Journal of Econometrics 99 (2000) 39–61

JOURNAL OF Econometrics

www.elsevier.nl/locate/econbase

Consistent cross-validatory model-selection for dependent data: hv-block cross-validation

Jeff Racine*

Department of Economics, University of South Florida, 4202 East Fowler Avenue, Tampa, FL 33620, USA

Abstract

This paper considers the impact of Shao's (1993) recent results regarding the asymptotic inconsistency of model selection via leave-one-out cross-validation on h-block cross-validation, a cross-validatory method for dependent data proposed by Burman, Chow and Nolan (1994, Journal of Time Series Analysis 13, 189–207). It is shown that h-block cross-validation is inconsistent in the sense of Shao (1993, Journal of American Statistical Association 88(422), 486–495) and therefore is not asymptotically optimal. A modification of the h-block method, dubbed 'hv-block' cross-validation, is proposed which is asymptotically optimal. The proposed approach is consistent for general stationary observations in the sense that the probability of selecting the model with the best predictive ability converges to 1 as the total number of observations approaches infinity. This extends existing results and yields a new approach which contains leave-one-out cross-validation, leave-n_v-out cross-validation, and h-block cross-validation as special cases. Applications are considered. © 2000 Elsevier Science S.A. All rights reserved.

JEL classification: C5; C51; C52

Keywords: Cross-validation; Dependence; Model-selection; Prediction

* Corresponding author. Tel.: 813-974-6555; fax: 813-974-6510.

E-mail address: jracine@coba.usf.edu (J. Racine).

PII: S0304-4076(00)00030-0

1. Introduction

In a recent series of articles (Shao, 1993, 1996), Shao addresses the issue of model selection in a simple linear regression context assuming independent identically distributed (*iid*) errors, with extensions to nonlinear settings. In these articles the author demonstrates the result that a number of existing data-driven methods of model selection are 'asymptotically inconsistent in the sense that the probability of selecting the model with the best predictive ability does not converge to 1 as the total number of observations $n \to \infty$ (Shao, 1993, p. 486)' and are 'inconsistent in the sense that the probability of selecting the optimal subset of variables does not converge to 1 as $n \to \infty$' (Shao, 1996, p. 655). The primary implications of these findings for applied researchers are twofold; a loss of power when testing hypotheses, and forecasts which are not as accurate as they otherwise could be even when one possesses an abundance of data.

This paper considers the impact of Shao's results on h-block cross-validation, a cross-validatory method for dependent data recently proposed by Burman et al. (1994) which has widespread potential application ranging from selection of the order of an autoregression for linear time-series models to selection of the number of terms in nonlinear series approximations for general stationary processes.

Burman et al. (1994, p. 354, Remark 2) state in their paper that 'Conditions for asymptotic optimality for this proposal remain open.'. In this paper it is demonstrated how h-block cross-validation is not asymptotically optimal in the sense of Shao (1993). A modification of the h-block method is proposed which is asymptotically optimal, and this new method is dubbed 'hv-block' cross-validation. This new approach provides a general framework for cross-validation which contains leave-one-out cross-validation (Allen, 1974; Stone, 1974, 1977; Geisser, 1975; Whaba and Wold, 1975), leave-n_v-out cross-validation[1] (Breiman et al., 1984, Chapter 3,8; Burman, 1989, 1990; Burman and Nolan, 1992; Geisser, 1975; Zhang, 1993), and h-block cross-validation (Burman et al., 1994; Racine, 1997) as special cases. Simulations are conducted to demonstrate the nature of the inconsistency and the value added by the proposed method. A modest application is conducted for determining the order of an autoregression for G7 exchange rate data.

2. Background

Following the notation of Shao (1993), consider a linear model of the form

$$\begin{aligned} y_i &= \mathrm{E}[y|x_i] + \varepsilon_i \\ &= x_i'\beta + \varepsilon_i, \quad i = 1, \ldots, n, \end{aligned} \tag{1}$$

[1] Similar variants include 'r-fold', 'v-fold', and 'multifold' cross-validation, and the 'repeated learning-testing criterion'.

where y_i is the response, x_i is a $p \times 1$ vector of predictors, β is a $p \times 1$ vector of unknown parameters, ε_i is a mean zero disturbance term with constant variance σ^2, and n denotes the sample size.

Let $\alpha \in \mathbb{N}^{p_\alpha}$ denote a subset of $\{1, \dots, p\}$ of size p_α and let $x_{i\alpha}$ be the subvector of x_i containing all components of x_i indexed by the integers in α. A model corresponding to α shall be called 'model α' for simplicity, and is given by

$$y_i = x'_{i\alpha}\beta_\alpha + \varepsilon_{i\alpha}, \quad i = 1, \dots, n. \tag{2}$$

If the true data generating process is linear in that the conditional mean can be expressed as $\mathrm{E}[y|x_i] = x'_i\beta$, then the problem of model/variable selection becomes one of finding the 'correct model', that is, the model α for which

$$y_i = \mathrm{E}[y|x_i] = x'_{i\alpha}\beta_\alpha + \varepsilon_{i\alpha}, \quad i = 1, \dots, n. \tag{3}$$

Starting from the set of predictors $x_i \in \mathbb{R}^{\mathrm{p}}$ assumed to contain those for the correct model, Shao (1993) considers all $2^{\mathrm{p}} - 1$ possible models of form (2) corresponding to a subset α and calls the class of such models $\mathcal{M}_\alpha$. Shao (1993) divides all models in this class into two categories:

1. Category I: Those for which a predictor belongs in the set of conditioning predictors for $\mathrm{E}[y|x_i]$ but does not appear in $x_{i\alpha}$.[2]
2. Category II: Those for which the set of conditioning predictors includes all relevant predictors and in addition may include predictors which do not belong in the set of conditioning predictors for $\mathrm{E}[y|x_i]$ but which appear in $x_{i\alpha}$.[3]

Shao (1993, p. 487) defines the 'optimal model' $\mathcal{M}_*$ as the 'model in Category II with the smallest dimension', and this optimal model $\mathcal{M}_*$ will possess the smallest expected prediction error of any model α. Cross-validation estimates the expected prediction error for a model α, and cross-validatory model selection proceeds by selecting that model α with smallest estimated expected prediction error.

One of the most widely used variants of cross-validation is leave-one-out cross-validation. The leave-one-out cross-validation function is given by

$$\begin{aligned} CV_1 &= \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_{i(-i)})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} (y_i - x'_i\hat{\beta}_{(-i)})^2, \end{aligned} \tag{4}$$

[2] 'At least one non-zero component of β is not in β_α' (Shao, 1993, p. 487).

[3] 'β_α contains all non-zero components of β' (Shao, 1993, p. 487).

where $\hat{y}_{i(-i)}$ is the prediction of y_i when the ith observation is deleted from the data set, and where $\hat{\beta}_{(-i)}$ is the least-squares estimator obtained by deleting the ith observation from the data set.

