CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Cardiff Economics Working Papers

Konstantinos Theodoridis

Dynamic Stochastic General Equilibrium (DSGE) Priors for Bayesian Vector Autoregressive (BVAR) Models: DSGE Model Comparison

E2007/15

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

ISSN 1749-6101
June 2007

Dynamic Stochastic General Equilibrium (DSGE) Priors for Bayesian Vector Autoregressive (BVAR) Models: DSGE Model Comparison

Konstantinos Theodoridis *†

June 6, 2007

Abstract

This *Paper* describes a procedure for constructing *theory* restricted prior distributions for *BVAR* models. The *Bayes Factor*, which is obtained without any additional computational effort, can be used to assess the plausibility of the restrictions imposed on the VAR parameter vector by competing DSGE models. In other words, it is possible to rank the amount of abstraction implied by each DSGE model from the *historical* data.

JEL *Classification:* C11, C13, C32, C52

Keywords: BVAR, DSGE Model Evaluation, Gibbs Sampling, Bayes Factor

1 Introduction

Bayesian inference relies on the properties of the *posterior* distribution, $p(\theta/Y_t, \mathcal{T})$, where $\theta \in \mathbf{\Theta}$[1] is the parameter vector of the econometric model, $\mathcal{T}$, and $Y_t \equiv \{y_1, y_2, ..., y_T\}$ denotes the historical data[2]. This is proportional to the product of the *Likelihood* of θ, $L(\theta/Y_t)$, and the prior distribution of θ, $p(\theta)$ ($p(\theta, Y_t/\mathcal{T}) \propto L(\theta/Y_t)\, p(\theta)$). The well-known weak point of the *Bayesian* methodology is the specification of the prior distribution regarding θ because such selection is essentially arbitrary[3]. In addition although the posterior distribution coincides asymptotically with the likelihood the influence of the prior can be substantial in small samples or when the likelihood is flat. In other words the inference about θ can be dominated by the selection of $p(\theta)$ in a manner which is difficult to interpret.

In the *DSGE* framework prior distribution selection for VAR models can be theoretically founded. Ingram & Whiteman (1994) were the first to construct

*Cardiff Business School, Cardiff University

†Corresponding Author: Cardiff Business School, Colum Drive, Cardiff, CF10 3EU, Tel: +44 (0)29 2087 0558, E-mail:Theodoridisk2@cardiff.ac.uk

[1] $\mathbf{\Theta}$ is a compact subset of the $\mathbb{R}^{d\theta}$ where $\mathbb{R}$ is the real line and da denotes the dimension of vector a.

[2] $y_t \in \mathbb{R}^{dy}$.

[3] See, Canova (2005, Chapter 9) for a detailed discussion and alternative methods of selecting priors.

theoretical priors for a BVAR using the basic neoclassical model (King et al., 1988). In order to see this notice that the *state-space* format of a solved DSGE model, linearized around a nonstochastic steady state, $\mathcal{M}$, is

$$
\begin{aligned}
y_t &= A(\gamma)\,\xi_t & (1)\\
\xi_t &= B(\gamma)\,\xi_{t-1} + \Gamma(\gamma)\,\varepsilon_t & (2)
\end{aligned}
$$

