General–to–Specific Reductions of Vector Autoregressive Processes

Hans-Martin Krolzig *
Department of Economics and Nuffield College, Oxford University.
hans-martin.krolzig@nuf.ox.ac.uk

November 12, 2000

Abstract

Unrestricted reduced form vector autoregressive (VAR) models have become a dominant research strategy in empirical macroeconomics since Sims (1980) critique of traditional macroeconometric modeling. They are however subjected to the curse of dimensionality. In this paper we propose general-to-specific reductions of VAR models and consider computer-automated model selection algorithms embodied in *PcGets* (see Krolzig and Hendry, 2000) for doing so. Starting from the unrestricted VAR, standard testing procedures eliminate statistically-insignificant variables, with diagnostic tests checking the validity of reductions, ensuring a congruent final selection. Since jointly selecting and diagnostic testing eludes theoretical analysis, we evaluate the proposed strategy by simulation. The Monte Carlo experiments show that *PcGets* recovers the DGP specification from a large unrestricted VAR model with size and power close to commencing from the DGP itself. The application of the proposed reduction strategy to a US monetary system demonstrates the feasibility of *PcGets* for the analysis of large macroeconomic data sets.
JEL Classification: C51, C32, E52.
Keywords: Model selection; Vector autoregression; Subset model; Lag order determination; Data mining.

1 Introduction

Since Sims (1980) critique of traditional macroeconometric modeling, vector-autoregressive (VAR) models are widely used in econometrics. Their popularity is due to the flexibility of the VAR framework and the ease of producing macroeconomic models with useful descriptive characteristics, within statistical tests of economically meaningful hypothesis can be executed. Over the last two decades VARs have been applied to numerous macroeconomic data sets providing an adequate fit of the data and fruitful insight on the interrelations between economic data.

Many estimation problems in the (unrestricted) VAR have been solved by Sims (1980) (see also the overview in Lütkepohl, 1991). The serious problem the VAR approach is faced with is the so-called curse of dimensionality: In a vector autoregression of dimension K, each additional lag adds K^2 coefficients. In words of Sims (1980, p.16)

> "If every variable is allowed to influence every other with a distributed lag of reasonable length, without restriction, the number of parameters grows with the square of the number of variables and quickly exhausts the degree of freedom".

*Financial support from the UK Economic and Social Research Council under the grant L138251009 is gratefully acknowledged. Address of the author: Hans-Martin Krolzig, University of Oxford, Department of Economics, Manor Road, Oxford OX1 3UL, UK.

With sample sizes commonly used in applied modeling, the available degrees of freedom are often small. This could affect the power of cointegration tests, for example. Similarly, monthly data might require a longer lag length than available data allow, and large VARs might have not been considered because of the curse of dimensionality.

This paper considers strategies to select alternative specifications which are more parsimonious. Already Sims (1980, p.33) pointed out that,

> "In expanding the list of variables in the model, practical methods for limiting the growth in number of parameters as sample size increases will have to be developed".

In this context, Sims conjectured the use of index models.

The literature on model selection has mainly focused on the selection of lag order, p, of an otherwise unrestricted VAR. In these selection procedures a model is usually selected by an information criterion which penalizes the likelihood function for the number of parameters.[1] Lütkepohl (1991, ch.5) discusses various strategies for the specification of subset VAR models. Subset VAR models are VARs with zero constraints on the coefficients. In most subset VAR modeling strategies the model choice is again based on the optimization of a specified model selection criterion. For a given maximal order p of VAR, a full search over all possible candidates is computationally unfeasible: in a VAR(p) without deterministic terms there are K^2p coefficients, any full search requires the estimation of a total of 2^{K^2p} subset models. Therefore various strategies have been proposed to overcome this problem (search over complete VAR matrices, top-down and bottom-up specification of the distributed lag lengths etc.). Brüggemann and Lütkepohl (2000) consider step-wise regression type single-equation reduction paths where the critical value is chosen such that an acceptance of the null hypothesis guarantees a marginal increase in a given information criterion.

In this study we consider General-to-specific (*Gets*) reductions of the unrestricted and, hence, highly parameterized VAR. The *Gets* reduction process is designed to ensure that the parsimonious subset VAR will convey all the information embodied in the unrestricted VAR. This is achieved by a joint selection and diagnostic testing process: starting from the unrestricted, congruent general model, standard testing procedures are used to eliminate statistically-insignificant variables, with diagnostic tests checking the validity of reductions, ensuring a congruent final selection. By reducing the complexity of the unrestricted VAR and checking the contained information, the selected simpler, more compact model provides an improved statistical description of the economic world (see Hendry, 1993, for an overview of the so-called 'LSE' methodology).

While the joint issue of sequential variable selection and diagnostic testing using multiple criteria has eluded most attempts at theoretical analysis, an evaluation of the properties of the model-selection process can be achieved by simulation. To implement a model-selection procdure approach in a computer algorithm, all decisions have to be mechanized. For the **Ge***neral*-**t***o*-**s***pecific* approach, Krolzig and Hendry (2000) developed *PcGets*. In this paper we use *PcGets* to analyze the *Gets* selection of subset VARs from a computer-automation perspective.

[1] The information criteria considered in the literature are defined as follows:

$$\begin{aligned} AIC &= -2\log L/T + 2n/T, \\ SC &= -2\log L/T + n\log(T)/T, \\ HQ &= -2\log L/T + 2n\log(\log(T))/T, \end{aligned}$$

where L is the maximized likelihood, n is the number of parameters and T is the sample size: see Akaike (1985), Schwartz (1978), and Hannan and Quinn (1979).

PcGets is a computer-automated approach to econometric modeling focusing on general-to-specific reduction approaches for linear, dynamic, regression models. The development of *PcGets* was stimulated by Hoover and Perez (1999), who had sought to evaluate the performance of *Gets*. *PcGets* mimics a researcher following the *general-to-specific* approach to econometric modeling. Hendry and Krolzig (1999) and Krolzig and Hendry (2000) have shown in Monte Carlo experiments that *PcGets* recovers the DGP specification from a general model with size and power close to commencing from the DGP itself. *PcGets* has been designed for modeling univariate time-series data when the precise formulation of the economic system under analysis is not known *a priori*. In this paper we investigate the application of the single-equation model-selection procedures automated in *PcGets* to VAR models.

Section 2 discusses the econometrics of model selection. In generalization of the *PcGets* algorithm for single-equation models, *Gets* reduction strategies for the specification of subset VARs are proposed. In section 3 we investigate by simulation whether the *PcGets* model-selection process works well or fails badly in the VAR framework. Results are presented for a Monte Carlo experiment where the data generating process (DGP) is a highly restricted bivariate VAR(2) and the general unrestricted model (GUM) is a VAR (4). The empirical illustration with a US monetary system presented in section 4 evaluates the usefulness of *PcGets* for the analysis of large macroeconomic data sets.

2 General-to-specific VAR model reductions

2.1 The vector autoregressive model

The basic model considered in the following is a vector autoregressive (VAR) model possibly including deterministic terms and with independent Gaussian errors: the K-dimensional time series vector y_t is generated by a vector autoregressive process of order p, denoted VAR(p) model,

$$y_t = A_1 y_{t-1} + \cdots A_p y_{t-p} + B d_t + \varepsilon_t \tag{1}$$

where $t = 1, \ldots, T$, the A_i and B are coefficient matrices and the initial values of $Y_0 = (y_0, \ldots, y_{1-p})$ are fixed. The innovation process ε_t is an unobservable zero-mean white noise process with a time-invariant positive-definite variance-covariance matrix Σ,

$$\varepsilon_t = y_t - \mathsf{E}[y_t | Y_{t-1}].$$

which is assumed to be Gaussian:

$$\varepsilon_t \sim \mathsf{NID}(\mathbf{0}, \Sigma).$$

Thus the expectation of y_t conditional on the information set $Y_{t-1} = (y_{t-1}, y_{t-2}, \ldots, y_{1-p})$ is given by

$$\mathsf{E}[y_t | Y_{t-1}] = B d_t + \sum_{j=1}^{p} A_j y_{t-j}.$$

In the following we will focus on processes where the only deterministic term is an intercept, $B d_t = \nu$.