Letting $\mathscr{M}_{cv_1}$ denote a model selected via leave-one-out cross-validation and assuming that standard regularity conditions hold, Shao (1993, p. 488) demonstrates that

$$\lim_{n\to\infty} pr(\mathscr{M}_{cv_1} \text{ is in Category I}) = 0, \tag{5}$$

that is, asymptotically the probability of selecting a model which excludes a relevant predictor when the assumed set of predictors nests the relevant set of predictors is zero. Furthermore, he demonstrates that

$$\lim_{n\to\infty} pr(\mathscr{M}_{cv_1} \text{ is in Category II but is not } \mathscr{M}_*) > 0 \tag{6}$$

and therefore that

$$\lim_{n\to\infty} pr(\mathscr{M}_{cv_1} = \mathscr{M}_*) \neq 1, \tag{7}$$

that is, asymptotically the probability of including an irrelevant predictor is not equal to zero unless the dimension of the relevant set of predictors p_α is equal to the dimension of the full set of assumed predictors p. This result has the direct implication that leave-one-out cross-validation will tend to select unnecessarily large models. Finally, it is noted that since numerous criteria such as the widely used Akaike Information Criterion (AIC) of Akaike (1974) are asymptotically equivalent to leave-one-out cross-validation, the same critique also applies to these criteria.

Result (5) states that the probability of asymptotically selecting a model which will yield biased predictions is zero given the assumptions in Shao (1993). Result (6) states that leave-one-out cross-validation will not be able to discriminate between models which include irrelevant variables but which otherwise are properly specified in that they include all relevant predictors even as $n \to \infty$. That is, many data-driven methods of model selection will select mean square error sub-optimal models with positive probability even asymptotically. Note that this holds only in the case where the number of conditioning variables is fixed as $n \to \infty$ (Shao, 1993, p. 486, second to last paragraph) thereby absolving series approximations and some other semi-parametric models whose dimension p increases with n from this critique.

3. Consistent cross-validatory model-selection for independent data: Leave-n_v-out cross-validation

Shao (1993) considers leave-n_v-out cross-validation in which the model is fit on $n_c = n - n_v$ observations and then the prediction error is determined using the remaining n_v observations not used to fit the model, and this is conducted for all ${}_nC_{n_v}$ possible validation sets. Letting v denote the collection of such subsets, the leave-n_v-out cross-validation function[4] is given by

$$CV_v = \frac{1}{{}_nC_{n_v} n_v} \sum_{\text{all } v \in \mathsf{v}} \|Y_v - \hat{Y}_{v(-v)}\|^2$$
$$= \frac{1}{{}_nC_{n_v} n_v} \sum_{\text{all } v \in \mathsf{v}} \|Y_v - X_v \hat{\beta}_{(-v)}\|^2, \tag{8}$$

where $\|a\| = \sqrt{a'a}$, $\hat{\beta}_{(-v)}$ is the least-squares estimator obtained by deleting v observations from the data set, and (Y_v, X_v) is the deleted sample. As a practical matter this approach becomes intractable as $n \to \infty$ since $\lim_{n\to\infty} {}_nC_{n_v} = \infty$ for finite n_v, and numerous simplifications have been proposed to circumvent this problem such as Shao (1993, p. 488) 'balanced-incomplete cross-validation', Zhang (1993, p. 307) 'r-fold cross-validation', and the 'repeated learning-testing method' (Burman, 1989, 1990; Zhang, 1993).

The conditions required for the consistency of CV_v are given by Shao (1993, pp. 488–489) and are

$$\lim_{n\to\infty} n_c = \infty,$$
$$\lim_{n\to\infty} \frac{n_c}{n} = 0,$$
$$\lim_{n\to\infty} \frac{n_v}{n} = 1. \tag{9}$$

One candidate for n_c is to let n_c be the integer part of n^δ. Letting the resulting degrees of freedom be strictly positive it is possible to obtain the following bounds on δ:

$$\frac{\log(p)}{\log(n)} < \delta < 1, \tag{10}$$

which follows directly from the condition $n_c - p > 0$ where n_c is the integer part of n^δ. Clearly δ satisfying these restrictions satisfies the conditions given in Eq. (9).

It is important to note that the size of the training set n_c must not grow too fast, and the size of the validation set n_v must increase as the sample size increases and cannot be a singleton. We now consider the impact of these results

[4] Zhang (1993, p. 300) refers to this as 'deleting-d multifold cross-validation'.

on a cross-validatory method for dependent data proposed by Burman et al. (1994) known as h-block cross-validation.

4. Cross-validation for dependent data: h-block cross-validation

Let the matrix of n observations on the response and the p predictors be given by $Z = (Y, X)$. We assume that Z denotes a set of jointly dependent stationary observations. Letting z_i denote the ith row (observation) of Z we note that, for stationary observations, the covariance between z_i and z_{i+j} depends only on j and approaches 0 as $|j - i| \to \infty$. Removing h observations either side of z_i will yield a new series which will be nearly independent of z_i as h increases. We remove the ith (vector-valued) observation and h observations on either 'side' of the ith thereby removing $2h + 1$ observations from the sample. When these $2h + 1$ observations are removed from a data set, the resulting data matrix will be denoted by $Z_{(-i:h)} = (Y_{(-i:h)}, X_{(-i:h)})$, while the matrix of removed observations will be denoted by $Z_{(i:h)} = (Y_{(i:h)}, X_{(i:h)})$. If $h = 0$ this will have the effect of removing only the ith observation from the sample.

The h-block cross-validation function is given by

$$\begin{aligned} CV_h &= \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_{i(-i:h)})^2 \\ &= \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i' \hat{\beta}_{(-i:h)})^2. \end{aligned} \tag{11}$$

where $\hat{\beta}_{(-i:h)} = (X'_{(-i:h)} X_{(-i:h)})^{-1} X'_{(-i:h)} Y_{(-i:h)}$ is the least-squares estimator obtained by removing the ith observation and the h observations either side of the ith. Note that when $h = 0$ this simplifies to leave-one-out cross-validation.

Burman et al. (1994) consider both this h-block cross-validation function and a modified h-block cross-validation function. Their modification was motivated by cases in which the ratio of the number of parameters to the sample size (p/n) is not negligible and corrects for underuse of the sample in small-sample settings. Since this paper is concerned with asymptotic behavior the ratio p/n is negligible and their small-sample modification is not addressed in this setting.

