900	1.000
3	547.8717011872933	0.500	0.000
...
398	612.3154009309173	0.900	0.000
399	583.5503587812454	0.500	1.000
n=400	624.8176359396298	1.000	1.000

(there is an *unpleasant* excess of decimal digits for the dependent variable; the compromise between excess of digits in the dataset and precision of the *round estimates* will be discussed in section 7.1).

For this particular case there is no need of the calibration procedures typical of the indirect estimation methods. It is enough to generate a vector $u \sim N(0, I_n)$, then project it orthogonally to all regressors and multiply it by a prefixed σ (for instance = 30.00000; also the correction for degrees of freedom must be taken into account). This produces a *tricky* pseudo-random vector

$$\varepsilon = \sigma (I_T - X(X'X)^{-1}X') u;$$

being orthogonal to all regressors, it can be used in the equation to simulate the vector y , and OLS estimation will leave unchanged the coefficients. With a simple rescaling of ε that takes into account n and k for the degrees of freedom correction, also the estimation of σ would coincide with the desired value.

So if we fix

$$\delta_1=500.00000 \quad \delta_2=100.00000 \quad \delta_3=2000.00000 \quad \sigma=30.00000$$

to simulate the data, OLS estimates of coefficients and σ will be *exactly equal* to the pre-fixed δ and σ .

There is no need of any iterative calibration technique; the dataset is produced in *one-shot* and the student that obtains estimates like those displayed above is sure to have done all the first computations correctly.

3 Regression model with conditional heteroskedasticity

The model is quite similar to the previous one, also with a moderately large dataset, but data are time series rather than cross-section data.

$$y_t = \delta_1 + \delta_2 x_t + \delta_3 d_t + \varepsilon_t \quad t = 1, 2, \dots, T.$$

The error terms are affected by conditional heteroskedasticity and are supposed to be well modelled as a Gaussian GARCH(1,1) (e.g. Bollerslev, 1986, Engle, 1982)

$$\varepsilon_t \sim \text{Gaussian GARCH}(1,1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

If the student applies correctly the Gaussian maximum likelihood estimation, regression coefficients should be, for instance,

$$\hat{\delta}_1=500.00000 \quad \hat{\delta}_2=100.00000 \quad \hat{\delta}_3=2000.00000$$

while GARCH parameters should be

$$\hat{\alpha}_0=100.00000 \quad \hat{\alpha}_1=0.10000 \quad \hat{\beta}_1=0.80000$$

thus with unconditional variance

$$\frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\beta}_1}=1000.00000$$

The dataset can be produced with the calibration procedure typical of the indirect estimation techniques applied *correctly*, but without any inferential purpose!

3.1 Calibration procedure and production of the datasets

Using the terminology typical of indirect estimation, the *true* model $y = M(\theta)$ and the *auxiliary* model $y = M(\beta)$ are exactly the same model, apart from the parameter names. It is a case of just-identification, with a strict one-to-one correspondence between parameters. Among the various indirect estimation procedures available in the literature, the one which is applied here is closer to the Gouriéroux et al. (1993), rather than Smith (1993), or Gallant et. al (1996). The length of the simulated data and the *sample size* will always be equal (using symbols familiar to the practitioners of indirect estimation methods, $H = 1$, thus length $HT = T$).

The *observed* y do not exist, thus no estimation of the auxiliary model has to be done on the observed variables. The usual phase of estimation of the auxiliary model on the observed variables (usually producing the pseudo-maximum likelihood estimates $\hat{\beta}$) is replaced by fixing $\bar{\beta}$ at the *target* values (round numbers, meaningful from the economic viewpoint).

The followings are the steps of the procedure.

1) Fix *target* $\bar{\beta}$ for each student: each element of $\bar{\beta}$ is chosen at random from a menu of *meaningful* round numbers (for instance, those of the example above, where in particular the GARCH parameters must satisfy the usual stationarity conditions).

2) Initially choose $\tilde{\theta} = \bar{\beta}$. It is surely a *good* initial choice of the parameters to be used for producing

the simulated data; thus, the Newton-Raphson solution algorithm should not encounter convergence problems.

3) Use $\tilde{\theta}$ to produce simulated data.