2.2 Model selection: problems and chances

The key issue for any model-selection procedure is the cost of search, since there are always bound to be mistakes in statistical inference: specifically, how bad does it get to search across many alternatives?

On the one hand, the conventional statistical analysis of repeated testing provides a pessimistic background: every test has a non-zero null rejection frequency (or size, if independent of nuisance

parameters), and so type I errors accumulate. Setting a small size for every test can induce low power to detect the influences that really matter. The study by Lovell (1983) of trying to select a small relation from a large database suggested search had very high costs, leading to an adverse view of 'data mining'. Although Lovell did not consider a structured reduction approach among his methods, *Gets* has been criticized by Pagan (1987) on the grounds that the selection path may matter, and so the result is not 'path independent'. Indeed, Leamer (1983) claimed that 'the mapping is the message'. Moreover, 'pre-testing' is known to bias estimated coefficients, and may distort inference: see *inter alia*, Bock, Yancey and Judge (1973) and Judge and Bock (1978).

On the other hand, White (1990) showed that with sufficiently-rigorous testing, the selected model will converge to the data generating process (DGP). Thus, any 'overfitting' and mis-specification problems are primarily finite sample. Moreover, Mayo (1981) emphasized the importance of diagnostic test information being effectively independent of the sufficient statistics from which parameter estimates are derived. Also, Hendry (1995) argued that congruent models are the appropriate class within which to search, that encompassing resolves many instances of 'data mining', and that in econometrics, theory dependence has as many drawbacks as sample dependence, so modeling procedures are essential. Finally, Hoover and Perez (1999) reconsidered the Lovell (1983) experiments to evaluate the performance of *Gets*. Most important is their notion of commencing from the congruent general model by following a number of reduction search paths, terminated by either no further feasible reductions or significant diagnostic tests occurring. Hoover and Perez select among the surviving models the one which fits best. They show how much better a structured approach is than any method Lovell considered, suggesting that modeling *per se* need not be bad. Indeed, overall, the size of their selection procedure is close to that expected, and the power is reasonable. Moreover, re-running their experiments using *PcGets*, Hendry and Krolzig, 1999 found substantively better outcomes. Thus, the case against model selection is far from proved.

2.3 The *PcGets* model selection algorithm for single-equation models

There is little research on how to design model-search algorithms in econometrics. To reduce search costs, any model-selection process must avoid getting stuck in a search path that initially inadvertently deletes variables that really matter in the DGP, thereby retaining other variables as proxies. Thus, it is imperative to explore multiple paths. To meet this requirement, *PcGets* builds on the multi-path approach to *Gets* model selection in Hoover and Perez (1999). The number of paths is increased to try all single-variable deletions, as well as various block deletions from the general unrestricted model (GUM). Different critical values can be set for multiple and single selection tests, and for diagnostic tests.

Equally, the search procedure must have a high probability of retaining variables that do matter in the DGP. To achieve that, *PcGets* uses encompassing tests between alternative reductions. Balancing these objectives of small size and high power still involves a trade-off, but one that is dependent on the algorithm. The diagnostic tests require careful choice to ensure they characterize the salient attributes of congruency, are correctly sized, and do not overly restrict reductions.

Details of the algorithm are shown in Table 1. In the following we will discuss the econometrics of the different stages of the *PcGets* model-selection algorithm.

2.3.1 The GUM and pre-search tests (*Stage 0*)

The starting point for *Gets* model-selection is the general unrestricted model, so the key issues concern its specification and congruence. In the case of the VAR, the researcher has to specify the order and the

Table 1 The *PcGets* algorithm .

Stage 0. **Estimation and testing of the GUM**

Pre-search tests

Stage I. **Multiple model reduction paths:**

Sequential estimation and test of reductions

(1) Remove insignificant variables.
(2) Model reductions are subjected to a wide range of tests to evaluate their validity:
  – Chow tests for structural stability;
  – residual autocorrelation;
  – ARCH effects in the residuals;
  – normality;
  – heteroscedasticity.

Encompassing

Stage II. **Union testing**

(1) Estimation and test of the new GUM;
(2) Multiple model reduction paths;
(3) Encompassing and final model selection.

Stage III. **Sub-sample evaluation**

(1) Test the significance of every variable in the final model from *Stage II* in two overlapping sub-samples;
(2) Penalize variable accordingly.

dimension of the process. The larger the initial regressor set, the more likely adventitious effects will be retained; but the smaller the GUM, the more likely key variables will be omitted. Further, the less orthogonality between variables, the more 'confusion' the algorithm faces, leading to a proliferation of mutual-encompassing models, where final choices may only differ marginally (e.g., lag 2 versus 1). Finally, the initial specification of the unrestricted VAR must be congruent, with no mis-specification tests failed at the outset. Empirically, the GUM would be revised if such tests rejected, and little is known about the consequences of doing so.

PcGets undertakes 'pre-search' simplification F-tests to exclude variables from the general unrestricted model (GUM), after which the GUM is reformulated. Since variables found to be irrelevant on such tests are excluded from later analyses, this step uses a loose significance level (such as 50%). The step consists of block (F) tests of groups of variables, ordered by their t-values in the GUM . This set includes the overall F-test of all regressors to check that there is something to model. Variables that are insignificant at this step, usually at a liberal critical value, are eliminated from the analysis, and a smaller GUM is formulated.

2.3.2 Multi-stage multi-path search (*Stages I and II*)

The *PcGets* reduction path relies on a classical, sequential simplification and testing approach designed to reduce the complexity of the model by ensuring the congruency of the reduction. Many possible paths from that GUM are investigated: reduction paths considered include both multiple deletions as well as

single, so t and/or F test statistics are used as simplification criteria.

Along each path the least significant variable having a t-values less than the critical value is eliminated.[2] If any diagnostic tests fail, that path is terminated, and the algorithm returns to the last accepted model of the search path: if the last accepted model cannot be further reduced, it becomes the terminal model of the particular search path; otherwise, the last removed variable is re-introduced, and the search path continues with a new reduction by removing the next least-insignificant variable of the last accepted model. If all tests are passed, but one or more variables are insignificant, the least significant variable of those is removed. If that specification has already been tested on a previous path, the current search path is terminated. Finally, if all diagnostic tests are passed, and all variables are significant, the model is the terminal model of that search path.

Some of these searches may lead to different terminal specifications, between which a choice must be made. Consequently, the reduction process is inherently iterative. Should multiple congruent contenders eventuate after a reduction round, encompassing can be used to test between them, with only the surviving – usually non-nested – specifications retained. At the encompassing step, all distinct contending valid reductions are collected, and encompassing is used to test between these (usually non-nested) specifications. Models which survive encompassing are retained; all encompassed equations are rejected.

If multiple models survive this *'testimation'* process, their union forms a new general model, and selection path searches recommence. Such a process repeats till a unique contender emerges, or the previous union is reproduced, then stops. Should that union repeat and mutually-encompassing distinct models survive the encompassing step, a final selection is made using information criteria, otherwise a unique congruent and encompassing reduction has been located.

2.3.3 Sub-sample evaluation (*Stage III*)

As a check for potential over-selection in *Stage II*, *PcGets* exploits sub-sample information by investigating split samples for significance (as against constancy). This mimics the idea of recursive estimation: Since non-central 't'-values diverge with increasing sample size, whereas central 't's fluctuate around zero, the latter have a low probability of exceeding any given critical value in two sub-samples, even when those sample overlap. Thus, adventitiously-significant variables may be revealed by their insignificance in one or both of the sub-samples.

Consequently, a progressive research strategy can gradually eliminate 'adventitiously-significant' variables and tilt the size-power balance favorably. Hoover and Perez (1999) found that by adopting a progressive search procedure (as in *Stage III*), the number of spurious regressors can lowered (inducing a lower overall size), without losing much power. The sub-sample information is used to accord a 'reliability' score to variables, which investigators may use to guide their model choice.