4.1. Inconsistency of h-block cross-validation

For h-block cross-validation, the size of the training set is $n_c = n - 2h - 1$. Györfi et al. (1989) require that $h/n \to 0$ as $n \to \infty$, which therefore implies that

$$\lim_{n \to \infty} n_c = \infty \tag{12}$$

and that

$$\lim_{n\to\infty} \frac{n_c}{n} = 1, \tag{13}$$

while Burman et al. (1994) recommend taking h as a fixed fraction of n, that is, $h/n = \gamma$ for some $0 < \gamma < \frac{1}{2}$, which again implies that

$$\lim_{n\to\infty} \frac{n_c}{n} = 1 - 2\gamma > 0 \text{ since } 0 < \gamma < \frac{1}{2}. \tag{14}$$

Note that h-block cross-validation retains a leave-one-out aspect in that the validation set is a singleton. If we denote the validation set again by n_v, then $n_v = 1$ for h-block cross-validation. Therefore, it is observed that for h-block cross-validation,

$$\lim_{n\to\infty} \frac{n_v}{n} = 0. \tag{15}$$

Given that $\lim_{n\to\infty} n_c/n > 0$ and $\lim_{n\to\infty} n_v/n = 0$ the results of Shao (1993) imply that h-block cross-validation is inconsistent for model selection. Also, since the h-block algorithm reduces to leave-one-out cross-validation for *iid* observations when $h = 0$, then in this case $\lim_{n\to\infty} n_c/n = 1$ and $\lim_{n\to\infty} n_v/n = 0$ again implying inconsistency. The results of Shao (1993) therefore imply that h-block cross-validation is inconsistent for linear model selection and will tend to select unnecessarily large models. This result is borne out by simulations summarized in Section 6. In addition, leave-n_v-out cross-validation is inappropriate for dependent data since the training and validation sets will be highly dependent since they are contiguous, and randomization is clearly inappropriate since the temporal ordering of the data contains the stochastic structure of the model.

Note that h-block cross-validatory model selection must be modified in such a way as to satisfy the conditions in Eq. (9) while maintaining near-independence of the training and validation data for dependent observations.

5. Consistent cross-validatory model-selection for dependent data: *hv*-block cross-validation

Fortunately, there is a straightforward modification of the h-block approach which will be consistent in the sense of Shao (1993). Instead of the validation set being a singleton we will train on a set of observations of size n_c and validate on a set of size n_v while maintaining near-independence of the training and validation data via h-blocking. By placing restrictions on the relationship between the training set, validation set, the size of the h-block, and the sample size, we can thereby obtain a consistent cross-validatory model selection procedure for general stationary processes.

We first remove v observations either side of z_i yielding a validation set of size $2v+1$, and then we remove another h observations on either side of this validation set with the remaining $n-2v-2h-1$ observations forming the training set. The value of v controls the size of the validation set with $n_v = 2v+1$, and the value of h controls the dependence of the training set of size $n_c = n - 2h - n_v$ and the validation set of size n_v. We thereby remove the ith (vector-valued) observation and $v+h$ observations on either 'side' of the ith thereby removing $2h+2v+1$ observations from the sample. When these $2h+2v+1$ observations are removed from a data set, the resulting data matrix will be denoted by $Z_{(-i:h,v)} = (Y_{(-i:h,v)}, X_{(-i:h,v)})$, while the matrix of removed observations will be denoted by $Z_{(i:h,v)} = (Y_{(i:h,v)}, X_{(i:h,v)})$ and the matrix of validation observations will be denoted by $Z_{(i:v)} = (Y_{(i:v)}, X_{(i:v)})$.

The resultant cross-validation function will then be given by

$$CV_{hv} = \frac{1}{(n-2v)n_v} \sum_{i=v}^{n-v} \|Y_{(i:v)} - \hat{Y}_{(i:v)(-i:h,v)}\|^2$$

$$= \frac{1}{(n-2v)n_v} \sum_{i=v}^{n-v} \|Y_{(i:v)} - X_{(i:v)}\hat{\beta}_{(-i:h,v)}\|^2, \tag{16}$$

where $\|a\| = \sqrt{a'a}$ and where $\hat{\beta}_{(-i:h,v)} = (X'_{(-i:h,v)}X_{(-i:h,v)})^{-1}X'_{(-i:h,v)}Y_{(-i:h,v)}$.

The parameter h controls the dependence of the validation and training sets and is set to insure near independence of these sets, and these sets need not be completely independent for cross-validation to work (Burman et al., 1994, p. 354). The parameter v controls the relationship between the training set, validation set, and the sample size.

Consider again setting n_c to be the integer part of n^δ. For positive degrees of freedom ($n_c - p > 0$) we obtain a lower bound on δ, $\log(p)/\log(n) < \delta$. If $\delta < 1$ then $n_c/n = n^{\delta-1}$ approaches 0 as n increases. As well, $n_v/n = 1 - n^\delta/n - 2h/n$ approaches 1 if $\delta < 1$ for fixed h and for h less than $o(n)$. Thus,

$$\lim_{n\to\infty} n_c = \infty, \tag{17}$$

and

$$\lim_{n\to\infty} \frac{n_c}{n} = 0, \tag{18}$$

and

$$\lim_{n\to\infty} \frac{n_v}{n} = 1. \tag{19}$$

Therefore, *hv*-block cross-validatory model-selection is consistent if n_c is the integer part of n^{δ} and where $\log(p)/\log(n) < \delta < 1$, $v = (n - n^{\delta} - 2h - 1)/2$, and the conditions of Burman et al. (1994) or Györfi et al. (1989) are met.

An attractive feature of the proposed *hv*-block approach is that it contains many existing forms of cross-validation as special cases. For instance, if $h = v = 0$ this will have the effect of removing only the ith observation from the sample and this simplifies to leave-one-out cross-validation. If $h = 0$ and $v > 0$, this simplifies to leave-n_v-out cross validation in which the number of validation samples is $n - n_v$ rather than the standard ${}_nC_{2v+1}$ possible samples. This is attractive since it results in feasible leave-n_v-out cross-validation for large samples, but note that the data must be pointwise randomized prior to applying this approach for *iid* data.[5] If $h > 0$ and $v = 0$ this simplifies to *h*-block cross-validation, while if $h > 0$ and $v > 0$, this new approach will be dubbed '*hv*-block' cross-validation.

5.1. Optimality of hv-block cross-validation for consistent model selection

It has been noted that, when the data form a dependent stationary sequence, the leave-one-out cross-validated estimate of prediction error may provide a 'very poor estimate' of the true expected prediction error (Burman et al., 1994, p. 352). They demonstrate this for the case of ordinary (leave-one-out) cross-validation ($n_v = 1$). Burman and Nolan (1992, p. 190) demonstrate that ordinary (leave-one-out) cross-validation approximates expected prediction error well if the errors are such that

$$\mathrm{E}(\varepsilon_i \varepsilon_j | X_1, \ldots, X_j) = 0 \quad \text{for } i < j, \tag{20}$$

where 'the training sample is $\{(X_j, Y_j): j \neq i\}$ and the test sample is $\{(X_i, Y_i)\}$'. They note that 'in dependent settings where (20) does not hold, the technique of cross-validation does not work. Instead, it yields a biased estimate of the prediction error' (Burman and Nolan, 1992, p. 190). This bias need not vanish asymptotically giving rise to the potential inconsistency of the ordinary cross-validation estimate of expected prediction error when the data form a dependent sequence.