4) On the simulated data, estimate the auxiliary model's parameters $\tilde{\beta}$ (that will be obviously $\tilde{\beta} = \beta(\tilde{\theta})$, a function of the parameters used to simulate the data) and compare it with the target $\bar{\beta}$.

5) Calibrate $\tilde{\theta}$ and iterate the process from step 3, till $\tilde{\beta} = \bar{\beta}$. From a mathematical viewpoint, this calibration phase is the solution of the equations system $\beta(\theta) = \bar{\beta}$, where θ are the unknowns, $\bar{\beta}$ have been fixed as desired values, $\beta(\theta)$ is the system of functions implicitly constructed by the estimation of the auxiliary model. Taking advantage of the good initial choice of the unknowns θ , the iterative Newton-Raphson algorithm is applied, providing at the $k + 1$ th iteration $\theta_{k+1} = \theta_k - J_k^{-1}[\beta(\theta_k) - \bar{\beta}]$, with numerical computation of derivatives in the Jacobian matrix $J = \partial\beta(\theta)/\partial\theta'$.

6) Ignore all values used for θ ; only save the last set of simulated variables. These will be given to the student, who has to treat them as *observed* values of the y . When the student estimates parameters of the $y = M(\beta)$ model, the results will be exactly the prefixed *round* $\bar{\beta}$.

3.2 Conditional heteroskedasticity and forecast

For the same model of section 3, $y_t = \delta_1 + \delta_2 x_t + \delta_3 d_t + \varepsilon_t$

$\varepsilon_t \sim$ Gaussian GARCH(1,1)

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

the student also receives the values of the exogenous variables *predicted for tomorrow* (x_{T+1}, d_{T+1}). They must be used to forecast tomorrow's y_{T+1} and the volatility σ_{T+1} .

These are, for example, values that might be produced

$$\hat{\delta}_1=500.00000 \quad \hat{\delta}_2=100.00000 \quad \hat{\delta}_3=2000.00000$$

$$\hat{\alpha}_0=100.00000 \quad \hat{\alpha}_1=0.10000 \quad \hat{\beta}_1=0.80000 \quad \frac{\hat{\alpha}_0}{1-\hat{\alpha}_1-\hat{\beta}_1}=1000.00000$$

$$\hat{y}_{T+1}=2600.00000 \quad \hat{\sigma}_{T+1}=30.00000$$

Round number for \hat{y}_{T+1} is automatically produced, with a simple choice of x_{T+1} and d_{T+1} .

For the volatility, the usual indirect estimation techniques are applied with a *tricky* modification.

The vector θ is augmented with one more *parameter* to be calibrated simultaneously with the previous ones: the last random error term u_T (or ε_T).

The vector β is augmented with one more *target*: the volatility of tomorrow σ_{T+1} .

Now θ has 7 elements: 3 coefficients, 3 GARCH parameters and u_T (or ε_T).

Also β has 7 elements, the last one being σ_{T+1} .

As in section 3.1, θ will be iteratively calibrated till estimated $\tilde{\beta} = \bar{\beta}$. Precision in the pseudo-maximum likelihood estimation of the GARCH parameters is ensured by the analytical computation of first and second order derivatives, as in Fiorentini et al. (1996).

All the values used for θ will be ignored, and only the last set of simulated variables will be saved. These will be given to the student, who has to treat them as *observed* values of the variables.

The student receives also the values of the exogenous variables predicted for the day after tomorrow (x_{T+2}, d_{T+2}) and has to use the model to forecast also y_{T+2} and the conditional variance σ_{T+2}^2 .

In addition to the estimated values displayed above, with a simple choice of x_{T+2} and d_{T+2} , round values could be obtained also for

$$\hat{y}_{T+2} = 2600.00000 \text{ and } \hat{\sigma}_{T+2}^2 = \frac{\hat{\alpha}_0}{1 - \hat{\alpha}_1 - \hat{\beta}_1} + (\hat{\alpha}_1 + \hat{\beta}_1) \left(\hat{\sigma}_{T+1}^2 - \frac{\hat{\alpha}_0}{1 - \hat{\alpha}_1 - \hat{\beta}_1} \right) = 910.00000.$$

No further calibration is required to get these last round values.