Equation (2) describes the evolution of the state vector, ${\xi_t}$[4], of the model, while (1) gives the relation between the state vector of the model and the observable variables. Finally, ε_t is normally distributed with mean zero and I_{dy}[5] covariance matrix, $\mathbf{N}(0, I_{dy})$. Obviously, the elements of the matrices $A(\gamma)$, $B(\gamma)$, and $\Gamma(\gamma)$ are nonlinear functions of the structural parameter vector γ. The key element in Ingram & Whiteman's analysis is that the dimension of the observable vector exceeds the dimension of the state vector, implying that the matrix $A(\gamma)$ has a full column rank. Through the use of the *Generalized Inverse*[6] the above system can be rewritten as VAR(1), $y_t = A(\gamma)\,B(\gamma)\left[\left(A(\gamma)'\,A(\gamma)\right)^{-1} A(\gamma)'\right] y_{t-1} + A(\gamma)\,\Gamma(\gamma)\,\varepsilon_t$. Clearly, θ is a function of γ, say $\theta = \phi(\gamma)$, and, given that γ is normally distributed with μ_γ mean and Σ_γ covariance matrix, then the prior distribution of θ is again normal with a mean and variance given by $\mu_\theta = \phi(\mu_\gamma)$ and $\Sigma_\theta = \nabla_\gamma \phi(\mu_\gamma)\,\Sigma_\gamma \nabla_\gamma \phi(\mu_\gamma)'$[7], respectively, $\mathbf{N}(\mu_\theta, \Sigma_\theta)$.

This *structural* VAR (SVAR) representation, which holds under the condition that the number of the observable variables exceed those in the state vector, can provide information only for the first lag of a VAR(p). At this point Ingram & Whiteman (1994) make the assumption that the rest $(p-1)$ blocks of autoregressive parameters are normally distributed with mean zero and a covariance matrix that is proportional to the covariance matrix of the first lag with the proportion parameter being an inverse function of the number of lags.

Recently, Del Negro & Schorfheide (2004) have suggested an alternative methodology of producing *theoretically* founded priors. They assume implicitly that the condition of the *Proposition* 2.1 of Christiano et al. (2006) hold[8], implying that the structural model described by equations (1) and (2) can be written as a SVAR of infinite order

$$y_t = \Delta_1(\gamma)\,y_{t-1} + \Delta_2(\gamma)\,y_{t-2} + ... + \Psi(\gamma)\,\varepsilon_t \tag{3}$$

where

$$\Delta_j(\gamma) = A(\gamma)\,B(\gamma)\,M(\gamma)^{j-1}\,\Gamma(\gamma)\,(A(\gamma)\Gamma(\gamma))^{-1} \tag{4}$$

for $j = 1, 2, ...,$ $\Psi(\gamma) = A(\gamma)\Gamma(\gamma)$, $M(\gamma) = \left[I_{dx} - \Gamma(\gamma)\Psi(\gamma)^{-1} A(\gamma)\right] B(\gamma)$. Since *historical* data is finite, however, the above time series process must be approximated by finite one, e.g. a SVAR(p) such as $y_t = \Delta_1(\gamma)\,y_{t-1} + ... +$

[4] $\xi_t \in \mathbb{R}^{d\xi}$

[5] I_{da} denotes the identity $(da \times da)$ matrix.

[6] See, for example Magnus & Neudecker (2002) for a detailed discussion.

[7] $\nabla_\alpha R(\alpha)$ denotes the diffrentiation of the vector function $R(\cdot)$ with respect to the vector α.

[8] These conditions are: (i) the number of the structural shocks in $\mathcal{M}$ equals to the number of the observable variable, and, (ii) the maximum *eigenvalue* of the matrix $M(\gamma) = \left[I_{dx} - \Gamma(\gamma)\,(A(\gamma)\Gamma(\gamma))^{-1} A(\gamma)\right] B(\gamma)$ is less than one in absolute terms.

$\Delta_p(\gamma) y_{t-p} + \Psi(\gamma)\varepsilon_t$. Therefore, the structural model $\mathcal{M}$ implies certain restrictions on the VAR(p) process, $y_t = \Phi_1 y_{t-1} + ... + \Phi_p(\gamma) y_{t-p} + v_t$, which are taken seriously by Del Negro & Schorfheide (2004), however, they are not, *"dogmatically"* imposed as in the Ingram & Whiteman's analysis. In this case the prior distribution of the VAR parameters, ($\beta = \left(\theta' = vec(\Phi_1, ..., \Phi_p)', vech(\Sigma_v)'\right)'$[9], where Σ_v is the covariance matrix of v_t), which is assumed to have a *typical* Normal-Wishart format, is obtained by fitting the VAR(p) on data simulated from the structural model, whose length is equal to a fraction, λ, of the length of the actual data. The beauty of this methodology is that posterior distribution of β has a closed form format and λ, which is a data driven *hyperparameter*, is as a measure of fit that could be used for the evaluation of the structural model; it shows how much the data likes the restrictions implied by $\mathcal{M}$[10].

