ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

STEP-INDICATOR SATURATION

Jurgen A. Doornik, David F. Hendry and Felix Pretis

Number 658
June 2013

Manor Road Building, Manor Road, Oxford OX1 3UQ

Step-indicator Saturation

Jurgen A. Doornik, David F. Hendry, and Felix Pretis*
Economics Department and Institute for New Economic Thinking
at the Oxford Martin School, Oxford University, UK.

May 31, 2013

Abstract

Using an extension of general-to-specific modelling, based on the recent developments of impulse-indicator saturation (IIS), we consider selecting significant step indicators from a saturating set to capture location shifts. The approximate non-centrality of the test is derived for a variety of shifts using a 'split-half' analysis, the simplest specialization of a multiple-block search algorithm. Monte Carlo simulations confirm the accuracy of the nominal significance levels under the null, and show rejections when location shifts occur, improving in non-null rejection frequency compared to the corresponding IIS-based and to Chow (1960) tests.

JEL classifications: C51, C22.
Keywords: General-to-specific; Step-indicator saturation; Test power; Location shifts; Model Selection; *Autometrics*.

1 Introduction

Unmodeled location shifts can have a pernicious effect on the constancy of models and on forecast performance. In-sample, a shift in the previous unconditional mean induces a systematic bias in forecasts, and estimated models will be mis-specified, potentially distorting parameter estimation and inference. Moreover, an unanticipated location shift at the forecast origin can lead to forecast failure: see Clements and Hendry (1998).

Consequently, we consider a method that can detect and model location shifts, extending the developments behind impulse-indicator saturation (IIS). Hendry, Johansen and Santos (2008) and Johansen and Nielsen (2009) derive the null distribution of IIS, and Castle, Doornik and Hendry (2012) provide simulations under the alternative. Also Hendry and Santos (2010) propose an IIS-based test of super exogeneity, and Hendry and Mizon (2011) and Ericsson and Reisman (2012) present empirical applications of IIS. We call the extension proposed here step-indicator saturation (SIS) (or super saturation when combined with IIS). IIS and SIS are feasible because (e.g.) *Autometrics* can handle more candidate variables N than observations T using a combination of expanding and contracting multiple block searches as described in Doornik (2009a) and Doornik and Hendry (2009). Bergamelli and Urga (2013) undertake extensive simulations of IIS, SIS, and their extensions to trend breaks, as well as comparisons with the

*Financial support from the Open Society Foundations and the Oxford Martin School is gratefully acknowledged. We are indebted to Jennifer L. Castle, Neil R. Ericsson, Søren Johansen, Bent Nielsen, and Giovanni Urga for helpful comments on an earlier draft. The numerical results are based on simulating in *Ox* and *Autometrics*.

sequential break tests proposed in De Peretti and Urga (2005), and using different block partitions find similar results to those reported below. The analytic derivations here use the split-half approach first discussed in Hendry *et al.* (2008).

When the locations, durations, magnitudes and signs of location shifts are unknown, a series of analyses of selection by SIS are required. The size of a test denotes its (false) null rejection frequency, but we are concerned with selection here, and as indicators that are insignificant by the pre-assigned criterion may nevertheless be retained, possibly to offset what would otherwise be a significant mis-specification test, the null retention frequency of indicators is called the gauge. Power is also a well defined concept relating to rejecting a false null hypothesis, so under selection of indicators, we refer to the non-null retention frequency as potency (see Castle, Doornik and Hendry, 2011).

The structure of this paper is as follows. Section 2 describes step-indicator saturation, then section 3 derives its gauge. The extensive set of Monte Carlo experiments is reported following each theory section, so the gauge of SIS is simulated in §3.1. Section 4 investigates the power of a step indicator to detect a known mean shift, and relates that to well-known procedures such as the Chow (1960) test. We next consider selecting indicators for unknown shifts in section 5, first for a shift that can be matched by a single step indicator in §5.1, which develops the basic analytical tools, and is simulated in §5.2. That setting is generalized in §6 to unknown shifts requiring two step indicators (simulated in §6.1); two opposite-signed shifts where one lies in each half (simulated in §6.2), with same-signed shifts in each half in §6.3; then to an unknown break spanning both splits (simulated in §6.4), with a summary of the simulation results in §6.5. Comparisons with IIS are noted in section 7 first for a known break then in §7.1 for an unknown break period, and simulation comparisons with IIS are discussed in §7.2. A generalization to retained regressors is simulated in §8. Section 9 provides an application, and section 10 concludes.