2.3.4 Calibration

The *'testimation'* process of *PcGets* depends on a number of decisions regarding the specification of the algorithm. The choice of mis-specification alternatives determines the number and form of the diagnostic tests. Their individual significance levels in turn determine the overall significance level of the test battery. Since significant diagnostic-test values terminate search paths, they act as constraints on moving away from the GUM. Thus, if a search is to progress towards an appropriate simplification, such

[2] *PcGets* allows the pre-selection of regressors: for example, one might want to fix economically-interesting spill-overs, then apply *Gets* to the remaining regressors.

tests must be well 'focused' and have the correct size. The choice of critical values for pre-selection, selection and encompassing tests is also important for the success of *PcGets*: the tighter the size, the fewer the 'spurious inclusions of irrelevant', but the more the 'false exclusions of relevant' variables. In the final analysis, the calibration of *PcGets* depends on the characteristics valued by the user: if *PcGets* is employed as a first 'pre-selection' step in a user's research agenda, the optimal significance level may be higher than when the focus is on controlling the overall size of the selection process. For single-equation models, Krolzig and Hendry (2000) investigated the calibration of *PcGets* for the operational characteristics of the diagnostic tests, the selection probabilities of DGP variables, and the deletion probabilities of non-DGP variables. Research has been undertaken to investigate the impact of these choices on model selection in order to provide the user with 'optimized' search strategies. The calibration of *PcGets* used in the Monte Carlo experiments and the empirical modeling example is reported in the appendix.

So far section we described the model selection algorithm of *PcGets* as it has been developed by Hendry and Krolzig (2000) for linear singe-equation models. In the following we discuss how to generalize the algorithm for the analysis of multiple time series models.

2.4 General-to-specific reductions of stationary VARs

We investigate reductions of VAR(p) processes as defined in equation (1), where the GUM is an unrestricted VAR(p) model and the unknown DGP is a subset of the unrestricted VAR.

First, consider the case where the variance-covariance matrix of the system is diagonal, i.e. all $\sigma_{ij} = 0$ for $i \neq j$. Contemporaneous non-causality implies that the equations of the VAR are unrelated to each other. Thus the probability density function (pdf) of y_t conditional on its past Y_{t-1} is given by

$$f(y_t|Y_{t-1};\theta) = f(y_{1t}|Y_{t-1};\theta_1) \cdot \ldots \cdot f(y_{Kt}|Y_{t-1};\theta_K)$$

where the parameter vectors θ_k of the equations $k = 1,\ldots,K$ of the system can be varied freely. Consequently, all possible reductions of the system can be efficiently estimated by OLS, and model-selection procedures can be applied equation-by-equation without a loss in efficiency. Hence, *PcGets* can be used to model the system as in the single-equation framework it has been designed for. In section 3 this case is studied in a Monte Carlo experiment.

The situation is different in case of contemporaneous causality between the variables, i.e. some $\sigma_{ij} \neq 0$ for $i \neq j$. As weak exogeneity is lost, the equations of the VAR are only seemingly unrelated to each other. Since eliminating a variable in one equation effects the others, single-equation model selection procedures are inefficient. Hence, in this case, *PcGets* in its recent form does not offer an optimal implementation of the *Gets* methodology, though it still might deliver reasonable results (see the empirical illustration in section 4). An (asymptotically) efficient estimation procedure is provided by Estimated Generalized Least Squares (EGLS), see Lütkepohl (1991) for details. This has strong implications for model selection procedures.

So how could *PcGets* be extended to become a system procedure? Assuming that the VAR is covariance-stationary and the variance-covariance matrix is unrestricted, the proposed strategy is sketched in table 2. Note that n is the number of regressors in the system, which is Kp for a p-th order vector autoregression without deterministic terms, and m is the number of regressors excluded at the system-reduction step.

In the case of a vector system we first have to distinguish between (i) joint reductions of the system and (ii) reductions of the individual equations. In case of the reductions of the system we are interested

Table 2 *Gets* Algorithm for stationary VARs .

(1) **Reductions of the system**

System analysis of joint restrictions (OLS): #state vector $= n$

- Presearch for the exclusion of blocks of variables from the system
- F-test search for the exclusion of single variables from the system
- Diagnostics for the vector of residuals

(2) **Reductions of the equations**

Contemporaneous causality?

(a) System analysis (EGLS):
state vector $= K(n-m)$

(b) Single-equation analysis (OLS or EGLS with Σ from stage 1):
k^{th} state vector $= (n-m)$

PcGets-style multi-stage multi-path model selection

- Pre-search tests
- Multiple model reduction paths:
  - Sequential estimation and test of reductions
  - Encompassing
- Union testing
- Sub-sample evaluation

in a system analysis of cross-equation restrictions of the kind

$$\left[\begin{array}{ccc} a_{1j,i} & \cdots & a_{Kj,i} \end{array} \right] = \mathbf{0}_3',$$

where $a_{kj,i}$ is the (k,j)-th element of A_i. An acceptance of the restriction would exclude the regressor $y_{j,t-i}$ from all equations of the system. The system-reduction process can be implemented corresponding to the reduction process of *PcGets* of single-equation models: Starting with an OLS estimation of the unrestricted VAR(p) the search involves $n = Kp$ regressors. The multi-stage multi-path reduction process consists of a presearch for the exclusion of blocks of regressors from the system (cf. *Stage 0*), and an F-test search for the exclusion of single regressors from the system (cf. *Stage I and II*).

After imposing the system reductions of the first step, the single-equation reductions are then the results of *PcGets*-style multi-stage multi-path model selection procedures. The outcome of the test on contemporaneous causality (see Lütkepohl, 1991) decides whether the following reduction approach considers the system as a whole or each equation separately. In the former case, along each path the coefficient with the lowest remaining t-value of the *system* is checked

$$(k^*, j^*, i^*) := \arg \min_{k=1,\ldots,K} \min_{j=1,\ldots,K} \min_{i=1,\ldots,p} \mathtt{t}_{kj,i}.$$

If the coefficient $a_{k^*j^*,i^*}$ of regressor $y_{j,t-i}$ in equation k is insignificant, the coefficient is restricted to zero and the complete system is re-estimated by EGLS.

In the later case, each equation is reduced separately as in *PcGets*. Thus the coefficient with the lowest remaining t-value of the k-th *equation* is checked

$$(j^*, i^*) := \arg \min_{j=1,\ldots,K} \min_{i=1,\ldots,p} \mathtt{t}_{kj,i} \text{ for } k = 1, \ldots, K.$$

Table 3 *Gets* Algorithm for Cointegrated VAR(p) Processes .

(1) **Determination of the lag order p of the VAR** (AIC or matrix F-tests)

(2) **Johansen cointegration analysis**

- Testing for the cointegration rank r
- Estimation of the cointegration matrix $\tilde{\beta}$
- Mapping into the VECM representation

(3) **Multi-path search for redundant variables of the VECM system**

(4) **Multi-path search for redundant variables of the VECM equations**

Contemporaneous causality?
(a) System analysis (EGLS)
(b) Single-equation analysis (OLS or EGLS with Σ from stage 1)
PcGets-style multi-stage multi-path model selection

If the coefficient a_{kj^*,i^*} of regressor $y_{j,t-i}$ in equation k is insignificant, the coefficient is restricted to zero and the equation is re-estimated by OLS. Alternatively, EGLS could be used whereby the variance-covariance matrix is taken from the reduced, but otherwise unrestricted system.

As in *PcGets* the reduction process is a sequential simplification and testing procedure, where the diagnostics are constructed to test for the properties of the vector of residuals. The computer-automation of this algorithm appears to be a straightforward extension of the *PcGets* algorithm.

2.5 General-to-specific reductions of cointegrated VARs

The procedure proposed for stationary VARs can be easily extended to the analysis of cointegrated processes. An outline for such a procedure is presented in table 3. The important point is to introduce a cointegration step at the beginning of the procedure and then to map the cointegrated VAR into its vector equilibrium correction (VEC) representation. The *Gets* reduction process is then applied to the VECM.

Various methods for the cointegration analysis of multiple time series have been proposed in the literature. As *PcGets* itself is likelihood based, Johansen's concentrated-likelihood-function approach (see Johansen, 1995) is apparently the natural choice. Johansen's reduced rank procedures are based on an unrestricted VAR. Therefore a lag selection procedure to determine the order of the VAR precedes the cointegration analysis. This step involves a liberal sequential F-test procedure of block restrictions $A_i = 0$ for $i = h, h-1, \ldots$ or AIC model comparisons. Given the outcome of the cointegration analysis (cointegration rank r and cointegration matrix β), the analysis then focuses on reductions α, Γ_i in the corresponding (stationary) VECM($p-1$)

$$\Delta y_t = \alpha\left(\beta' y_{t-1}\right) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{NID}(\mathbf{0}, \Sigma).$$

The VECM becomes the new GUM. the *general-to-specific* reductions of stationary VARs discussed above are then applied to the VECM. Possible extensions could involve (identified) simultaneous equation models.