2 The Proposed Method

This paper suggests a way of constructing priors for a VAR(p), which combines the above two procedures. Similar to Del Negro & Schorfheide (2004), the conditions that allow the structural model to be expressed as an infinite SVAR(p) (Equation 3), hold here as well. However, the restrictions implied by the structural model $\mathcal{M}$ are imposed in a manner similar to Ingram & Whiteman (1994). Therefore, under the assumption that γ is normally distributed with mean and variance equal to μ_γ and Σ_γ, respectively, θ is also normally distributed with a mean and variance given by $\mu_\theta = \phi(\mu_\gamma)$ and $\Sigma_\theta = \nabla_\gamma \phi(\mu_\gamma)\Sigma_\gamma \nabla_\gamma \phi(\mu_\gamma)'$, respectively. However, the additional assumptions posed here are that Σ_θ is block diagonal (its block is given by $\Sigma_\theta^j = \nabla_\gamma vec\left[\Delta_j(\mu_\gamma)\right]\Sigma_\gamma \nabla_\gamma vec\left[\Delta_j(\mu_\gamma)\right]'$, i.e. $\Sigma_\theta = diag\left(\Sigma_\theta^1, ..., \Sigma_\theta^p\right)$), and, the dimension of γ is greater equal to the square of the number of the observable variables, dy^2. These two assumptions ensure that the distribution of θ is non-degenerate.

From Proposition 2.1 of Christiano et al. (2006) it is known that $\Sigma_v = \Psi(\gamma)\Psi(\gamma)'$. Given that $\Psi(\gamma)$ is normally distributed, it can also be concluded that Σ_v^{-1} follows the *Wishart* distribution with a scale matrix given by $\Pi = \Psi(\mu_\gamma)\Psi(\mu_\gamma)'$ and degrees of freedom equal to η, $\mathbf{W}(\Pi, \eta)$. The priors in this case regarding θ and Σ_v^{-1} conditional on the structural model $\mathcal{M}$ are, therefore, given by

$$p(\theta/\mathcal{M}) \propto |\Sigma_\theta|^{-0.5d\theta} \exp\left\{-0.5\left[\Sigma_\theta^{-0.5}(\theta - \mu_\theta)\right]'\left[\Sigma_\theta^{-0.5}(\theta - \mu_\theta)\right]\right\} \quad (5)$$

and

$$p\left(\Sigma_v^{-1}/\mathcal{M}\right) \propto |\Pi|^{0.5\eta} |\Sigma_v|^{-0.5(\eta+dy-1)} \exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\} \quad (6)$$

Proposition 2.1 *Given expressions (5) and (6) the posterior distribution of θ and Σ_v^{-1}, $p\left(\theta, \Sigma_v^{-1}/Y_t, \mathcal{T}\right)$, can be written as the product of the conditional densities, $p\left(\theta/\Sigma_v^{-1}, Y_t, \mathcal{T}, \mathcal{M}\right) = \mathbf{N}\left(\tilde{\theta}, \tilde{\Sigma}_\theta\right)$ and $p\left(\Sigma_v^{-1}/\theta, Y_t, \mathcal{T}, \mathcal{M}\right) = \mathbf{W}\left(\tilde{\Pi}, T + \eta\right)$,*

[9] The *vec* and *vech* operators transform a matrix with dimensions $m \times m$, respectively, into an $m^2 \times 1$ vector by stacking the columns and to $m(m+1)/2 \times 1$ vector by stacking the elements of and below the main diagonal.