2 Step-indicator saturation (SIS)

Using a modified general-to-specific procedure, Hendry *et al.* (2008) and Johansen and Nielsen (2009) analytically establish the null distribution of the IIS estimator of regression parameters after adding T impulse indicators when the sample size is T. An impulse indicator is defined by $1_{\{t=j\}} = 1$ for observation $t = j$, and zero otherwise. A two-step process is investigated, where half the indicators are added, significant indicators are selected and their times recorded, and then dropped so that the other half can be examined in the same way; finally the significant indicators from the two sets are combined and re-selected. The average retention rate of impulse indicators under the null (gauge) is αT when the significance level of all individual tests is set at α. Moreover, Hendry *et al.* (2008) show that other splits, such as using m splits of size T/m, or unequal splits, do not affect the gauge, or the simulation-based distributions.

Johansen and Nielsen (2009) generalize the analysis to dynamic models (possibly with unit roots) and establish that for small α (e.g., $\alpha \leq 0.01$), the inefficiency of conducting IIS is small despite testing T impulse indicators: intuitively, retained indicators correspond to eliminating individual observations, so αT data points are 'lost'. Their approach applies to regression models with n conditioning variables $\mathbf{z}_t$ which are retained without selection, so when m impulse indicators are selected:

$$y_t = \beta_0 + \boldsymbol{\beta}_1' \mathbf{z}_t + \sum_{i=1}^{m} \phi_i 1_{\{t=t_i\}} + v_t \tag{1}$$

where $v_t \sim \mathsf{IN}\left[0, \sigma_v^2\right]$ and the coefficients of the significant impulses are denoted ϕ_i when the signifi-

cance level α is used in testing their retention. Equation (1) is selected to be congruent.

Consider instead adding a complete set of step indicators $\mathcal{S}_1 = \left\{1_{\{t \leq j\}}, j = 1, \ldots, T\right\}$ to a regression model, where $1_{\{t \leq j\}} = 1$ for observations up to j, and zero otherwise. Step indicators are the cumulation of impulse indicators up to each next observation. As whole-sample vectors, step indicators take the form $\boldsymbol{\iota}_1' = (1,0,0,\ldots,0)$, $\boldsymbol{\iota}_2' = (1,1,0,\ldots,0)$, ..., $\boldsymbol{\iota}_T' = (1,1,1,\ldots,1)$, which is the intercept dummy. Following the analysis in Hendry and Johansen (2013), other regressors could be retained without selection (if not too numerous), so here we just consider the simplest case of a model without regressors.

Thus, we add the saturating set of step indicators $\left\{1_{\{t \leq j\}}, j = 1, \ldots, T\right\}$ to the initial specification:

$$y_t = \sum_{j=1}^{T} \delta_j 1_{\{t \leq j\}} + u_t \tag{2}$$

where $u_t \sim \mathsf{IN}\left[0, \sigma_u^2\right]$. It is infeasible to estimate (2), but the split-half approach to understanding IIS applies to SIS. For T indicators, add the first $T/2$ and select at significance level α, recording which indicators have significant coefficients. Eliminate those, and instead add the second block of $T/2$ to the original model, again selecting at significance level α, and recording which are significant in that subset. Finally, combine the recorded variables (if any) from the two stages, and select again at significance level α. Under the null, setting $\alpha = 1/T$, on average at both sub-steps, $\alpha T/2$ (namely $1/2$ an indicator) will be retained by chance, so on average αT indicators will be retained from the combined stage, so one degree of freedom is lost on average. When there are more regressors plus indicators than T, the procedure can be extended by dividing the total set of N candidate variables into smaller sub-blocks setting $\alpha = 1/N$ overall.

Despite some similarities between the procedures in IIS and SIS, there are important differences necessitating a new analysis. First, while impulse indicators are mutually-orthogonal candidate regressors, step indicators are far from orthogonal, overlapping increasingly as the second index increases. Second, for a location shift that is not at either end, say from T_1 to T_2, two indicators are required to characterize it: $1_{\{t \leq T_2\}} - 1_{\{t < T_1\}}$. Third, the ease of detection must be affected by whether or not T_1 and T_2 lie in the same split. Next, location shifts may occur in both halves. Consequently, instead of just the split-half approach, the combination of expanding and contracting multiple block searches as implemented in *Autometrics* will need to be used in practice (which is why forward-only searches can fail).