It is beyond the scope of this paper to present a computer implementation of the algorithm outlined in tables 2 and 3. However it should not be to difficult to extend *PcGets* accordingly, such that some first results can be expected for the near future. In the next section we will use simulation techniques to investigate the properties of *Gets* reductions of VARs in a Monte Carlo experiment. The DGP is chosen to allow the efficient use of *PcGets*. The simulation study will therefore give insights into the usefulness of *PcGets* for the analysis of multiple time series.

3 Monte Carlo results

3.1 Aim of the Monte Carlo

Although the sequential nature of *PcGets* and its combination of variable-selection and diagnostic testing has eluded most attempts at theoretical analysis, the properties of the *PcGets* model-selection process can be evaluated in Monte Carlo (MC) experiments. In the MC considered here, we aim to measure the 'size' and 'power' of the *PcGets* model-selection process, namely the probability of inclusion in the final model of variables that do not (do) enter the DGP.

3.2 Design of the Monte Carlo

To produce unbiased estimates of the properties of *PcGets*, we shall work with a VAR model with a diagonal variance-covariance matrix. The DGP is a Gaussian VAR(2) model

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 0.02 \\ 0.03 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.1 \\ 0.4 & 0.5 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0.25 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix},$$

where ε_t is a Gaussian white-noise process with variance-covariance matrix:

$$\Sigma = \begin{bmatrix} 0.09 & 0 \\ 0 & 0.04 \end{bmatrix}.$$

This DGP has been proposed by Lütkepohl (1991) and was reconsidered by Brüggemann and Lütkepohl (2000).

As the processes of y_{1t} and y_{2t} are contemporaneously uncorrelated, the parameters of the two equations can be estimated independently by OLS. It also implies that single-equation model-specification strategies are optimal and dominate system approaches which do not impose the $\sigma_{12} = \sigma_{21} = 0$ restriction.

The GUM is an unrestricted VAR(4) model (with intercept). The sample size T is 30 or 100 and the number of replications M is 1000. The model will be specified equation-by-equation using *PcGets*. Note that the k-th equation of the GUM is given by

$$y_{k,t} = \beta_{k,0} + \sum_{i=1}^{p} \sum_{j=1}^{K} \beta_{k,(i-1)*K+j} y_{j,t-i} + \varepsilon_{k,t},$$

or in matrix notation

$$\mathbf{y}_k = \mathbf{X}\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}_k.$$

Simplification can at best eliminate the nuisance regressors all or most of the time (size), yet retain the substance nearly as often as the DGP (power). The probability is low of detecting an effect that has a scaled population t-value less than the empirical selection criterion. When compared to missing an

effect with $|\mathsf{t}| > 4$ (say), 'missing' a variable with $|\mathsf{t}| < 2$ is attributable as a cost of inference, not a flaw of *Gets* type searches.

Recall that the t-test of the null $\beta_{kr} = 0$ versus the alternative $\beta_{kr} \neq 0$ is given by:

$$\mathsf{t}_{kr} = \frac{\widehat{\beta}_{kr}}{\widehat{\sigma}_{\beta_{kr}}} = \frac{\widehat{\beta}_{kr}}{\sqrt{\widehat{\sigma}^2_{\varepsilon_k} \left(\mathbf{X}'_k \mathbf{X}_k\right)^{-1}_{rr}}} = \frac{\widehat{\beta}_{kr} / \sqrt{\sigma^2_{\varepsilon_k} \left(\mathbf{X}'_k \mathbf{X}_k\right)^{-1}_{rr}}}{\sqrt{\widehat{\sigma}^2_{\varepsilon_k} / \sigma^2_{\varepsilon_k}}}.$$

The population value of the t-statistic is:

$$\mathsf{t}^*_{kr} = \frac{\beta_{kr}}{\sigma_{\beta_{kr}}} = \frac{\beta_{kr}}{T^{-\frac{1}{2}} (Q_k)^{-1/2}_{rr} \sigma_{\varepsilon_k}},$$

where the moment matrix $Q_k = \lim_{T \to \infty} \left(T^{-1} \mathbf{X}'_k \mathbf{X}_k\right)$ is assumed to exist.

Note that for the unrestricted VAR(2) we have that

$$Q = \begin{bmatrix} 1 & \boldsymbol{\mu}' \\ \boldsymbol{\mu} & \boldsymbol{\Gamma} + \boldsymbol{\mu}\boldsymbol{\mu}' \end{bmatrix}$$

where $\boldsymbol{\mu} = \mathbf{1}_2 \otimes \nu$ and $\boldsymbol{\Gamma}$ is the covariance matrix of (y_t, y_{t-1})

$$\boldsymbol{\Gamma} = \begin{bmatrix} \Gamma(0) & \Gamma(1) \\ \Gamma(-1) & \Gamma(0) \end{bmatrix},$$

with $\Gamma(i)$ as the autocovariance function of y_t at lag i. Thus the population t-values are for $T = 30$:

$$\mathsf{t}_\nu = \begin{bmatrix} 0.34 \\ 0.77 \end{bmatrix}, \mathsf{t}_{A_1} = \begin{bmatrix} 2.74 & 0.38 \\ 3.29 & 2.87 \end{bmatrix}, \mathsf{t}_{A_2} = \begin{bmatrix} 0 & 0 \\ 1.64 & 0 \end{bmatrix}.$$

For $T = 100$, the t-values are given by

$$\mathsf{t}_\nu = \begin{bmatrix} 0.62 \\ 1.40 \end{bmatrix}, \mathsf{t}_{A_1} = \begin{bmatrix} 5.00 & 0.70 \\ 6.00 & 5.23 \end{bmatrix}, \mathsf{t}_{A_2} = \begin{bmatrix} 0 & 0 \\ 3.00 & 0 \end{bmatrix}.$$

If the VAR is estimated with the zero-restrictions imposed, the following population t-values result for $T = 30$:

$$\mathsf{t}_\nu = \begin{bmatrix} 0.34 \\ 0.77 \end{bmatrix}, \mathsf{t}_{A_1} = \begin{bmatrix} 2.99 & 0.70 \\ 3.30 & 4.30 \end{bmatrix}, \mathsf{t}_{A_2} = \begin{bmatrix} . & . \\ 1.69 & . \end{bmatrix},$$

and for $T = 100$:

$$\mathsf{t}_\nu = \begin{bmatrix} 0.63 \\ 1.41 \end{bmatrix}, \mathsf{t}_{A_1} = \begin{bmatrix} 5.46 & 1.28 \\ 6.02 & 7.85 \end{bmatrix}, \mathsf{t}_{A_2} = \begin{bmatrix} . & . \\ 3.09 & . \end{bmatrix}.$$

Thus, for conventional critical values, four of the seven coefficients of the true model are insignificant for $T = 30$ and three for $T = 100$. Obviously, this will quite dramatically reduce the chance of finding the truth when starting from a general model.

3.3 Evaluation of the Monte Carlo

The evaluation of Monte Carlo experiments always involves measurement problems: see Hendry (1984) and Hendry and Krolzig (1999). A major decision concerns the basis of comparison: the 'truth' seems to be a natural choice, and both Lovell (1983) and Hoover and Perez (1999) measure how often the search finds the DGP exactly – or nearly. However, 'finding the DGP exactly' does not appeal to be a good choice of comparator, because it implicitly entails a basis where the truth is known, and one is *certain* that it is the truth. Rather, to isolate the costs of selection *per se*, we seek to match probabilities with the same procedures applied to testing the DGP. In each replication, the correct DGP equation is fitted, and the same selection criteria applied: we then compare the retention rates for DGP variables from *PcGets* with those that occur when no search is needed, namely when inference is conducted once for each DGP variable, and additional (non-DGP) variables are never retained.

Table 4 *PcGets* Properties.