[10] Actually, the most attractive element of this methodology, which is not related to the analysis undertaken here, is that it leads to a posterior estimate of γ for which the distance between the SVAR(p) and VAR(p) is small enough.

where $\tilde{\theta} = \left[\Sigma_\theta^{-1} + \left(\Sigma_v^{-1} \otimes X'X\right)\right]^{-1} \left[\Sigma_\theta^{-1}\mu_\theta + \left(\Sigma_v^{-1} \otimes X\right)' \mathbf{y}\right]$, $\tilde{\Sigma}_\theta = \left[\Sigma_\theta^{-1} + \left(\Sigma_v^{-1} \otimes X'X\right)\right]^{-1}$, *and* $\tilde{\Pi} = \left[\Pi + (Y_t - X\theta)' (Y_t - X\theta)\right]$. $\square$

The benefit of the above Proposition is twofold. Firstly, the *Gibbs* sampler[11] allows us to approximate the posterior distribution of θ and Σ_v^{-1}, $p\left(\theta, \Sigma_v^{-1}/Y_t, \mathcal{T}\right)$. Given the current state of PCs this technique is only marginally computationally more demanding than the one introduced by Del Negro & Schorfheide. However, in contrast to their method, the one suggested here offers a measure of fit regarding the DSGE model $\mathcal{M}$, without any additional computational effort. This convenience arises from the work of Chib (1995) who shows that a *marginal likelihood* of the time series model $\mathcal{T}$, $\mathcal{L}(Y_t/\mathcal{T}, \mathcal{M}) = \int L\left(Y_t/\beta, \mathcal{T}\right) p\left(\beta/\mathcal{M}\right) d\beta$, can be obtained from the output of the above *Markov Chain Monte Carlo* (MCMC) resampling scheme.

Given the *marginal likelihood*, the *Bayes Factor*, $\mathcal{B}_{ij}=\frac{\mathcal{L}(Y_t/\mathcal{T},\mathcal{M}_i)}{\mathcal{L}(Y_t/\mathcal{T},\mathcal{M}_j)}$, of comparing κ time series models ($\mathcal{T}_j$, where $j \neq i = 1, ..., \kappa$), is readily calculated. In the present framework the time series model remains the same and the likelihood does not change. However, the priors about θ and Σ_v, $p(\theta/\mathcal{M}_j)$ and $p(\Sigma_v^{-1}/\mathcal{M}_j)$ where $j = 1, ..., \kappa$, are derived from κ different competing DSGE models and therefore this methodology provides us with a way of assesing the plausibility of the various restrictions implied by different *economic* theories to the *historical* data. This process provides substantial computational advantages, as mentioned, and it has a very good intuition, which is explained in the next paragraph.

DSGE models have been developed to explain certain economic phenomena and not all the features of the *historical* data. This implies that they are *parsimonious,* and, by construction, misspecified. The natural way to assess their performance is, therefore, based on these phenomena, which are summarized by the restrictions imposed on the time series parameter vector β. This is the logic behind the proposed evaluation process and it coincides with the one suggested by Del Negro & Schorfheide, whose measure of fit evaluates the extent to which these restrictions could be relaxed to make the time series model (again with priors coming from the structural model, however, as it has been discussed above, they are derived differently) to fit well in the *historical* data.

3 Conclusion

This *Paper* describes an alternative methodology of obtaining prior distributions for BVAR models, which can be applied when the number of the structural shock coincides with the number of the observable variables, and, the *eigenvalues* of the matrix $M(\gamma)$ are less than one in absolute terms. It is shown that the posterior distribution of the time series parameter vector can be written as the product of a conditional Normal and a conditional Wishart density, implying that the former distribution can be approximated through the use of the Gibbs sampler. Existing results in the MCMC literature show that the *marginal likelihood* and consequently the *Bayes Factor*, can be obtained from the output of the above sampling scheme. In other words, there is no additional computa-

[11] For a detailed discussion see Gilks et al. (1996); Robert & Casella (2004); Canova (2005).

tional cost of getting these values, which can be used to rank competing DSGE models.