2.1 The specification of step indicators

The forms of step indicators are not limited to the above case of $\mathcal{S}_1 = \left\{1_{\{t \leq j\}}, j = 1, \ldots, T\right\}$, but rather can be varied in terms of the shape of the step-indicator matrix. Additionally, instead of using 0/1 indicators, each step-indicator variable could be centred to be mean zero and used in deviations from its mean.

Considering first the shape, alternatively to $\mathcal{S}_1$, the set $\mathcal{S}_2 = \left\{1_{\{t \geq j\}}, j = 1, \ldots, T\right\}$ of the form (0,0,0,...,1), (0,0,0,...,1,1), ..., (1,1,1,...1), could be used; indeed we also consider including both $\mathcal{S}_1$ and $\mathcal{S}_2$ to detect location shifs. Similarly, the set of $\mathcal{S}_3 = \{1_{\{t \leq j\}}, j = T...1\}$ of the form (1,1,1....1), (0,1,1,...1), ... (0,0,...,0,1) could be used. The varying shape has no impact on selection under the null hypothesis of no break. Equally, in the case of a break ocurring, while the regression matrices using $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$ etc. will vary (e.g. forward vs. backward differences), the estimators for a break point occurring in the middle of samples will generally be unbiased and estimated correctly, regardless of the specified shape of $\mathcal{S}$.

The same results extend to using step-indicators in deviations from their mean. Consider $\mathcal{S}_{1,0}$ where the second subscript 0 now indicates mean-zero variables. Then we have $(T-1)$ mean-zero indicators such that $\boldsymbol{\iota}_1' = (1-1/T, -1/T, -1/T, \ldots, -1/T)$, $\boldsymbol{\iota}_2' = (1-2/T, 1-2/T, -2/T, \ldots, -2/T)$, ..., $\boldsymbol{\iota}_{T-1}' = (1-(T-1)/T, 1-(T-1)/T, \ldots, 1-(T-1)/T, -(T-1)/T)$. As before, the shape of these indicators, $\mathcal{S}_{1,0}$, $\mathcal{S}_{2,0}$, $\mathcal{S}_{3,0}$ etc. will determine whether the resulting matrices are equivalent to forward/backward differences, however, the mid-sample break-point estimators will be unbiased. Since here the indicators are mean-zero, the intercept needs to be added as a separate variable to the estimation and selection of each sample-split.

Overall the specification of step indicators has little impact on the estimates and detection of breaks, with the main variation being the resulting forward or backwards difference application. For shifts that fall within a sample, there is no difference in terms of bias or consistency: working with $\mathcal{S}_1$ indicators is equivalent to $\mathcal{S}_{1,0}$ mean-zero indicators. Similarly, a simulation comparison suggests there is no difference in whether mean-zero or original step indicators are used. For simplicity, therefore, the remaining analysis focuses on $\mathcal{S}_1$, 0/1 indicators, while acknowledging that this is only one possible way of implementing SIS.

3 Null rejection frequency of the step-indicator test

To investigate the null rejection frequency of the proposed step indicator test, consider the constant DGP:

$$y_t = \mu + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{3}$$

Under the null hypothesis of no step shift, no indicators are required, so the analysis is invariant to orthogonal transformations and selection over irrelevant variables, one of which reduces SIS back to IIS, to which the analysis in Hendry *et al.* (2008) is applicable. Alternatively, the analysis in Hendry and Johansen (2013) can be used. Either way, as with IIS, add the first half of the step indicators, recording the significant indicators, drop these and add the second half, again recording significant indicators, and finally add the combined set of retained indicators and select the significant indicators from that joint set. Each of the m significant step indicators $\{1_{\{t \leq j\}}\}$ retained in (3) has:

$$\left| \mathsf{t}_{\widehat{\delta}_j} \right| > c_{\alpha_1} \tag{4}$$

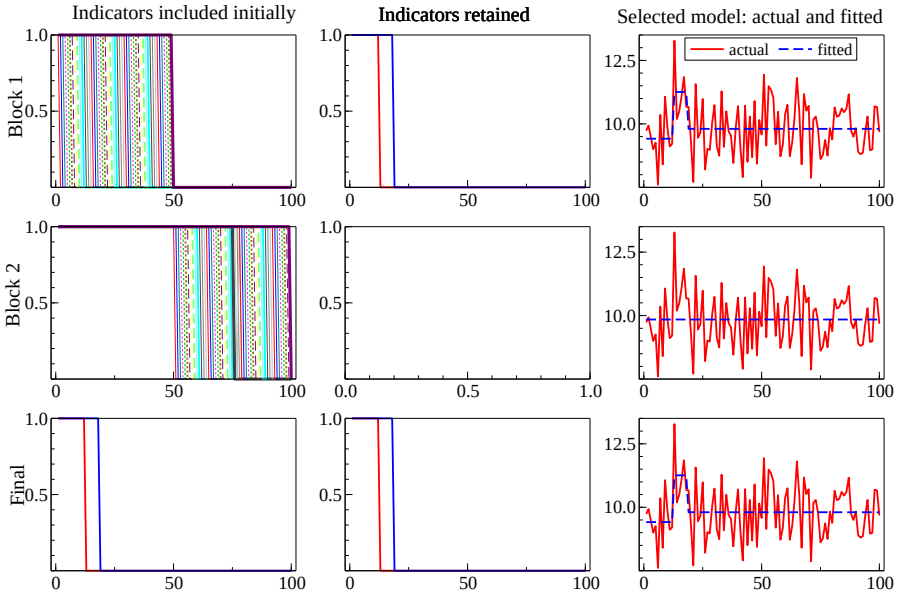
where c_{α_1} is the critical value for significance level α_1, so a subset $\alpha_1 T$ of the indicators will be retained by chance on average. An F-test for the significance of the retained $\left(\delta_{(1)} \ldots \delta_{(m)}\right)$ can be conducted.

Figure 1 illustrates the 'split-half' approach to SIS for (3) when $\mu = 10$ and $\sigma_\epsilon^2 = 1$. The three rows correspond to the three stages: add the first half of the indicators, the second half, then the selected indicators combined. The three columns report the indicators entered, the indicators retained, and the fitted and actual values of the selected model. Fifty indicators are added, and two are retained in row 1. When the second half are entered (row 2), none is retained. Selecting over the two retained indicators again retains them.

Different and multiple splits or unequal divisions entered into (3) should not affect the retention probability under the null: a step shift should only be retained when it is present, here by chance due to a collection of sufficient magnitude, same–signed $\{\epsilon_t\}$ over a sub-sample. In practice, multiple block searches as in *Autometrics* will be needed, as the null may be false.

Non-normal but continuous symmetric distributions $\mathsf{f}(\cdot)$ with at least 8 finite moments entail using the appropriate critical value c_α, where $\alpha = 1 - \int_{-c_\alpha}^{c_\alpha} \mathsf{f}(u)\mathrm{d}u$. If conventional critical values are used for

Figure 1: Illustrating SIS under the null of no shift in (3)



selection, Castle *et al.* (2012) show that (e.g.) IIS will retain indicators corresponding to what are judged 'outliers' relative to the Normal, and we anticipate SIS will do likewise.

3.1 Retention of step indicators under the null hypothesis of no break

We investigated the properties of step indicator saturation using a range of Monte Carlo experiments. All simulations are coded using the *Ox* programming language (see Doornik, 2009b) and replicated $M = 1000$ times at $T = 100$ for $\alpha = 0.01$.

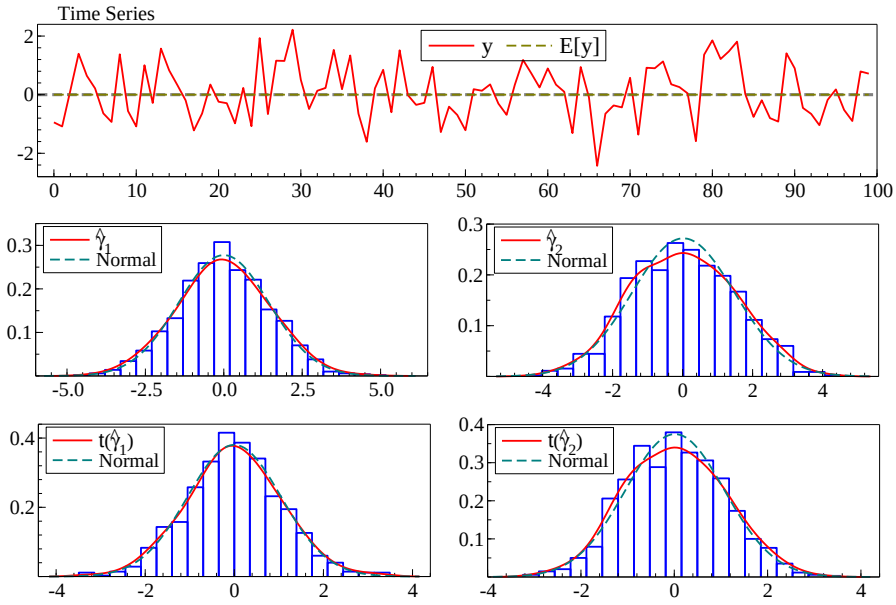
First, the model given in equation (2) is estimated using the split-half approach outlined in section 2. For T=100, initially the first $T/2$ are included and selected at significance level α. Subsequently, these steps are repeated for the second set of $T/2$ indicators. Under the null hypothesis of no break, we expect on average a retention frequency of irrelevant indicators (gauge) equal to α. The DGP is given by equation (3), $y_t = \mu + \epsilon_t$ where $\mu = 0$, $\epsilon \sim \mathsf{IN}[0, \sigma_\epsilon^2]$ and $\sigma_\epsilon^2 = 1$.