4 Logit model

For a problem of *discrete choice*, the student receives the individual dataset and is requested to estimate the logit model

$$prob[y_i = 1] = \frac{\exp(\delta_1 + \delta_2 x_i + \delta_3 z_i + \delta_4 q_i)}{1 + \exp(\delta_1 + \delta_2 x_i + \delta_3 z_i + \delta_4 q_i)}$$

as discussed in econometric textbooks (e.g. Greene, 2008, ch.23).

i	y	x	z	q
1	0.	1.0	8.0	-0.736072071285318
2	1.	2.0	5.0	0.396914342224885
3	1.	4.0	3.0	1.145589102963721
...
198	0.	0.0	4.0	0.080969978915930
199	0.	0.0	2.0	-0.705199321511834
200	0.	1.0	2.0	-1.079599537479198

These are, for example, values that might be obtained from the estimation

$$\hat{\delta}_1 = -3.00000 \quad \hat{\delta}_2 = 0.50000 \quad \hat{\delta}_3 = 0.40000 \quad \hat{\delta}_4 = 1.00000$$

To produce the dataset, there is no need of the calibration technique typical of indirect estimation methods. Analogously to OLS (sect. 2), we take advantage of the orthogonality relationships. But since the dependent variable y_i is *integer*, the variable with many decimal digits will be one of explanatory variables.

5 Linear panel data model with random effects

$$y_{i,t} = \delta_1 + \delta_2 x_{i,t} + \delta_3 z_{i,t} + \delta_4 q_{i,t} + \alpha_i + e_{i,t}$$

i	t	y	x	z	q
1	1	14.6903076350614	1.00	5.00	1.00
1	2	50.7317760110699	1.00	5.00	2.00
1	3	27.9682784991741	1.00	6.00	2.00
1	4	30.5478373176867	1.00	7.00	2.00
....	
99	1	5.00000000000000	1.00	3.00	0.00
99	2	0.3539053628556	0.00	0.00	1.00
99	3	0.00000000000000	0.00	0.00	1.00
99	4	9.6645145817598	0.00	0.00	1.00

The standard hypotheses are assumed (e.g. Baltagi, 1998, sect. 12.2.2, or Wooldridge, 2002, sect. 10.4). The individual effects α_i and the idiosyncratic errors $e_{i,t}$ are assumed to be normally distributed, mutually independent, homoskedastic and without autocorrelations of any order.

When the student applies Gaussian maximum likelihood, estimates would be like

$$\hat{\delta}_1=1.00000 \quad \hat{\delta}_2=3.00000 \quad \hat{\delta}_3=1.00000 \quad \hat{\delta}_4=8.00000$$

$$\hat{\sigma}_\alpha^2=64.00000 \quad \hat{\sigma}_e^2=81.00000$$

To produce the dataset, indirect estimation is applied *correctly*, as in section 3.1. calibrating the 6 parameters.

6 Linear time series model

The student has to deal with one of the following models:

MA(1), MA(2), AR(1), AR(2), ARMA(1,1).

t	y
1	-306.7935808544725
2	-311.9435496768569
3	-291.9642620525142
...
398	-347.2085655352504
399	-376.1035906386596
400	-375.4223070201451

The student is requested to apply, first of all, the Box-Jenkins (1979) identification rules (see, for instance, Granger, 1989, ch. 3). For example, the student realises that autocorrelations of order 1

and 2 are significantly different from zero, while all other autocorrelations are inside the confidence band, while partial autocorrelations are significantly non-zero till some moderately large order. Thus the student should select

$$\text{MA}(2): y_t = \delta_1 + \varepsilon_t + \delta_2\varepsilon_{t-1} + \delta_3\varepsilon_{t-2}$$

and estimate the parameters by maximum likelihood, obtaining, for instance

$$\hat{\delta}_1 = -400.00000 \quad \hat{\delta}_2 = 0.40000 \quad \hat{\delta}_3 = 0.50000 \quad \hat{\sigma}^2 = 900.00000$$

To produce the dataset, indirect estimation procedures are applied *correctly* as in section 3.1, calibrating the δ and σ parameters (here 4 parameters). Their computer implementation is moderately simple and always runs in a smooth and fast way.