Equation	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$
Sample size T	30	30	30	30	100	100
Nominal size α	0.10	0.10	0.05	0.05	0.05	0.05
Truth: DGP found	0.0150	0.0730	0.0030	0.0220	0.0030	0.2290
PcGets: DGP found	0.0020	0.0200	0.0020	0.0160	0.0020	0.1700
Truth: Dominated	0.6640	0.5990	0.7230	0.6820	0.7160	0.6250
PcGets: Dominated	0.0170	0.0300	0.0050	0.0120	0.0050	0.0450
Size	0.1058	0.1334	0.0870	0.1078	0.0713	0.0682
Power	0.3550	0.5843	0.3217	0.5643	0.4447	0.7750
Non-deletion prob.	0.4060	0.4490	0.3450	0.3720	0.3460	0.2570
Non-selection prob.	0.9820	0.9600	0.9940	0.9730	0.9910	0.7720

PcGets settings are reported in appendix A for $\alpha = 0.05$. For $\alpha = 0.1$, the nominal size of t and F specification test is increased to 10%.

Table 4 clarifies the 'success' and 'failure' of *PcGets*. In our experiments the probability to find the truth is between 17% and 0.2% – depending on the specification of the DGP and the number of observations. These figures seem to be small, but have to be compared to the probability of finding the DGP when starting the search from the true model which is in between 22.9% and 0.3%. Instead of focusing whether the DGP has been found or not, we prefer to check whether the deviation of the 'specific' model found by *PcGets* from the 'true' model nevertheless results in a sound model that, based on statistically criteria, could not have been improved by knowing the truth. We consider an encompassing test between the 'true' model and the 'specific' model found by *PcGets*. As long as *PcGets* is able to find a model that is not dominated by the 'true' model, the reduction process has been a success. If the specific model is dominated by the 'true' model, then the search algorithm has failed. Our results indicate that the risk to find a model which is dominated by the DGP is extremely small. In the case of the first variable and a nominal size of 5%, the risk is consistently less than one percent and for the second variable it is just 1.2% for $T = 30$ and 4.5% for $T = 100$. In contrast the model found by *PcGets* dominates the true model in between 59.9% and 72.5% of the cases. It remains to be said that by construction the outcome of *PcGets* always beats the unrestricted VAR(4) model. However, there might be a scope for further improvements by future developments.

The 'size' of *PcGets* (the average probability of selecting a Non-DGP variable) is with 6.8% to 10.8% slightly higher than the nominal size of 5%. Hendry (1995, p.490) suggested to make the significance level of the specification tests dependent on the sample size. We therefore replicated the Monte Carlo with *PcGets* for a nominal size of 10%. Indeed we get an empirical size of 10.6% and 13.3% which is much closer to the nominal size. We conclude that there is some support for Hendry's sugges-

tion, further investigations are required.

The 'power' of *PcGets* (the average probability of selecting a DGP variable) is in between 32.2% and 77.5%. However, the overall probability to miss an DGP variable is not very informative as the DGP variables have distinctively different population t-values and, hence, chances to be found. The reader is referred to table 5 which reports the probabilities of inclusion for the nine regressors of the GUM.

Table 5 Inclusion Probabilities.

Equation	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$
Sample size T	30	30	30	30	100	100
Nominal size α	0.10	0.10	0.05	0.05	0.05	0.05
$y_{1,t-1}$	0.645	0.874	0.614	0.842	0.995	1.000
$y_{1,t-2}$	0.109	0.538	0.091	0.493	0.080	0.852
$y_{1,t-3}$	0.088	0.150	0.067	0.134	0.074	0.060
$y_{1,t-4}$	0.104	0.128	0.088	0.097	0.069	0.065
$y_{2,t-1}$	0.141	0.725	0.121	0.754	0.166	0.994
$y_{2,t-2}$	0.118	0.126	0.092	0.099	0.083	0.073
$y_{2,t-3}$	0.109	0.117	0.098	0.100	0.065	0.075
$y_{2,t-4}$	0.107	0.146	0.086	0.109	0.057	0.068
Constant	0.279	0.200	0.230	0.168	0.173	0.254

PcGets settings are reported in appendix A for $\alpha = 0.05$. For $\alpha = 0.1$, the nominal size of t and F specification test is increased to 10%.

In the worst case scenario, $y_{1,t}$ and $T = 30$, the DGP involves the variable $y_{2,t-1}$ and the constant whose t-values (evaluated in the true model) are in the population 0.34 and 0.7, respectively. Even if one would start the truth, based on statistical criteria, these variables would be removed and only in 0.3% to 22.9% of the cases the estimated true model would be accepted. In a world like $y_{1,t}$, there is no way how a data-driven approach could detect that $y_{2,t-1}$ and the constant are part of the DGP. So, even for $T = 100$, the probability missing at least one of the DGP variables is 99.1% in the first equation and not much better in the second.

Overall, *PcGets* works more than satisfactory despite the presence of collinearity among the regressors. In table 6 we compare the statistical properties of *PcGets* with some of the single-equation 'Sequential Elimination of Regressors' (SER) strategy proposed by Brüggemann and Lütkepohl (2000).

Table 6 Comparison of *PcGets* and Brüggemann and Lütkepohl (2000) .

	PcGets				SER	
Equation	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$	$y_{1,t}$	$y_{2,t}$
Sample size T	30	30	30	30	30	30
Nominal size α	0.10	0.10	0.05	0.05	—	—
power						
$y_{1,t-1}$	0.645	0.874	0.614	0.842	0.652	0.868
$y_{2,t-1}$	0.141	0.725	0.121	0.754	0.171	0.743
$y_{1,t-2}$		0.538		0.493		0.589
size	0.1058	0.1334	0.0870	0.1078	0.1328	0.1584

Source: Table 2 in Brüggemann and Lütkepohl (2000);
selection procedure: SER/TP; criterion: SC.
PcGets settings are reported in appendix A.

With empirical sizes of 13.28% and 15.84%, the SER strategy is clearly more liberal than *PcGets* is. For the latter, the empirical size is 8.7% in the first equation and 10.78% in the second (with $\alpha = 5\%$ and $T = 30$). At a nominal size of 10%, the empirical size is 10.58% and 13.33%, respectively. Given the liberal significance levels, SER achieves a slightly higher power. However the size-power trade-off

is in favor of *PcGets*: While the average probability of including a DGP variable increases by 13.7%, the average probability of including a nuisance variable surges by 31.8%. At a nominal size of 10%, SER gains 2 percentage points in power, but has a size which is higher by 2 percentage points. We conclude that *PcGets* leads to reasonable results which dominate the outcome of stepwise-regression type procedures. To investigate this issue further, it would be desirable to look larger sample sizes, well-specified DGPs that offer a fair chance to be found, and other alternative strategies.

4 Empirical Illustration

To illustrate the *Gets* procedures for subset VARs proposed in the foregoing section, we will now use *PcGets* to analyze the US monetary system considered by Christiano, Eichenbaum and Evans (1996) and Brüggemann and Lütkepohl (2000). Christiano *et al.* (1996) analyzed the effects of monetary policy shocks in an unrestricted VAR(4) of the following variables:

$$y_t = (\mathrm{dgp}_t, \mathrm{p}_t, \mathrm{pcom}_t, \mathrm{ff}_t, \mathrm{nbrd}_t, \mathrm{tr}_t, \mathrm{m1}_t),$$

where dgp_t is the log of real GDP, p_t the log of the GDP deflator, pcom_t the log of a commodity price index, ff_t the fed funds rate, nbrd_t the negative log of unborrowed reserves, tr_t the log of total reserves and $\mathrm{m1}_t$ the log of M1.

The data are in levels and, therefore, show trending behavior. This could potentially cause a problem as, to date, *PcGets* conducts all inferences as I(0). But most selection tests will in fact be valid even when the data are I(1), given the results in, say, Sims, Stock and Watson (1990). Only t- or F-tests for an effect that corresponds to a unit root require non-standard critical values. Similarly, Wooldridge (1999) shows that diagnostic tests on the GUM (and presumably simplifications thereof) remain valid even for integrated time. The empirical example on integrated data considered here does not reveal problems, but in principle it would be useful to implement cointegration tests and appropriate transformations (see the discussion in section 2.5).