A Appendix

Proof: *of Proposition 2.1*

$$\begin{aligned}
p\left(\theta,\Sigma_v^{-1}/Y_t,\mathcal{T}\right) &= p\left(\theta/\mathcal{M}\right)p\left(\Sigma_v^{-1}/\mathcal{M}\right)L\left(Y_t/\theta,\Sigma_v^{-1},\mathcal{T}\right)\\
&\propto \left|\Sigma_\theta\right|^{-0.5d\theta}\exp\left\{-0.5\left[\Sigma_\theta^{-0.5}\left(\theta-\mu_\theta\right)\right]'\left[\Sigma_\theta^{-0.5}\left(\theta-\mu_\theta\right)\right]\right\}\\
&\quad\times\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\}\\
&\quad\times\left|\Sigma_v\right|^{-0.5T}\exp\left\{\begin{array}{c}-0.5\left[\left(\Sigma_v^{-0.5}\otimes I_T\right)\mathbf{y}-\left(\Sigma_v^{-0.5}\otimes X\right)\theta\right]'\\ \left[\left(\Sigma_v^{-0.5}\otimes I_T\right)\mathbf{y}-\left(\Sigma_v^{-0.5}\otimes X\right)\theta\right]\end{array}\right\}\\
&= \left|\Sigma_\theta\right|^{-0.5d\theta}\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(T+\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\}\\
&\quad\exp\left\{-0.5\left(\begin{array}{c}\left[\left(\Sigma_v^{-0.5}\otimes I_T\right)\mathbf{y}-\left(\Sigma_v^{-0.5}\otimes X\right)\theta\right]'\\ \left[\left(\Sigma_v^{-0.5}\otimes I_T\right)\mathbf{y}-\left(\Sigma_v^{-0.5}\otimes X\right)\theta\right]\\ +\left[\Sigma_\theta^{-0.5}\left(\theta-\mu_\theta\right)\right]'\left[\Sigma_\theta^{-0.5}\left(\theta-\mu_\theta\right)\right]\end{array}\right)\right\}
\end{aligned}\tag{7}$$

Similar to Canova (2005, Chapter 10, page: 354) let $\mathcal{Y}\equiv\left[\begin{array}{cc}\Sigma_\theta^{-0.5}\mu_\theta & \left(\Sigma_v^{-0.5}\otimes I_T\right)\mathbf{y}\end{array}\right]'$ and $\mathcal{X}\equiv\left[\begin{array}{cc}\Sigma_\theta^{-0.5} & \left(\Sigma_v^{-0.5}\otimes X\right)\end{array}\right]'$, then (7) becomes

$$\begin{aligned}
p\left(\theta,\Sigma_v^{-1}/Y_t,\mathcal{T}\right) &\propto \left|\Sigma_\theta\right|^{-0.5d\theta}\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\}\\
&\quad\times\left|\Sigma_v\right|^{-0.5T}\exp\left\{-0.5\left(\mathcal{Y}-\mathcal{X}\theta\right)'\left(\mathcal{Y}-\mathcal{X}\theta\right)\right\}\\
&= \left|\Sigma_\theta\right|^{-0.5d\theta}\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\}\\
&\quad\times\left|\Sigma_v\right|^{-0.5T}\exp\left\{\begin{array}{c}-0.5\left[\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)-\mathcal{X}\left(\theta-\tilde{\theta}\right)\right]'\\ \left[\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)-\mathcal{X}\left(\theta-\tilde{\theta}\right)\right]\end{array}\right\}\\
&= \left|\Sigma_\theta\right|^{-0.5d\theta}\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\} \qquad (8)\\
&\quad\times\left|\Sigma_v\right|^{-0.5T}\exp\left\{-0.5\left[\begin{array}{c}\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)'\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)\\ -\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)'\mathcal{X}\left(\theta-\tilde{\theta}\right)\\ -\left(\theta-\tilde{\theta}\right)'\mathcal{X}'\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)\\ +\left(\theta-\tilde{\theta}\right)'\mathcal{X}'\mathcal{X}\left(\theta-\tilde{\theta}\right)\end{array}\right]\right\}
\end{aligned}$$