Inconveniences may easily arise after completion of the procedure. A first check of coherence is performed on the series produced. Autocorrelations and partial autocorrelations, till a prefixed order, should satisfy the identification rules for that particular model. If *not all* the rules are respected, model and dataset are rejected, a new random choice of a model is performed, and the construction of its corresponding dataset starts again.

One more type of inconvenience may arise when computing the standardized residuals. It frequently happens that they cannot be regarded as a white noise, because the autocorrelation of some order is significantly non-zero (lags till 12 are usually checked). Also in this case, the model and its dataset are rejected, a new random choice of a model is performed, and the procedure starts again.

The number of rejections is quite large (it obviously becomes larger when enlarging the number of lags). Checking correlations till 12 lags, the number of *accepted* models and datasets is about 5% of the models and datasets produced. Nevertheless the procedure is very fast, as the average computation time to produce an acceptable model is about 1 second on a PC with a modern i7 processor.

7 Simultaneous equations

Klein-I model (Klein, 1950) has been used for several decades as the benchmark for experimenting econometric techniques on systems of simultaneous equations. Among many others (very many others indeed), one can cite in the fifties Chernoff et al. (1953); in the sixties Goldberger et al. (1961); in the seventies Hendry (1971); in the eighties Bianchi et al. (1982); till the nineties and later Calzolari et al. (1993), Renfro (2009).

The model is presented in Greene (2008, ch. 10) as

$$\left[\begin{array}{l} C_t = \alpha_{1,1} + \alpha_{1,2}P_t + \alpha_{1,3}P_{t-1} + \alpha_{1,4}(W_t^p + W_t^g) + u_{1,t} \\ I_t = \alpha_{2,1} + \alpha_{2,2}P_t + \alpha_{2,3}P_{t-1} + \alpha_{2,4}K_{t-1} + u_{2,t} \\ W_t^p = \alpha_{3,1} + \alpha_{3,2}X_t + \alpha_{3,3}X_{t-1} + \alpha_{3,4}A_t + u_{3,t} \\ X_t = C_t + I_t + G_t \\ P_t = X_t - T_t - W_t^p \\ K_t = K_{t-1} + I_t \end{array} \right. \begin{array}{l} \text{consumption} \\ \text{investment (net)} \\ \text{private wages} \\ \text{equilibrium demand} \\ \text{profits} \\ \text{capital stock} \end{array}$$

Besides the 6 endogenous variables, the model also includes some exogenous variables: W_t^g = government wages, T_t = business taxes, A_t = *proxy* for bargaining power of labour (or union strength; during the sample period it is a linear time trend, measured as annual deviations from 1931, positive or negative), G_t = government nonwage expenditure..

Each student receives a *modified* model. While accounting equations (or identities, i.e. the last three equations) are always as above, each of the first three behavioural stochastic equations can be modified, with insertion (or cancellation, or modification) of some regressors. Needless to say, each modified equation must be meaningful from the economic viewpoint. For instance, the consumption equation could be without lagged profits

$$C_t = \alpha_{1,1} + \alpha_{1,2}P_t + \alpha_{1,3}(W_t^p + W_t^g) + u_{1,t}$$

or with taxes as additional regressor

$$C_t = \alpha_{1,1} + \alpha_{1,2}P_t + \alpha_{1,3}P_{t-1} + \alpha_{1,4}(W_t^p + W_t^g) + \alpha_{1,5}T_t + u_{1,t}$$

or include lagged value of wages and salaries

$$C_t = \alpha_{1,1} + \alpha_{1,2}P_t + \alpha_{1,3}(W_t^p + W_t^g) + \alpha_{1,4}(W_{t-1}^p + W_{t-1}^g) + \alpha_{1,5}T_t + u_{1,t}$$

or it can be modified in other ways.

The investment equation could be modified with the inclusion of taxes as additional regressor

$$I_t = \alpha_{2,1} + \alpha_{2,2}P_t + \alpha_{2,3}P_{t-1} + \alpha_{2,4}K_{t-1} + \alpha_{2,5}T_t + u_{2,t}$$

or with the inclusion of the lagged dependent variable

$$I_t = \alpha_{2,1} + \alpha_{2,2}P_t + \alpha_{2,3}P_{t-1} + \alpha_{2,4}K_{t-1} + \alpha_{2,5}I_{t-1} + u_{2,t}$$

or in other ways.