Starting with the unrestricted VAR(4) as the GUM, *PcGets* sets 115 zero restrictions, it finds 44 coefficients that are significant at the 1% percent level and 6 that are significant at the 5% percent level. Seven insignificant coefficients are included as setting them zero would result in an invalid reduction. The details are reported in table 7. In contrast, from the 203 coefficients of the unrestricted VAR(4) only 18 are significant at the 1% percent level and 4 at the 5% percent level.

It is worth comparing the outcome of *PcGets* to other reduction procedures. We consider again the SER strategy of Brüggemann and Lütkepohl (2000). The results here confirm the findings of the Monte Carlo: the SER strategy is very liberal: 115 null restrictions are set, 30 coefficients are insignificant, 24 coefficients are significant at the 5% percent level and 34 coefficients are significant at the 1% percent level. Interestingly, the SER strategy does not find all the effects identified by *PcGets*. Thus the *PcGets* reduction is not nested within the SER reduction. Further research is required on the pros and cons of the procedures.

As pointed out earlier, *PcGets* has originally been developed for single-equation models. It is efficient only if the equations of the VAR are unrelated, i.e. the variance-covariance matrix is diagonal. But this is unlikely the case here. Therefore it is interesting to see how strongly the results of single-equation estimation (OLS as in *PcGets*) and system estimations (EGLS,FIML) differ. Significant changes would indicate that the single-equation selection procedure used here is problematic and that a system approach should be used instead.

Table 8 reports the OLS, EGLS and FIML estimation results of the dominant valid reduction found by *PcGets*. The results show that the system and single-equation estimations are consistently close,

Table 7 Zero restrictions set by *PcGets* and Brüggemann and Lütkepohl (2000).

		gdp	p	pcom	ff	nbrd	tr	m1	ν
	lag	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	
VAR(4)	gdp	* . . .	. . . .	. . . .	. * . .	. . . .	. . . .	. * . .	.
	p	. . . .	* . . .	* . . .	. . . .	. . . .	. . . .	. . . .	*
	pcom	. . . .	. . . .	* . . .	. . * .	. . . .	. . . .	. . . .	.
	ff	. . . .	. . . .	* . . .	* . . .	. . . .	. . . .	. . + .	.
	nbrd	. . . .	. . . .	+ . . .	. . . .	* * . .	. + . .	. . . .	.
	tr	. . . .	. . . .	. . . .	. . . .	. . . .	* . . .	* . . .	.
	m1	. . . .	. . . .	. + . .	* * . .	. . . .	. . . .	* . . .	*
PcGets	gdp	* 0 0 0	0 0 0 0	0 0 0 0	0 * 0 0	0 0 0 0	0 0 0 *	0 * 0 0	0
	p	0 0 0 *	* * 0 0	* 0 * 0	0 0 0 0	0 0 0 0	0 0 0 +	0 0 0 *	*
	pcom	0 0 0 .	+ + 0 0	* 0 0 *	0 0 0 0	0 . 0 0	0 + 0 0	. 0 0 0	0
	ff	* 0 * 0	0 0 * 0	* * 0 0	* 0 * 0	0 * 0 0	0 * 0 0	* 0 0 *	0
	nbrd	* 0 0 *	0 0 0 0	* 0 0 *	0 0 * 0	* * 0 0	0 * 0 0	* * 0 0	0
	tr	* 0 0 *	0 0 0 0	+ . 0 0	0 0 0 .	0 0 0 0	* 0 0 0	* * 0 0	0
	m1	0 0 0 *	0 0 0 *	+ . 0 0	* * 0 0	0 0 0 0	. 0 0 0	* * 0 0	*
SER	gdp	* 0 0 0	0 0 0 0	0 0 0 0	0 * 0 0	0 0 0 0	. 0 0 +	0 * 0 0	.
	p	* 0 * 0	* . . +	* 0 * 0	. . 0 0	0 + . 0	+ 0 . .	0 0 0 .	*
	pcom	0 0 0 0	+ . 0 0	* 0 0 +	0 0 0 0	0 0 0 0	0 . 0 .	0 + 0 +	.
	ff	+ 0 * 0	0 0 + .	* + 0 0	* 0 * 0	. + 0 0	. * 0 0	. 0 + +	.
	nbrd	+ 0 0 *	0 0 + 0	* 0 + 0	0 0 * 0	* * 0 0	* * 0 0	0 * 0 0	.
	tr	* 0 0 *	* + 0 0	0 0 0 0	+ + . 0	0 0 . 0	* 0 . 0	* . . 0	.
	m1	0 0 0 *	. 0 0 +	+ + . 0	* * . 0	0 0 0 .	+ 0 0 0	* 0 * 0	*

Legend: 0 Coefficient is set to zero.
. Coefficient is insignificant at the 5% percent level.
+ Coefficient is significant at the 5% percent level.
* Coefficient is significant at the 1% percent level.

indicating that the loss in efficiency by modeling the system equation-by-equation is very limited.[3] In other words, a system model selection procedure would presumably have come up with a similar set of zero-restrictions. It remains to report that the LR test of over-identifying restrictions is clearly accepted with $\chi^2(144) = 39.5[0.5904]$ for the FIML estimates and $\chi^2(144) = 139.6[0.5881]$ in case of the EGLS estimation.

The variance-covariance matrix of the residuals of the selected model as estimated by EGLS is reported in table 9. It is clearly non-diagonal rejecting any hypothesis of contemporal non-causality.[4]

Figure 1 shows the response of all system variables to a monetary shock in the unrestricted VAR and the model selected by *PcGets*. Plotted are the orthogonalized responses to a unit shock in the federal funds rate ff, where in case of the reduced system the variance-covariance matrix of EGLS estimation has been used. The solid line represents the response in the unrestricted VAR and the dashed line the response in the reduced VAR. The increase in the federal funds rate causes a persistent drop in GDP and, with some delay, a smooth decline in the aggregate and commodity price indices. This is quite in line with predictions of economic theory and the findings of Christiano *et al.* (1996) and Brüggemann and Lütkepohl (2000). Interestingly, the responses of the VAR and its reduction show a very similar pattern. There is no indication of a bias caused by the reduction. When compared to the impulse response of the system found by Brüggemann and Lütkepohl (2000), the similarity of the results indicates that it is

[3]Note that there is strong indication that the VAR(4) itself is misspecified. Four of the seven equations show a break at the middle of the sample. Some of the unrestricted equations also non-normality and ARCH effects of the residuals.

[4]The corresponding matrix of the correlations of the FIML residuals is very similar (deviations are all less 0.01) and, therefore, not reported here.

Table 8 Model selected by *PcGets* .