Bias arises due to the fact that the training and test sets are contiguous, and Burman et al. (1994) solution is to insert a block of length h between the training and test sets. They note that 'blocking allows near independence between these two sets' (Burman et al., 1994). By employing *h*-blocking with appropriate

---

[5] I am grateful to Jun Shao for pointing out that this is very similar to his proposed 'Monte Carlo *CV*' method in Shao (1993).

choice of h, Burman et al. (1994) therefore ensure that

$$\mathrm{E}(\varepsilon_{i-h}\varepsilon_j | X_1, \ldots, X_j) \approx 0 \quad \text{for } i < j, \tag{21}$$

holds, and they demonstrate the advantages of this approach over ordinary cross-validation via simulation. When Eq. (20) holds, h-block cross-validation is unbiased and consistent for the estimation of expected prediction error.

However, consistency of cross-validation for expected prediction error is not sufficient for consistent model-selection as Shao (1993, p. 488) clearly demonstrates for ordinary cross-validation assuming that the data form an independent sequence. This again follows since, just like leave-one-out cross-validation for *iid* data, h-block cross-validation will not be able to discriminate between models which include irrelevant variables but which otherwise are properly specified in that they include all relevant predictors even as $n \to \infty$. This paper's contribution is simply to combine Shao (1993) solution of v-blocking on independent data with Burman et al. (1994) h-blocking on dependent data for improved model-selection.

A rigorous theoretical proof along the lines of Shao (1993) JASA article is not attempted here, and remains the subject of future theoretical work in this area. However, it is briefly noted that conditions required for the validity of Shao (1993) results are indeed met by the proposed hv-block method. In order for Shao (1993) results to follow when h-blocking is used, we must address the notion of 'balance' which is central to his proofs. For independent data, there are ${}_nC_{n_v}$ different subsets of size n_v, but Shao (1993) noted that it is impractical and unnecessary to carry out the validation for all different splits when $n_v > 1$. Shao (1993) therefore proposed 'balanced incomplete cross-validation (BICV)' employing a collection of subsets $\mathscr{B}$ that have size n_v selected according to 'balance' conditions: a) every i, $1 \leqslant i \leqslant n$, appears in the same number of subsets in $\mathscr{B}$; and b) every pair (i, j), $1 \leqslant i < j \leqslant n$, appears in the same number of subsets in $\mathscr{B}$, and his proofs for n_v cross-validation are those for BICV. However, for dependent data we can only use consecutive observations for validation. Fortunately, it is observed that v-block cross-validation using consecutive observations is balanced in the sense of Shao (1993) by letting $n_v = 2v + 1$ and noting that every i, $1 + v \leqslant i \leqslant n - v$, appears in the same number of subsets in $\mathscr{B}$ and b) every pair (i, j), $1 + v \leqslant i < j \leqslant n - v$, appears in the same number of subsets in $\mathscr{B}$. Thus, h-blocking in this context is balanced, and proofs along the lines of those contained in Shao (1993) would be expected to follow, though again a rigorous proof remains the subject of future work in this area.

5.2. Appropriate choice of h

An issue remains as to whether h-blocking is strictly necessary for consistent model-selection for dependent processes. The short answer is that there exist dependent processes for which h-blocking is not necessary, however, for many common dependent processes the incorporation of h-blocking significantly improves the process of model-selection.

Burman and Nolan (1992) demonstrate how, when Eq. (20) holds, '[leave-one-out] cross-validation can carry over, without modification, to the dependent setting'. For example, Burman and Nolan (1992) consider a stationary Markov process with one-step prediction error as one such process, however, they also consider a number of common dependent processes for which h-blocking is strictly necessary. It appears that the rule of thumb given by Burman et al. (1994) works well in practice for the proposed hv-block method. They recommend taking h as a fixed fraction of n ($h/n = \gamma$) and suggest that $\gamma = 0.25$ appears to work well in a wide range of situations.

In the simulations which follow we investigate a range of values for h and conclude that existing rules of thumb appear to provide sensible results.

6. Simulations

Shao (1993) considered a limited simulation experiment with the sample size held fixed at $n = 40$. The purpose of those experiments was to investigate the relative performance of leave-n_v-out cross-validation and leave-one-out cross-validation for *iid* data. He considered two variants of leave-n_v-out cross-validation, Monte Carlo cross-validation (bootstrap resampling and averaging the squared predicting errors), and analytic approximate cross-validation (See Shao, 1993, p. 490 for details). He found a negligible difference between the performance of leave-one-out cross-validation and analytic approximate cross-validation, while Monte Carlo cross-validation performed much better in some situations but worse when the optimal model was the largest model considered. Similarly, Shao (1996) noted the same inconsistency in the context of bootstrap model-selection again for $n = 40$, and again his modified bootstrap model-selection procedure performed much better than the unmodified one in some situations but worse when the optimal model is the largest model considered.

In the following sections we address the issues raised by Shao (1993) from an empirical perspective using a richer set of experiments in a model-selection framework. We first consider the relevance of his findings as $n \to \infty$ for cross-validatory model-selection for *iid* data, then consider their impact on cross-validatory model-selection for dependent data using the h-block method, and then consider the performance of the proposed hv-block approach.

6.1. *Inconsistency of CV_1 for iid data*

In order to validate Shao (1993) results in an *iid* setting it is clearly necessary to examine the performance of leave-one-out cross-validation as $n \to \infty$. Following Shao (1993), linear models were considered with $p = 5$. Five models were compared, two from Category I, two from Category II, and one which was optimal. The data were simulated from

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \tag{22}$$

where the $x_{ij} \sim U[0, 1]$ for $j = 1, \ldots, 3$ and $i = 1, \ldots, n$, $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$ with $\sigma = 0.5$, and $(\beta_1, \beta_2, \beta_3) = (1, 1, 1)$. The sample size was increased to determine whether leave-one-out cross-validation displays any symptoms of the asymptotic inconsistency which would require Shao (1993) modification involving leave-n_v-out cross-validation whereby $n_c/n \to 0$ and $n_v/n \to 1$ as $n \to \infty$.

The five models were
Model 1:

$$y_i = \beta_1 x_{i1} + \varepsilon_i,$$

Model 2:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,$$

Model 3:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \tag{23}$$

Model 4:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i,$$

Model 5:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \varepsilon_i$$

and clearly models 1 and 2 are from Category I, 4 and 5 from Category II, and 3 is optimal.

The simulations were run using the feasible algorithm of Racine (1997), and results from 1,000 simulations are given in Table 1 below.

For a given sample size, the empirical probability of choosing a model from Category I is the sum of the entries from columns 2 and 3 (appearing in column 7 and labelled *pr*(I)), the probability of choosing a model from Category II is the sum of columns 5 and 6 (appearing in column 8 and labelled *pr*(II)), while the probability of choosing the optimal model is the entry in column 4.