Also the third equation can be modified in several ways; for instance, excluding the regressor A_t

$$W_t^p = \alpha_{3,1} + \alpha_{3,2}X_t + \alpha_{3,3}X_{t-1} + u_{3,t}$$

or including lagged value of wages and salaries

$$W_t^p = \alpha_{3,1} + \alpha_{3,2}X_t + \alpha_{3,3}X_{t-1} + \alpha_{3,4}A_t + \alpha_{3,5}(W_{t-1}^p + W_{t-1}^g) + u_{3,t}$$

Etcetera.

t	C	I	Wp	X
1961	95.70054522459304734	-9.67678579582512686	71.40132348840745213	88.423759450.....
1962	95.22062596580400958	-3.74407131601168135	69.92187641252546463	95.376554342.....
1963	105.12850618957142301	5.34134451905109693	79.45634139702196552	113.669851376.....
.....
2012	126.79974091720402202	11.26544320067497741	95.63361169572591585	144.665182971.....
2013	129.89585884936650070	7.79875257518685015	98.53541764346997334	145.094610918.....
2014	137.00585393833001646	10.12679862413780732	104.69845550509313529	160.932651759.....

There is clearly an unpleasant excess of decimal digits in the dataset. In a first phase, if the dataset is used with all decimal digits and the model includes 4 coefficients in each equation, like the original model, 2SLS estimates would be

$$\hat{\alpha} = \begin{bmatrix} 30.00000000 & 0.10000000 & 0.10000000 & 0.80000000 \\ 30.00000000 & 0.40000000 & 0.30000000 & -0.20000000 \\ 30.00000000 & 0.30000000 & 0.10000000 & 0.10000000 \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} 2.00000000 & -1.00000000 & 1.00000000 \\ -1.00000000 & 1.00000000 & -1.00000000 \\ 1.00000000 & -1.00000000 & 2.00000000 \end{bmatrix}$$

(alternatively, data could be produced in order to get round numbers from 3SLS estimation). To produce the dataset, indirect estimation is applied *correctly* as in section 3.1, calibrating the α and Σ parameters (here 12 + 6 parameters).

7.1 Compromise on the number of decimal digits

Too avoid the excess of decimal digits in the dataset, a compromise can be adopted between the number of decimal digits and the number of zeroes after the decimal point in the estimates.

A first tentative solution could be obtained truncating or rounding all dataset numbers at 6 decimal digits: thus, for instance, the integer part of each number would be *billions* of Euros or of US Dollars, the decimal digits would be millions and thousands.

t	C	I	Wp	X
1961	95.700545 22459304734	-9.676785 79582512686	71.401323 48840745213	88.423759 450....
1962	95.220625 96580400958	-3.744071 31601168135	69.921876 41252546463	95.376554 342....
1963	105.128506 18957142301	5.341344 51905109693	79.456341 39702196552	113.669851 376....
....
2012	126.799740 91720402202	11.265443 20067497741	95.633611 69572591585	144.665182 971....
2013	129.895858 84936650070	7.798752 57518685015	98.535417 64346997334	145.094610 918....
2014	137.005853 93833001646	10.126798 62413780732	104.698455 50509313529	160.932651 759....

If the resulting 2SLS estimates were like

$$\hat{\alpha} = \begin{bmatrix} 30.0000003272 & 0.1000002814 & 0.1000003165 & 0.8000003165 \\ 30.0000001488 & 0.4000003143 & 0.3000002879 & -0.2000003000 \\ 30.0000003334 & 0.3000002924 & 0.1000003155 & 0.1000003064 \end{bmatrix}$$

the procedure would be successful.

But unfortunately it may easily happen to obtain

$$\hat{\alpha} = \begin{bmatrix} 29.9999994762 & 0.1000002814 & 0.1000003165 & 0.8000003165 \\ 30.0000001488 & 0.3999995144 & 0.3000002879 & -0.2000003000 \\ 29.9999991256 & 0.3000002924 & 0.1000003155 & 0.1000003064 \end{bmatrix}$$

Believe it or not, asking the student to display results rounded at a couple of decimal digits, it will be frequent the case of a 29.99 displayed as result for the first coefficient, rather than 30.00!