by setting $\tilde{\theta}=\left(\mathcal{X}'\mathcal{X}\right)^{-1}\mathcal{X}'\mathcal{Y}=\left[\Sigma_\theta^{-1}+\left(\Sigma_v^{-1}\otimes X'X\right)\right]^{-1}\left[\Sigma_\theta^{-1}\mu_\theta+\left(\Sigma_v^{-1}\otimes X\right)'\mathbf{y}\right]$, (8) reduces to

$$\begin{aligned}
p\left(\theta,\Sigma_v^{-1}/Y_t,\mathcal{T}\right) &\propto \left|\Sigma_\theta\right|^{-0.5d\theta}\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(\eta+dy-1)}\exp\left\{-0.5tr\left(\Sigma_v^{-1}\Pi\right)\right\} \qquad (9)\\
&\quad\left|\Sigma_v\right|^{-0.5T}\exp\left\{-0.5\left[\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)'\left(\mathcal{Y}-\mathcal{X}\tilde{\theta}\right)+\left(\theta-\tilde{\theta}\right)'\mathcal{X}'\mathcal{X}\left(\theta-\tilde{\theta}\right)\right]\right\}
\end{aligned}$$

From (9), conditioning on Σ_v, $p\left(\theta/\Sigma_v^{-1},Y_t,\mathcal{T}\right)\propto\left|\Sigma_\theta\right|^{-0.5d\theta}\exp\left\{-0.5\left(\theta-\tilde{\theta}\right)'\mathcal{X}'\mathcal{X}\left(\theta-\tilde{\theta}\right)\right\}$ is a normal distribution with mean and variance equal to $\tilde{\theta}$ and $\left[\Sigma_\theta^{-1}+\left(\Sigma_v^{-1}\otimes X'X\right)\right]^{-1}$, respectively. While from equation (7) it is clear that, conditioning on θ, $p\left(\Sigma_v^{-1}/\theta,Y_t,\mathcal{T}\right)\propto\left|\Pi\right|^{0.5\eta}\left|\Sigma_v\right|^{-0.5(T+\eta+dy-1)}$ $\exp\left\{-0.5tr\left(\Sigma_v^{-1}\left[\Pi+\left(Y_t-X\theta\right)'\left(Y_t-X\theta\right)\right]\right)\right\}$, which is the Wishart distribution with a scale matrix given by $\left[\Pi+\left(Y_t-X\theta\right)'\left(Y_t-X\theta\right)\right]$ and with $T+\eta$ degrees of freedom. ■

References

Canova, F. (2005), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton.

Chib, S. (1995), 'Marginal likelihood form the gibbs output', *Journal of American Statistical Association* **90**, 1313–1321.

Christiano, L., Eichenbaum, M. & Vigfusson, R. (2006), 'Assessing structural vars', *Working Paper,* **NBER**, 12353.

Del Negro, M. & Schorfheide, F. (2004), 'Priors from general equilibrium models for vars', *International Economic Review* **45**, 643–673.

Gilks, W., Richardson, S. & Spiegelhalter, D. (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, New York.

Ingram, B. & Whiteman, C. (1994), 'Supplanting the 'minnesota' prior forecasting macroeconomic time series using real business cycle model priors', *Journal of Monetary Economics* **34**, 497–510.

King, R., Plosser, C. & Rebelo, S. (1988), 'Production, growth, and business cycles: I. the basic neoclassical model', *Journal of Monetary Economics* **21**, 195–232.

Magnus, J. & Neudecker, H. (2002), *Matrix Differential Calculus with Application in Statistics and Econometrics*, WILEY, New York.

Robert, C. & Casella, C. (2004), *Monte Carlo Statistical Methods*, Springer-Verlag, New York.