Equation	Misspecification of the GUM	Variable	OLS		EGLS		FIML	
			Coeff	t-value	Coeff	t-value	Coeff	t-value
gdp		gdp_1	0.9843	267.07	0.9851	266.	0.9851	266.
		ff_2	-0.2022	-8.00	-0.1991	-7.83	-0.1989	-7.83
		tr_4	-0.0716	-4.90	-0.0684	-4.66	-0.0681	-4.65
		$m1_2$	0.0631	4.94	0.0603	4.70	0.0601	4.69
		σ	0.0078		0.0079		0.0079	
p		gdp_4	0.0415	6.20	0.0416	6.62	0.0416	6.63
		p_1	1.2293	14.52	1.2308	15.2	1.2288	15.2
		p_2	-0.2412	-2.89	-0.2449	-3.08	-0.2433	-3.05
		$pcom_1$	0.0938	6.93	0.0921	7.02	0.0926	7.05
		$pcom_3$	-0.0720	-4.91	-0.0692	-4.92	-0.0695	-4.95
		tr_4	0.0233	2.04	0.0232	2.13	0.0229	2.10
		$m1_4$	-0.0465	-2.67	-0.0456	-2.75	-0.0451	-2.72
		Constant	-0.1795	-5.21	-0.1810	-5.64	-0.1818	-5.67
		σ	0.0023		0.0023		0.0023	
pcom	Chow(1976:3)	gdp_4	-0.0362	-1.87	-0.0197	-1.13	-0.0193	-1.11
	Normality	p_1	0.8799	2.42	0.8858	2.64	0.8886	2.64
	ARCH 1-4	p_2	-0.9467	-2.35	-0.9040	-2.45	-0.9051	-2.45
		$pcom_1$	1.1550	20.15	1.1305	21.0	1.1295	20.9
		$pcom_4$	-0.1716	-2.65	-0.1581	-2.65	-0.1578	-2.64
		$nbrd_2$	-0.0418	-1.55	-0.0420	-1.66	-0.0421	-1.67
		tr_2	-0.1391	-2.37	-0.0978	-1.82	-0.0969	-1.80
		$m1_1$	0.1559	1.89	0.0869	1.17	0.0850	1.14
		σ	0.0120		0.0121		0.0121	
ff	Chow(1976:3)	gdp_1	0.2725	2.68	0.1971	2.23	0.1865	2.12
	Normality	gdp_3	-0.3404	-3.40	-0.2556	-2.93	-0.2448	-2.82
	ARCH 1-4	p_3	-0.2030	-2.75	-0.1685	-2.67	-0.1667	-2.66
		$pcom_1$	0.3989	4.36	0.3528	4.34	0.3482	4.30
		$pcom_2$	-0.3496	-3.51	-0.3137	-3.55	-0.3101	-3.53
		ff_1	0.5487	7.67	0.5913	9.75	0.5978	9.95
		ff_3	0.4225	5.01	0.3608	4.87	0.3527	4.79
		$nbrd_2$	-0.0896	-2.76	-0.0765	-2.55	-0.0753	-2.51
		tr_2	-0.2661	-3.17	-0.2323	-3.19	-0.2314	-3.20
		$m1_3$	0.7013	3.86	0.6848	4.56	0.6903	4.66
		$m1_4$	-0.4139	-3.21	-0.4371	-4.12	-0.4430	-4.24
		σ	0.0129		0.0128		0.0128	
nbrd	Chow(1976:3)	gdp_1	0.6713	3.21	0.5769	3.18	0.5714	3.16
	Normality	gdp_4	-0.7145	-3.63	-0.6046	-3.51	-0.5979	-3.48
	ARCH 1-4	$pcom_1$	0.4684	3.25	0.4244	3.34	0.4139	3.28
		$pcom_3$	-0.5584	-3.42	-0.4896	-3.50	-0.4781	-3.45
		ff_3	0.5971	3.20	0.5149	3.22	0.5179	3.25
		$nbrd_1$	0.5533	5.86	0.5740	7.87	0.5797	8.08
		$nbrd_2$	-0.3044	-3.17	-0.2743	-3.36	-0.2780	-3.41
		tr_2	-0.9191	-4.27	-0.8193	-4.51	-0.8134	-4.50
		$m1_1$	-1.2318	-3.32	-1.2500	-3.87	-1.2482	-3.89
		$m1_2$	1.4344	3.09	1.3898	3.50	1.3837	3.51
		σ	0.0355		0.0353		0.0353	
tr	Chow(1976:3)	gdp_1	-0.3165	-3.47	-0.2780	-4.03	-0.2746	-4.08
		gdp_4	0.2401	2.86	0.1987	3.06	0.1952	3.08
		$pcom_1$	-0.2064	-2.36	-0.1911	-2.30	-0.1897	-2.28
		$pcom_2$	0.1302	1.39	0.1046	1.19	0.1025	1.17
		ff_4	-0.1336	-1.78	-0.0762	-1.32	-0.0711	-1.26
		tr_1	0.7584	11.89	0.7516	13.2	0.7517	13.2
		$m1_1$	0.8761	6.50	0.8565	6.94	0.8498	6.93
		$m1_2$	-0.5872	-4.31	-0.5539	-4.52	-0.5468	-4.50
		σ	0.0145		0.0145		0.0145	
m1		gdp_4	0.0848	3.98	0.0860	5.04	0.0874	5.19
		p_4	0.1618	2.93	0.1836	4.41	0.1872	4.59
		$pcom_1$	-0.1297	-2.37	-0.1366	-2.75	-0.1381	-2.77
		$pcom_2$	0.0860	1.35	0.0816	1.46	0.0818	1.46
		ff_1	-0.2040	-4.27	-0.1517	-4.09	-0.1445	-3.96
		ff_2	0.1537	3.09	0.1229	3.24	0.1178	3.17
		tr_1	0.0660	1.64	0.0733	2.11	0.0755	2.18
		$m1_1$	1.0477	13.16	1.0223	15.8	1.0163	15.8
		$m1_3$	-0.2307	-5.06	-0.2199	-5.56	-0.2182	-5.54
		Constant	-0.2695	-3.61	-0.2556	-4.55	-0.2578	-4.69
		σ	0.0074		0.0074		0.0074	

PcGets settings are reported in appendix A.

Table 9 EGLS: correlation of residuals (standard deviations on diagonal).

	gdp	p	pcom	ff	nbrd	tr	m1
gdp	0.0079	0.0591	0.1192	0.2224	0.1064	-0.0527	0.1040
p	0.0591	0.0023	0.3571	0.1053	0.0574	-0.0098	0.0943
pcom	0.1192	0.3571	0.0121	0.3082	0.2551	-0.0344	0.0628
ff	0.2224	0.1053	0.3082	0.0128	0.4409	0.0059	-0.0938
nbrd	0.1064	0.0574	0.2551	0.4409	0.0353	-0.4656	-0.2642
tr	-0.0527	-0.0098	-0.0344	0.0059	-0.4656	0.0145	0.6415
m1	0.1040	0.0943	0.0628	-0.0938	-0.2642	0.6415	0.0074

reasonable to be tougher on potential nuisance variables. As fewer parameters have to be estimated, the responses are estimated even more precisely. Overall, *PcGets* seems to be useful in specifying VARs, analyzing impulse responses and Granger causality.

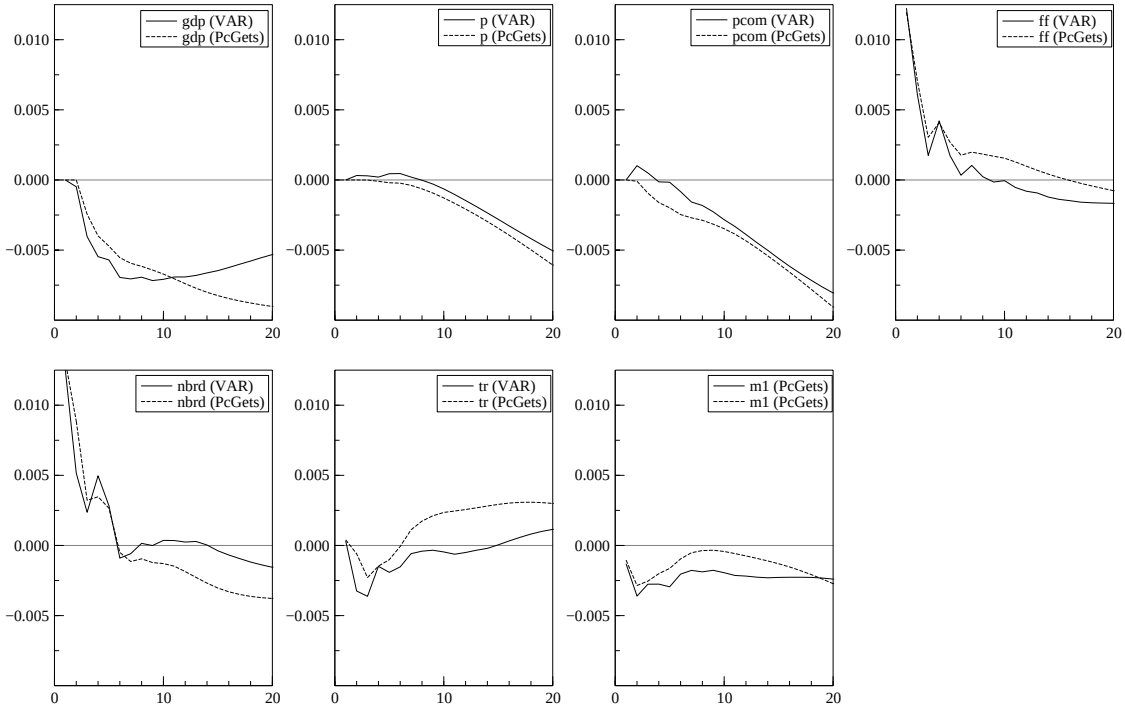

Figure 1 Response to a monetary policy shock.