For a sample size $T = 100$, Table 1 shows the retention frequency of irrelevant indicators (gauge) both for the first $T/2$ set (D_1) and the second $T/2$ set of indicators (D_2). Gauge values refer to the proportion of incorrectly retained irrelevant indicators for D_1 and D_2 respectively. The retention frequency of irrelevant indicators is closely calibrated to the level of significance α, so on average retain αT irrelevant step indicators under the null distribution that no break occurs. Figure 2 displays the properties of two out of the 100 step indicators ($T_1 = 20, T_2 = 35$). Both estimators $\widehat{\gamma}_{T1}$ and $\widehat{\gamma}_{T2}$ exhibit central t-statistics, and their densities are close to normal distributions centered around zero.

Table 1: Proportion of irrelevant retained indicators under no shift

	Gauge		
α	Overall	D_1	D_2
0.001	0.0018	0.0018	0.0018
0.01	0.013	0.013	0.013
0.05	0.056	0.057	0.054

Figure 2: Step indicator properties under no shift



4 Analytical power of a step-indicator test

We next investigate the power of a step indicator to detect a known mean shift from $\lambda \neq 0$ to $\lambda = 0$ at time $T_1 < T/2$. Consider the DGP in (5):

$$y_t = \lambda 1_{\{t \leq T_1\}} + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{5}$$

where $\lambda \neq 0$. An overall intercept $\mu \neq 0$ can be included to be retained without selection. For simplicity, here the intercept is assumed to have been removed by the Frisch and Waugh (1933) theorem, so implicitly the shift is from $\mu + \lambda$ to μ.

To determine the power of a step-indicator test for the above-given DGP, the simplest nesting model is when the break is known, so that:

$$y_t = \delta_{T_1} 1_{\{t \leq T_1\}} + v_t \tag{6}$$

Then:

$$\widehat{\delta}_{T_1} = \left(\sum_{t=1}^{T} 1_{\{t \leq T_1\}}\right)^{-1} \sum_{t=1}^{T} 1_{\{t \leq T_1\}} y_t = \lambda + T_1^{-1} \sum_{t=1}^{T} 1_{\{t \leq T_1\}} \epsilon_t \tag{7}$$

So, for $T_1/T \to \tau$:

$$\sqrt{T}\left(\widehat{\delta}_{T_1} - \lambda\right) \xrightarrow{D} \mathsf{N}\left[0, \tau \sigma_\epsilon^2\right] \tag{8}$$

Hence, neglecting the estimation uncertainty in $\widehat{\sigma}_\epsilon^2$:

$$\mathsf{t}_{\widehat{\delta}_{T_1}} = \frac{\sqrt{T}\widehat{\delta}_{T_1}}{\sqrt{\tau}\sigma_\epsilon} = \frac{\sqrt{T}(\widehat{\delta}_{T_1} - \lambda)}{\sqrt{\tau}\sigma_\epsilon} + \frac{\sqrt{T}}{\sqrt{\tau}\sigma_\epsilon}\lambda \xrightarrow{D} \mathsf{N}\left[\psi, 1\right] \tag{9}$$

where $\psi = \frac{\sqrt{T_1}}{\tau\sigma_\epsilon}\lambda$ is the non-centrality. However, the denominator approximation requires that any other shifts have been detected and their effects removed. Then (9) is $\sqrt{T_1}$ times the corresponding non-centrality for an individual impulse indicator in Hendry and Santos (2010) (see (34) below). The Chow (1960) predictive failure test is based on IIS (see Salkever, 1976), so SIS could dominate the Chow test, yet not require knowledge of the break point. IIS can already dominate Bai and Perron (1998) as shown in Castle *et al.* (2012).