The rightmost column in Table 1 contains *pr*($\mathcal{M}_{cv}$ is in Category II but is not $\mathcal{M}_*$) denoted by *pr*(II) in the table. For this example this probability remains constant at around $\frac{1}{5}$ as $n \to \infty$. It is worth noting that results were generated for $n = 100{,}000$ and this probability did fall below $\frac{1}{5}$. Therefore, $pr(\mathcal{M}_{cv_1} = \mathcal{M}_*) \neq 1$ as $n \to \infty$, hence the inconsistency result of Shao (1993) for leave-one-out cross-validation for *iid* data appears to hold up under simulation.

Table 1
Inconsistency of CV_1 for *iid* data. Each row represents the relative frequency of model-selection via leave-one-out cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathscr{M}_*$	Category II		$pr(\mathscr{M}_{cv} \neq \mathscr{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.000	0.000	0.799	0.116	0.085	0.000	0.201
250	0.000	0.000	0.781	0.134	0.085	0.000	0.219
500	0.000	0.000	0.783	0.135	0.082	0.000	0.217
1000	0.000	0.000	0.788	0.139	0.073	0.000	0.212
2500	0.000	0.000	0.798	0.124	0.078	0.000	0.202
5000	0.000	0.000	0.791	0.146	0.063	0.000	0.209

Table 2
Consistency of CV_v for *iid* data, $\delta = 0.5$. Each row represents the relative frequency of model-selection via leave-n_v-out cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathscr{M}_*$	Category II		$pr(\mathscr{M}_{cv} \neq \mathscr{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.000	0.065	0.871	0.056	0.008	0.065	0.064
250	0.000	0.016	0.945	0.038	0.001	0.016	0.039
500	0.000	0.002	0.958	0.037	0.003	0.002	0.040
1000	0.000	0.000	0.971	0.029	0.000	0.000	0.029
2500	0.000	0.000	0.986	0.012	0.002	0.000	0.014
5000	0.000	0.000	0.992	0.008	0.000	0.000	0.008

6.2. Consistency of CV_v for iid data

We now consider the performance of leave-n_v-out cross-validation for *iid* data using the proposed algorithm setting $h = 0$. We use the experimental design outlined Section 6.1 and arbitrarily set $\delta = 0.5$ so that conditions for consistency are met. The simulations were run using a modified version of the approach of Racine (1997), and results from 1000 simulations are given in Table 2 below.

The expectation is for $pr(\mathscr{M}_{cv}$ is in Category II but is not $\mathscr{M}_*)$ in Table 2 to approach 0 as $n \to \infty$, and for this example this clearly occurs. As well, $pr(\mathscr{M}_{cv_1} = \mathscr{M}_*)$ as $n \to \infty$ for leave-n_v-out cross-validation, hence the consistency result of Shao (1993) for *iid* data appears to hold up under simulation provided that n_c and n_v satisfy the conditions from Eq. (9).

Table 3
Inconsistency of CV_h for dependent data, $\gamma = 0.25$. Each row represents the relative frequency of model-selection via h-block cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathcal{M}_*$	Category II		$pr(\mathcal{M}_{cv} \neq \mathcal{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.000	0.001	0.672	0.176	0.151	0.001	0.327
250	0.000	0.000	0.622	0.225	0.153	0.000	0.378
500	0.000	0.000	0.657	0.213	0.130	0.000	0.343
1000	0.000	0.000	0.656	0.211	0.133	0.000	0.344
2500	0.000	0.000	0.660	0.194	0.146	0.000	0.340
5000	0.000	0.000	0.667	0.209	0.124	0.000	0.333

6.3. Inconsistency of CV_h for dependent data

One of the goals of this paper is to determine whether the inconsistency result of Shao (1993) also applies to h-block cross-validation, a cross-validatory method for dependent data. The following example considers the performance of h-block cross-validation as $n \to \infty$, and interest lies in $pr(\mathcal{M}_{cv}$ is in Category II) as $n \to \infty$. If this inconsistency result holds, we would expect that $pr(\mathcal{M}_{cv}$ is in Category II but is not $\mathcal{M}_*) > 0$ as $n \to \infty$.

We again use the models listed in Eq. (23), but introduce dependence via a simple autoregressive error structure. We generate the dependent error process via $\varepsilon_i = \rho\varepsilon_{i-1} + u_i$ with $u_i \sim \mathrm{N}(0, \sigma_u^2)$. Since $\sigma_\varepsilon^2 = \sigma_u^2/(1-\rho^2)$, we adjust σ_u^2 to maintain a constant variance for ε_i equal to that used in Section 6.1, and for this experiment we set $\rho = 0.95$.

The size of the h-block was set using $h = \gamma n$ with $\gamma = 0.25$ in accordance with the suggestion of Burman et al. (1994, p. 356).

The rightmost column in Table 3 contains $pr(\mathcal{M}_{cv}$ is in Category II but is not $\mathcal{M}_*)$ denoted by pr(II) in the table. Note that $pr(\mathcal{M}_{cv}$ is in Category II but is not $\mathcal{M}_*)$ does not go below roughly $\frac{1}{3}$ as $n \to \infty$. As well, $pr(\mathcal{M}_{cv_1} = \mathcal{M}_*) \neq 1$ as $n \to \infty$, hence the inconsistency result of Shao (1993) appears to also hold for h-block cross-validation with dependent data, as outlined in Section 4.1.

6.4. Consistency of CV_{hv} for dependent data

We now consider the performance of the proposed hv-block cross-validation in the same setting as Section 6.3 and arbitrarily set $\gamma = 0.25$ and $\delta = 0.5$ based upon rules-of-thumb found in the literature which satisfy conditions required for consistent model-selection. The simulations were run using a modified version of the algorithm of Racine (1997), and results from 1000 simulations are given in

Table 4
Consistency of CV_{hv} for dependent data, $\gamma = 0.25$, $\delta = 0.5$. Each row represents the relative frequency of model-selection via *hv*-block cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathcal{M}_*$	Category II		$pr(\mathcal{M}_{cv} \neq \mathcal{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.008	0.055	0.707	0.512	0.078	0.063	0.230
250	0.002	0.010	0.739	0.156	0.093	0.012	0.249
500	0.000	0.004	0.829	0.111	0.056	0.004	0.167
1000	0.000	0.000	0.870	0.095	0.035	0.000	0.130
2500	0.000	0.000	0.931	0.057	0.012	0.000	0.069
5000	0.003	0.000	0.971	0.022	0.004	0.003	0.026

Table 5
Consistency of CV_v for dependent data, $\gamma = 0.0$, $\delta = 0.5$. Each row represents the relative frequency of model-selection via *v*-block cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathcal{M}_*$	Category II		$pr(\mathcal{M}_{cv} \neq \mathcal{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.004	0.052	0.753	0.134	0.057	0.056	0.191
250	0.003	0.018	0.745	0.160	0.074	0.021	0.234
500	0.001	0.013	0.789	0.117	0.080	0.014	0.197
1000	0.000	0.005	0.827	0.128	0.040	0.005	0.168
2500	0.000	0.005	0.881	0.091	0.023	0.005	0.114
5000	0.000	0.000	0.874	0.105	0.021	0.000	0.126

Table 4 below. For what follows, we again consider the stationary disturbance process given by $\varepsilon_i = 0.95\varepsilon_{i-1} + u_i$.