5 Conclusions

The foregoing small example should have made clear that economically and statistically useful models can be obtained by *PcGets* without a great burden in computation time and a minimal loss in their statistical properties when compared to the true model. Even in the case of highly-dimensional systems, the very few, but significant parameters found by *PcGets* can be sufficient to describe the dynamics of the system. The parsimoniously specified VAR allows precise impulse-responses, delivers informative forecast intervals and provides the basis for powerful tests.

The aim of the paper was to propose and evaluate computerized model-selection strategies for subset

VARs, to see if they worked well, indifferently, or failed badly. The results come much closer to the first: the diagnostic-test operational characteristics are fine; selection-test probabilities match those relevant to the DGP; and deletion-test probabilities show 1% retention at a nominal 1% when no sub-sample testing is used. Thus, we deem *PcGets* successful, and deduce that the underlying methodology is appropriate for the reduction of VAR processes.

Nevertheless, this is a first attempt: consequently, we believe it is feasible to circumvent the baseline nominal selection probabilities. First, since diagnostic tests must be insignificant at every stage to proceed, *PcGets* avoids spurious inclusion of a variable simply because wrong standard errors are computed (e.g., from residual autocorrelation). Thus, it could attain the same lower bound as in a pure vector white-noise setting, since every selection must remain both congruent and encompassing. Secondly, following multiple paths reduces the overall size, relative to stepwise-regression-type strategies, despite the hugely increased number of selection (and diagnostic) tests conducted. Intuitively, the iterative loops around sequences of path searches could be viewed as 'sieves' of ever-decreasing meshes filtering out the relevant from the irrelevant variables: as an analogy, first large rocks are removed, then stones, pebbles, so finally only the gold dust remains. Thirdly, post-selection tests may further reduce the probability of including non-DGP variables below the nominal size of selection t-tests, at possible costs in the power of retaining relevant variables, and possibly the diagnostics becoming significant.

Further work on *Gets* and other strategies – such as just using information criteria to select – for stationary and cointegrated multiple time series are merited. More detailed Monte Carlo studies are required to investigate the impacts of contemporaneous causality, integration and cointegration. But the door is open – and we anticipate some fascinating developments will follow for model selection.

Acknowledgments

I am grateful to Ralf Brüggemann, David Hendry and Helmut Lütkepohl for helpful discussions. Special thanks to Ralf Brüggemann for providing me with the data and unpublished results from Brüggemann and Lütkepohl (2000). The computations reported in this paper were carried out with PcGets1, see Hendry and Krolzig (2000), and PcFiml10, see Doornik and Hendry (1999). Financial support from the UK Economic and Social Research Council under grant L138251009 is gratefully acknowledged.

References

Akaike, H. (1985). Prediction and entropy. In Atkinson, A. C., and Fienberg, S. E. (eds.), *A Celebration of Statistics*, C.H. 1. New York: Springer-Verlag.

Bock, M. E., Yancey, T. A., and Judge, G. C. (1973). Statistical consequences of preliminary test estimators in regression. *Journal of the American Statistical Association*, **68**, 109–116.

Brüggemann, R., and Lütkepohl, H. (2000). Lag selection in subset VAR models with an application to a U.S. monetary system. SFB 373 Discussion Paper 37/2000, Humboldt Universität zu Berlin.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1996). The effects of monetary policy shocks: Evidence from the flow of funds. *Review of Economics and Statistics*, **78**, 16–34.

Doornik, J. A., and Hendry, D. F. (1999). *Modelling Dynamic Systems using PcFiml: Volume II*. London: Timberlake Consultants Press.

Hannan, E. J., and Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, 190–195.

Hendry, D. F. (1984). Monte Carlo experimentation in econometrics. In Griliches, Z., and Intriligator, M. D. (eds.), *Handbook of Econometrics*, Vol. 2–3, C.H. 16. Amsterdam: North-Holland.

Hendry, D. F. (1993). *Econometrics: Alchemy or Science?* Oxford: Blackwell Publishers.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F., and Krolzig, H.-M. (1999). Improving on *'data mining reconsidered'* by K.D. Hoover and S.J. Perez. *Econometrics Journal*, **2**, 167–191.

Hendry, D. F., and Krolzig, H.-M. (2000). Comuter-automated econometric modelling using PcGets. Discussion Paper, Department of Economics, University of Oxford.

Hoover, K. D., and Perez, S. J. (1999). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. *Econometrics Journal*, **2**, 167–91.

Johansen, S. (1995). *Likelihood based Inference on Cointegration in the Vector Autoregressive Model*. Oxford: Oxford University Press.

Judge, G. G., and Bock, M. E. (1978). *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*. Amsterdam: North Holland Publishing Company.

Krolzig, H.-M., and Hendry, D. F. (2000). Computer automation of general-to-specific model selection procedures. Forthcoming, Journal of Economic Dynamics and Control.

Leamer, E. E. (1983). Let's take the con out of econometrics. *American Economic Review*, **73**, 31–43. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Lovell, M. C. (1983). Data mining. *Review of Economics and Statistics*, **65**, 1–12.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Mayo, D. (1981). Testing statistical testing. In Pitt, J. C. (ed.), *Philosophy in Economics*, pp. 175–230: D. Reidel Publishing Co. Reprinted as pp. 45–73 in Caldwell B. J. (1993), *The Philosophy and Methodology of Economics*, Vol. 2, Aldershot: Edward Elgar.

Pagan, A. R. (1987). Three econometric methodologies: A critical appraisal. *Journal of Economic Surveys*, **1**, 3–24. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 462–464.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1–48. Reprinted in Granger, C. W. J. (ed.) (1990), *Modelling Economic Series*. Oxford: Clarendon Press.

Sims, C. A., Stock, J. H., and Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, **58**, 113–144.

White, H. (1990). A consistent model selection. In Granger, C. W. J. (ed.), *Modelling Economic Series*, pp. 369–383. Oxford: Clarendon Press.

Wooldridge, J. M. (1999). Asymptotic properties of some specification tests in linear models with integrated processes. In Engle, R. F., and White, H. (eds.), *Cointegration, Causality and Forecasting*, pp. 366–384. Oxford: Oxford University Press.

Calibration of PcGets

Testimation algorithm	F presearch testing	FALSE
	Sample split analysis	FALSE
	Sample-size adjusted significance levels	FALSE
	Outlier correction	FALSE
	Model selection criterion	SC
Significance levels	t - tests	0.0500
	F - tests	0.0500
	F - test of GUM	0.7500
	Encompassing test	0.0500
	Diagnostics (high)	0.0100
	Diagnostics (low)	0.0050
F presearch tests	Significance level of F-test (step 1)	0.9000
	Significance level of F-test (step 2)	0.7500
	Marginal t-prob (step 1)	0.1000
	Marginal t-prob (step 2)	0.0500
	Two-step presearch testing	TRUE
Sample split analysis	Significance level	0.1000
	Size of the subsample (fraction)	0.7500
	Penalty for failed t-test in full sample	0.2000
	Penalty for failed t-test in subsample 1	0.4000
	Penalty for failed t-test in subsample 2	0.4000
Block search	Check groups with t-pvals > 0.90	TRUE
	Check groups with t-pvals > 0.70	TRUE
	Check groups with t-pvals > 0.50	TRUE
	Check groups with t-pvals > 0.25	TRUE
	Check groups with t-pvals > 0.10	TRUE
	Check groups with t-pvals > 0.05	TRUE
	Check groups with t-pvals > 0.01	TRUE
	Check groups with t-pvals > 0.001	TRUE
Diagnostic tests	Chow test 1	TRUE
	Chow test 2	TRUE
	Portmanteau	FALSE
	Normality	TRUE
	AR test	TRUE
	ARCH test	TRUE
	Hetero test	TRUE
Test options	Chow test breakpoint 1	0.50
	Chow rest breakpoint 2	0.90
	Portmanteau max lag	12
	AR test min lag	1
	AR test max lag	4
	ARCH test min lag	1
	ARCH test max lag	4