A solution to this problem can be obtained fixing the targets not at 30.000000, 0.1000000, etc. but fixing them at 30.000001, 0.1000001, etc. Having done this, now the indirect estimation procedure

can be applied *correctly*, calibrating the α and Σ parameters (here 12 + 6 parameters), and simulated variables are reported in the dataset rounded at 6 decimal digits. The unpleasant excess of decimal digits has been removed.

If this dataset is produced and given to the student,

t	C	I	Wp	X	P	K	Wg	T	A	G
1961	95.700545	-9.676785	71.401323	88.423759	13.622436	106.795056	2.2	3.4	-9.0	2.4
1962	95.220625	-3.744071	69.921876	95.376554	17.754678	103.050985	2.7	7.7	-8.0	3.9
1963	105.128506	5.341344	79.456341	113.669851	30.313510	108.392329	2.9	3.9	-7.0	3.2
.....
2012	126.799740	11.265443	95.633611	144.665182	40.131571	308.066428	7.8	8.9	5.0	6.6
2013	129.895858	7.798752	98.535417	145.094610	36.959192	315.865180	8.0	9.6	6.0	7.4
2014	137.005853	10.126798	104.698455	160.932651	44.634196	325.991978	8.5	11.6	7.0	13.8

2SLS estimates *rounded* at more than 2 decimal digits (here 5) would be

$$\hat{\alpha} = \begin{bmatrix} 30.00000 & 0.10000 & 0.10000 & 0.80000 \\ 30.00000 & 0.40000 & 0.30000 & -0.20000 \\ 30.00000 & 0.30000 & 0.10000 & 0.10000 \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} 2.00000 & -1.00000 & 1.00000 \\ -1.00000 & 1.00000 & -1.00000 \\ 1.00000 & -1.00000 & 2.00000 \end{bmatrix}$$

(In principle further improvements are possible, to increase the number of desired exact decimal digits in the estimates, by introducing more decimal digits into the exogenous variables. We feel it unnecessary, as the results seem *elegant* enough).

Inconveniences may easily arise after completion of the procedure. Due to the complexity of the model, many checks should be performed to verify that the model and the dataset are *economically coherent* and surely *manageable by the student*. The structural form coefficients have been chosen from a menu where all numbers are meaningful from the economic viewpoint. For instance, in the consumption equation, the marginal propensity to consume of profits and of wages and salaries are positive and less than one, coefficient of profits being smaller than the coefficient of wages and salaries; or, if taxes are inserted into the equation, the coefficient must be negative; etc. However, it may happen that, after structural form coefficients have been computed with the desired round values, impact or delay multipliers have the wrong sign (if, for instance, the Keynesian multiplier $\partial X_t / \partial G_t$ is negative, or it is positive but too large). Or it may happen that the value forecasted out of sample one-step-ahead for C_{T+1} is more than 20% larger or smaller than the last *observed* value in the sample, C_T , which would be unreasonable from the economic viewpoint.

If one of these checks (also other checks are performed) does not give correct results, the model and its dataset are rejected; a new random choice of a model is performed, and the construction of its corresponding dataset starts again.

For some exams, students are requested to estimate (in this case simply by OLS) the first equation (consumption) where all variables are replaced by their logarithms. In this case, the simultaneous solution of the model (and therefore also the computation of forecasts, goodness of fit indicators, multipliers) requires the implementation of some numerical solution method like Gauss-Seidel; and it may

happen that the Gauss-Seidel numerical solution method does not converge in one of the years, when performing the static or the dynamic solution of the model (unless specific *tricks* are experimented with, like equations reordering or introduction of relaxation parameters; these are however not requested to the students). If this happens, or any of the same checks performed on the linear model does not give positive results, also in this case the model and its dataset are rejected.

The number of rejections is moderately small, and computation is fast. If only linear models have to be produced, there are about 30% rejections, and the average computation time to produce an acceptable model is a bit more than 1 second on a PC with a modern i7 processor. If also the nonlinear modification of the model is requested, there are about 70% rejections, and the average computation time to produce an acceptable model is about two and a half seconds.