If the proposed *hv*-block method is consistent in the sense of Shao (1993), the expectation would be for $pr(\mathcal{M}_{cv}$ is in Category II but is not $\mathcal{M}_*)$ in Table 4 to approach 0 as $n \to \infty$, and for this example this clearly occurs. As well, $pr(\mathcal{M}_{cv_1} = \mathcal{M}_*) \to 1$ as $n \to \infty$, hence the consistency result of Shao (1993) for *hv*-block cross-validation for dependent data appears to hold up under simulation when γ and δ satisfy the conditions required for consistency as outlined in Section 5.

Of direct interest is the performance of *v*-block cross-validation in this setting in order to address benefits from *h*-blocking in addition to *v*-blocking for model-selection purposes for dependent processes. Table 5 presents directly comparable results again using $\delta = 0.5$, that is, $n_v = \sqrt{n}$, but with $h = 0$.

Table 6
Consistency of CV_{hv} for dependent data, $\gamma = 0.01$, $\delta = 0.5$. Each row represents the relative frequency of model-selection via v-block cross-validation for a given sample size, and 1000 simulations were considered for each sample size

	Category I		$\mathscr{M}_*$	Category II		$pr(\mathscr{M}_{cv} \neq \mathscr{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
100	0.006	0.057	0.750	0.133	0.054	0.063	0.187
250	0.001	0.014	0.784	0.145	0.056	0.015	0.201
500	0.000	0.007	0.811	0.121	0.061	0.007	0.182
1000	0.000	0.000	0.841	0.110	0.049	0.000	0.159
2500	0.000	0.000	0.882	0.095	0.023	0.000	0.118
5000	0.000	0.000	0.932	0.158	0.010	0.000	0.068

Tables 4 and 5 clearly indicate that hv-block cross-validation is much more effective than v-block cross-validation for model-selection in dependent settings for large data samples. The probability of selecting the optimal model $\mathscr{M}_*$ (the model in Category II with the smallest dimension) increases by up to 10% for large sample sizes ($n = 5000$) when h-blocking is used. The fact that the hv-block method performs slightly worse than the v-block approach for small samples is likely due to ignoring the modification of Burman et al. (1994) which corrects for underuse of the sample in small-sample settings.

Finally, we consider the effects of increasing h even slightly above zero, that is, of going incrementally from v- to hv-block cross-validatory model-selection with a small of $h > 0$ being used ($\gamma = h/n = 0.01$ is employed). Results for this experiment appear in Table 6.

The results presented in Table 6 indicate that even very small values of h can noticeably improve upon cross-validatory model-selection for dependent processes relative to that arising from v-block cross-validation ($h = 0$). In summary, we observed in Section 6.3 how h-block cross-validation is inconsistent, and we now observe how the proposed hv-block method is consistent and is capable of outperforming v-block methods in dependent settings over a wide range of values of h.

7. Discussion of simulation results

The above simulations validate the findings of inconsistency of Shao (1993) for leave-one-out cross-validation for *iid* data, and confirm that they are also relevant for h-block cross-validation for dependent data. In addition, the simulations validate the consistency of leave-n_v-out cross-validation for *iid* data and of the proposed hv-block method for dependent data.

Some regularities appear in the simulations which are noteworthy. First, for small sample sizes the consistent approaches CV_v and CV_{hv} have a larger probability of the selected model lying in Category I than the inconsistent approaches CV_1 and CV_h for the examples considered above. That is, in small samples the consistent approaches tend to choose underparameterized more frequently than their inconsistent counterparts. As well, it appears that this probability increases with δ and is therefore inversely related to the training data size n_v. However, it is likely that this is simply a reflection of not using the finite-sample corrections for leave-n_v-out cross-validation of Burman (1989) and that for h-block cross-validation found in Burman et al. (1994). As the focus of this paper lies with the asymptotic inconsistency of h-block cross-validation, this correction has not been incorporated here, and it remains an open question whether the corrections will improve the performance of the consistent approaches CV_v and CV_{hv} in very small samples. Second, the inconsistency result for h-block cross-validation manifests itself in moderate sample-size settings and is clearly more than an asymptotic curiosity. Third, the proposed consistent hv-block approach performs much better than it's inconsistent counterpart (h-block) for all sample sizes, while it performs much better than v-block cross-validation for moderate to large sample sizes.

What guidelines can be offered for optimal settings of γ and δ for the proposed approach? For the examples given, $\gamma = 0.25$ and $\delta = 0.5$ appear to be reasonable but clearly will be application-specific. It does appear, however, that even small values of h significantly enhance the process of cross-validatory model-selection in the presence of dependent data. Though no general solution to this problem is offered at this time, Appendix B examines the effects of various settings for γ and δ on the algorithm's performance. A general solution to this problem remains the subject of future work in the area.

8. Application: Autoregression order for G7 exchange rate series

The modest aim of this application is simply to gauge whether the proposed algorithm makes a difference, that is, will there be dramatic differences in the models selected in applied settings or will all approaches tend to select the same model.

It is of interest to compare the performance of the proposed hv-block approach with a traditional model-selection criterion and with variants such as leave-one-out, v-block, and h-block cross-validation. The traditional criterion used here for comparison purposes is the well-known Akaike Information Criterion (AIC) (Akaike, 1974). As well, it would be interesting to see how the different variants of cross-validation compare when in fact the data form a dependent sequence, and Appendix A considers their relative performance when the data form a simulated dependent sequence detailed in Section 6.3

Table 7
Autoregressive model order selected by various criteria

Country	AIC	Leave-one-out ($\gamma = 0$, $\delta = 1$)	v-Block ($\gamma = 0$, $\delta = 0.5$)	h-block ($\gamma = 0.25$, $\delta = 1$)	hv-block ($\gamma = 0.25$, $\delta = 0.5$)
Canada	AR(2)	AR(2)	AR(4)	AR(2)	AR(1)
England	AR(4)	AR(4)	AR(3)	AR(3)	AR(1)
France	AR(4)	AR(4)	AR(1)	AR(2)	AR(1)
Germany	AR(2)	AR(2)	AR(1)	AR(2)	AR(1)
Italy	AR(2)	AR(2)	AR(1)	AR(2)	AR(1)
Japan	AR(4)	AR(4)	AR(1)	AR(3)	AR(1)

again employing the disturbance process given by $\varepsilon_i = 0.95\varepsilon_{i-1} + u_i$. As is seen in Appendix A, these methods do not perform nearly as well as the proposed hv-block approach, tending to select overly parameterized models in large-sample settings as would be expected based on the results of Shao (1993).

We now consider the problem of selecting the order of an autoregression for exchange rate data using cross-validatory methods. Exchange rate modeling has a long history going back decades to early work on efficient markets by Fama (1965) and Cootner (1964), while Meese and Rogoff (1983) work remains the seminal reference for out-of-sample point prediction of nominal exchange rates. Monthly nominal exchange rate data for the G7 countries were obtained from Citibase for time periods 1972:2–1994:10 (all rates were quoted relative to the US$). The range of models from which a model was selected went from an AR(0) (unconditional mean of the series) through an AR(6) model (that is, $y_t = \phi_0 + \sum_{j=1}^{k} \phi_j y_{t-j} + \varepsilon_t$ for $k = 0, \ldots, 6$). The model-selection results for the various criteria are summarized in Table 7 below, while Appendix C presents the partial autocorrelation (PAC) functions for each series for comparison purposes.

The tendency for the nonhv-block approaches to select over-parameterized exchange rate models can be seen by an examination of Table 7. For example, when modeling the England/US series, the AIC and leave-one-out criteria select AR(4) models, the v- and h-block criteria select AR(3) models, while the proposed hv-block approach selects an AR(1) model.

The models selected via hv-block cross-validation given in Table 7 are consistent with model-selection based on the PAC function, as can be seen from the PAC functions presented in Appendix C. Based solely on the PAC function, one would conclude that an AR(1) specification is appropriate for Canada, Germany, Italy, and Japan, while at most an AR(2) specification would be appropriate for England and France. However, cross-validation goes farther than a simple examination of the PAC function by effectively comparing out-of-sample prediction errors across AR(P) models and would therefore be expected to be a preferable model-selection criterion for prediction purposes when dependence is taken into account.

This modest application suggests that the use of the proposed *hv*-block method for model-selection can make a noticeable difference in applied time-series settings and thus can constitute a valuable addition to the tools employed by applied researchers.

9. Conclusion

Cross-validatory model selection for dependent processes has widespread potential application ranging from selection of the order of an autoregression for linear time-series models to selection of the number of terms in nonlinear series approximations for general stationary processes. In this paper we consider the impact of Shao (1993) result of the inconsistency of linear model-selection via cross-validation on *h*-block cross-validation, a cross-validatory method for dependent data. It is shown that model-selection via *h*-block cross-validation is inconsistent and therefore not asymptotically optimal. A modification of the *h*-block method, dubbed '*hv*-block' cross-validation, is proposed which is asymptotically optimal. The proposed approach is consistent in the sense that the probability of selecting the model with the best predictive ability converges to 1 as the total number of observations approaches infinity.

The proposed *hv*-block cross-validatory approach to model selection extends existing results and contains numerous variants of cross-validation including leave-one-out cross-validation, leave-n_v-out cross-validation, and *h*-block cross-validation as special cases.

A number of simulations and applications are conducted to examine the inconsistency of leave-one-out cross-validation and *h*-block cross-validation and to validate consistency of leave-n_v-out cross-validation for *iid* data and the proposed *hv*-block cross-validation for general stationary data. The simulations and applications provide insight into the nature of the various methods and suggest that the proposed method outperforms existing variants of cross-validation in time-series contexts and can be of value in applied settings.

Much work remains to be done before a seamless approach towards model selection can exist, however. Optimal choice of δ and γ for *hv*-block cross-validation is clearly an important issue, and a sound framework regarding optimal settings for δ and γ remains the subject of future work in this area.

Acknowledgements

I would like to thank but not implicate Prabir Burman and Jun Shao for their valuable comments and suggestions. I would also like to thank two anonymous referees for their extremely useful comments and encouragement. Support from

Table 8

	Category I		$\mathscr{M}_*$	Category II		$pr(\mathscr{M}_{cv} \neq \mathscr{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
Leave-one-out ($\gamma = 0.0$, $\delta = 1.0$)							
100	0.000	0.000	0.595	0.209	0.196	0.000	0.405
250	0.000	0.000	0.610	0.193	0.197	0.000	0.390
500	0.000	0.000	0.600	0.218	0.182	0.000	0.400
1000	0.000	0.000	0.593	0.230	0.177	0.000	0.407
2500	0.000	0.000	0.595	0.214	0.191	0.000	0.405
5000	0.000	0.000	0.665	0.183	0.152	0.000	0.335
h-block ($\gamma = 0.25$, $\delta = 1.0$)							
100	0.002	0.019	0.582	0.224	0.173	0.021	0.397
250	0.000	0.000	0.660	0.178	0.162	0.000	0.340
500	0.000	0.000	0.640	0.197	0.163	0.000	0.360
1000	0.000	0.000	0.672	0.183	0.145	0.000	0.328
2500	0.000	0.000	0.668	0.171	0.161	0.000	0.332
5000	0.000	0.000	0.699	0.184	0.147	0.000	0.331
v-block ($\gamma = 0.0$, $\delta = 0.5$)							
100	0.004	0.052	0.753	0.134	0.057	0.056	0.191
250	0.003	0.018	0.745	0.160	0.074	0.021	0.234
500	0.001	0.013	0.789	0.117	0.080	0.014	0.197
1000	0.000	0.005	0.827	0.128	0.040	0.005	0.168
2500	0.000	0.005	0.881	0.091	0.023	0.005	0.114
5000	0.000	0.000	0.874	0.105	0.021	0.000	0.126
hv-block ($\gamma = 0.25$, $\delta = 0.5$)							
100	0.008	0.055	0.707	0.152	0.078	0.063	0.230
250	0.002	0.010	0.739	0.156	0.093	0.012	0.249
500	0.000	0.004	0.829	0.111	0.056	0.004	0.167
1000	0.000	0.000	0.870	0.095	0.035	0.000	0.130
2500	0.000	0.000	0.931	0.057	0.012	0.000	0.069
5000	0.003	0.000	0.971	0.022	0.004	0.003	0.026

the Division of Sponsored Programs at the University of South Florida is gratefully acknowledged. All errors remain, of course, my own.

Appendix A. Relative performance of algorithms for dependent data

Table 8 presents simulation results for leave-one-out, h-block, and v-block variants for the DGP given in Section 6.4.

It can be seen that the proposed hv-block method appears to outperform all variants of cross-validation in terms of its ability to asymptotically select the correct model when the data form a dependent sequence.