8 Computation

Algorithms have been programmed in Fortran-77. The basic structure of the main program is essentially still as in the stochastic simulation program that was announced many years ago in *Econometrica* (Bianchi et al. 1978). Then, a set of specific subroutines deals with each of the models discussed above, separately. For each model, the calibration procedure is programmed inside one of the specific subroutines, and it is essentially as it was implemented in Calzolari et al. (1998).

The version of the programs which is available on request from the author is *slightly more* user-friendly. No installation procedure is requested. The *executables* can directly run on any 32 or 64 bits Windows system. The *source* Fortran code is also supplied for each type of model (it can be modified by expert programmers, or it can undergo the usual *compile, link and go* procedure under other operating systems). Also an *ASCII data* file for each type of model contains some specifications that can be modified or adapted by the user, such as initial and final year of the time series, number of the cross-section units, etc. A simple *ASCII text* file must be provided by the user with the list of the student's names (a sample is provided).

For each model, all outputs are *ASCII text* files. There is one file for each student (containing the individual dataset, as well as a brief description of the model to be estimated) and one *Latex* file, for the examiner, containing all the *expected* results.

9 References

Baltagi, B. (1998): *Econometrics*. Berlin: Springer-Verlag.

Bianchi, C. and G. Calzolari (1982): "Evaluating Forecast Uncertainty Due to Errors in Estimated Coefficients: Empirical Comparison of Alternative Methods", in *Evaluating the Reliability of Macroeconomic Models*, ed. by G. C. Chow and P. Corsi. New York: John Wiley, 251-277.

- Bianchi, C., G. Calzolari, and P. Corsi (1978): “A Program for Stochastic Simulation of Econometric Models”, *Econometrica* 46, 235-236.
- Bollerslev, T. (1986): “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics* 31, pages 307-327.
- Box, G. E. P. and G. M. Jenkins (1979): *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Calzolari, G., and L. Sampoli (1993): “A Curious Result on Exact FIML and Instrumental Variables”, *Econometric Theory* 9, 296-309.
- Calzolari, G., F. Di Iorio, and G. Fiorentini (1998): “Control Variates for Variance Reduction in Indirect Inference: Interest Rate Models in Continuous Time”, *Econometrics Journal* 1, C100-C112.
- Chernoff, H. and N. Divinsky (1953): “The Computation of Maximum-Likelihood Estimates of Linear Structural Equations”, in *Studies in Econometric Method*. ed. by W. C. Hood and T. C. Koopmans. New York: John Wiley, 236-302.
- Engle, R. F. (1982): “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, *Econometrica* 50, 987-1007.
- Fiorentini, G., G. Calzolari, and L. Panattoni (1996): “Analytic Derivatives and the Computation of GARCH Estimates”, *Journal of Applied Econometrics* 11, 399-417.
- Gallant, A. R. and G. Tauchen (1996): “Which moments to match?”, *Econometric Theory* 12, 657-681.
- Goldberger, A. S., A. L. Nagar and H. S. Odeh (1961): “The Covariance Matrices of Reduced-Form Coefficients and of Forecasts for a Structural Econometric Model”, *Econometrica* 29, 556-573.
- Gourieroux, C., A. Monfort and E. Renault (1993): “Indirect Inference”, *Journal of Applied Econometrics* 8, S85-S118.
- Granger, C. W. J. (1989): *Forecasting in Business and Economics* (2nd edition). San Diego, CA: Academic Press.
- Greene, W. H. (2008): *Econometric Analysis* (6th edition). Upper Saddle River, NJ: Prentice-Hall.
- Hendry, D. F. (1971): “Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with Errors Generated by a Vector Autoregressive Process”, *International Economic Review* 12, 257-272.
- Johnston, J. (1984): *Econometric Methods* (3rd edition). New York: McGraw-Hill.
- Klein, L. R. (1950): *Economic Fluctuations in the United States, 1921-1941*. New York: John Wiley, Cowles Commission Monograph No. 11.
- Renfro, C. G. (2009): *Building and Using a Small Macroeconometric Model: Klein Model I as an Example*. Philadelphia: Modler Information Technologies Press.
- Smith, A. (1993): “Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions,” *Journal of Applied Econometrics* 8, S63-S84.
- Wooldridge, J. (2002): *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