Table 9

	Category I		$\mathscr{M}_*$	Category II		$pr(\mathscr{M}_{cv} \neq \mathscr{M}^*)$	
n	Model 1	Model 2	Model 3	Model 4	Model 5	pr(I)	pr(II)
hv-block ($\gamma = 0.1$, $\delta = 0.75$)							
100	0.005	0.031	0.648	0.189	0.127	0.036	0.316
250	0.001	0.011	0.724	0.147	0.117	0.012	0.264
500	0.000	0.000	0.774	0.149	0.077	0.000	0.226
1000	0.000	0.000	0.819	0.133	0.048	0.000	0.181
2500	0.000	0.000	0.903	0.073	0.024	0.000	0.097
5000	0.000	0.000	0.921	0.067	0.012	0.000	0.079
hv-block ($\gamma = 0.1$, $\delta = 0.5$)							
100	0.006	0.046	0.701	0.181	0.066	0.052	0.247
250	0.001	0.013	0.777	0.147	0.062	0.014	0.209
500	0.000	0.007	0.826	0.119	0.048	0.007	0.167
1000	0.000	0.001	0.871	0.100	0.028	0.001	0.128
2500	0.000	0.000	0.949	0.044	0.007	0.000	0.051
5000	0.000	0.000	0.975	0.022	0.003	0.000	0.025
hv-block ($\gamma = 0.25$, $\delta = 0.75$)							
100	0.002	0.040	0.573	0.208	0.177	0.042	0.385
250	0.000	0.006	0.645	0.206	0.143	0.006	0.349
500	0.000	0.000	0.734	0.163	0.103	0.000	0.266
1000	0.000	0.000	0.757	0.163	0.080	0.000	0.243
2500	0.000	0.000	0.839	0.103	0.058	0.000	0.161
5000	0.000	0.000	0.877	0.095	0.028	0.000	0.123
hv-block ($\gamma = 0.25$, $\delta = 0.5$)							
100	0.008	0.055	0.707	0.152	0.078	0.063	0.230
250	0.002	0.010	0.739	0.156	0.093	0.012	0.249
500	0.000	0.004	0.829	0.111	0.056	0.004	0.167
1000	0.000	0.000	0.870	0.095	0.035	0.000	0.130
2500	0.000	0.000	0.931	0.057	0.012	0.000	0.069
5000	0.003	0.000	0.971	0.022	0.004	0.003	0.026

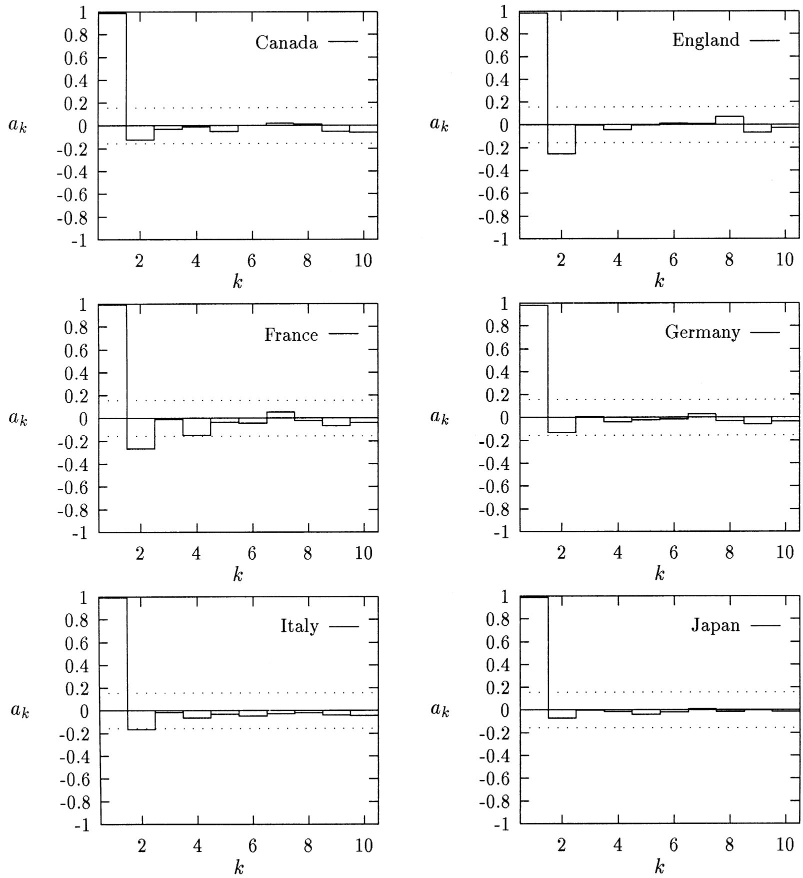
Appendix B. Performance of *hv*-block cross-validation: selection of γ and δ

It is of interest to consider the effect of changing the values of γ and δ upon the performance of the *hv*-block algorithm. Table 9 presents simulation results for the DGP given in Section 6.4.

It can be seen that the default values of $\gamma = 0.25$ and $\delta = 0.5$ suggested by existing literature on *h*-block and *v*-block cross-validation appear to be reasonable for the DGP considered.

Appendix C. Partial autocorrelation functions for G7 exchange rates

The partial autocorrelation (PAC) functions for the nominal exchange rate series modeled in Section 8 are graphed below along with 99% confidence intervals around $a_k = 0$. Based on the PAC function, one would conclude that an AR(1) specification is appropriate for Canada, Germany, Italy, and Japan, while *at most* an AR(2) specification would be appropriate for England and France.



References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions in Automatic Control 19 (6), 716–723.

Allen, D.M., 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 16, 125–127.

Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and Regression Trees. Wadsworth, Belmont, CA.

Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation, and the repeated learning-testing methods. Biometrika 76, 503–514.

Burman, P., 1990. Estimation of optimal transformations using v-fold cross-validation and repeated learning-testing methods. Sankhyā 52, 314–345.

Burman, P., Chow, E., Nolan, D., 1994. A cross-validatory method for dependent data. Biometrika 84, 351–358.

Burman, P., Nolan, D., 1992. Data dependent estimation of prediction functions. Journal of Time Series Analysis 13, 189–207.

Cootner, P., 1964. The Random Character of Stock Market Prices. MIT Press, Cambridge, MA.

Fama, E.F., 1965. The behavior of stock market prices. Journal of Business 38, 34–105.

Geisser, S., 1975. A predictive sample reuse method with application. Journal of the American Statistical Association 70, 320–328.

Györfi, L., Härdle, W., Sarda, P., Vieu, P., 1989. Nonparametric Curve Estimation from Time Series. Springer, Berlin.

Meese, R., Rogoff, K.D., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample?. Journal of International Economics 14, 3–24.

Racine, J., 1997. Feasible cross-validatory model selection for general stationary processes. Journal of Applied Econometrics 12, 169–179.

Shao, J., 1993. Linear model selection by cross-validation. Journal of the American Statistical Association 88 (422), 486–495.

Shao, J., 1996. Bootstrap model selection. Journal of the American Statistical Association 91 (434), 655–665.

Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions (with discussion). Journal of the Royal Statistical Society 36, 111–147.

Stone, M., 1977. Cross-validation: a review. Mathematics of Operations Statistics 9, 127–139.

Whaba, G., Wold, S., 1975. A completely automatic French curve: Fitting spline functions by cross-validation. Communications in Statistics 4, 1–17.

Zhang, P., 1993. Model selection via multifold cross-validation. Annals of Statistics 21, 299–313.