5 Potency of SIS for unknown breaks

When the locations, durations, magnitudes and signs of location shifts are unknown, a series of analyses of SIS are required. We first consider a shift that can be matched by a single step indicator in §5.1, which develops the basic analytical tools. That setting is generalized in §6 to an unknown break period requiring two step indicators. Subsections 6.1.1 and 6.2.1 consider the occurrence of two shifts where one lies in each half, first when opposite-signed, then when they are roughly equal magnitudes, signs and durations. Subsection 6.3.1 then considers the most difficult to detect case of an unknown break spanning both splits.

5.1 Unknown break period matched by a single step indicator

When the location shift dates are not known, the key issue is how close the probability of retention (potency) comes when searching by SIS to the above probability of rejection for a known step indicator.

We prove that detection of a location shift is using step-indicator saturation is feasible. In the simplest case of a single location shift falling entirely within a half-sample of the data ($T_1 < T/2$) as in (5), add the first half of the step indicators, assuming T is even, so the model becomes:

$$y_t = \sum_{j=1}^{T/2} \gamma_j 1_{\{t \leq j\}} + v_t \tag{10}$$

In matrix notation, the DGP is:

$$\mathbf{y} = \lambda \boldsymbol{\iota}_{T_1} + \boldsymbol{\epsilon} \tag{11}$$

with the model:

$$\mathbf{y} = \mathbf{D}_1 \boldsymbol{\gamma}_{(1)} + \mathbf{v} \tag{12}$$

where $\boldsymbol{\gamma}_{(1)} = (\gamma_1 \dots \gamma_{T/2})'$ and $\mathbf{D}_1 = (\boldsymbol{\iota}_1 \dots \boldsymbol{\iota}_{T/2})$. Then, from (11):

$$\widehat{\boldsymbol{\gamma}}_{(1)} = \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\mathbf{y} = \lambda \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\boldsymbol{\iota}_{T_1} + \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\boldsymbol{\epsilon}$$

where $(\mathbf{D}_1'\mathbf{D}_1)$ is:

$$\begin{aligned}
\begin{pmatrix} \boldsymbol{\iota}_1' \\ \boldsymbol{\iota}_2' \\ \vdots \\ \boldsymbol{\iota}_{T/2-1}' \\ \boldsymbol{\iota}_{T/2}' \end{pmatrix} (\boldsymbol{\iota}_1 \boldsymbol{\iota}_2 \dots \boldsymbol{\iota}_{T/2-1} \boldsymbol{\iota}_{T/2}) &= \begin{pmatrix} \boldsymbol{\iota}_1'\boldsymbol{\iota}_1 & \boldsymbol{\iota}_1'\boldsymbol{\iota}_2 & \dots & \boldsymbol{\iota}_1'\boldsymbol{\iota}_{T/2-1} & \boldsymbol{\iota}_1'\boldsymbol{\iota}_{T/2} \\ \boldsymbol{\iota}_2'\boldsymbol{\iota}_1 & \boldsymbol{\iota}_2'\boldsymbol{\iota}_2 & \dots & \boldsymbol{\iota}_2'\boldsymbol{\iota}_{T/2-1} & \boldsymbol{\iota}_2'\boldsymbol{\iota}_{T/2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\iota}_{T/2-1}'\boldsymbol{\iota}_1 & \boldsymbol{\iota}_{T/2-1}'\boldsymbol{\iota}_2 & \dots & \boldsymbol{\iota}_{T/2-1}'\boldsymbol{\iota}_{T/2-1} & \boldsymbol{\iota}_{T/2-1}'\boldsymbol{\iota}_{T/2} \\ \boldsymbol{\iota}_{T/2}'\boldsymbol{\iota}_1 & \boldsymbol{\iota}_{T/2}'\boldsymbol{\iota}_2 & \dots & \boldsymbol{\iota}_{T/2}'\boldsymbol{\iota}_{T/2-1} & \boldsymbol{\iota}_{T/2}'\boldsymbol{\iota}_{T/2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & \dots & 2 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & \dots & T/2-1 & T/2-1 \\ 1 & 2 & \dots & T/2-1 & T/2 \end{pmatrix}
\end{aligned} \tag{13}$$

The inverse of $(\mathbf{D}_1'\mathbf{D}_1)$ is the classic 'double difference' matrix:

$$\left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \tag{14}$$

so:

$$\left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1' = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \end{pmatrix} \tag{15}$$

which is the forward-difference matrix. Consequently, letting $\nabla\epsilon_t = \epsilon_t - \epsilon_{t+1}$:

$$\begin{aligned}
\widehat{\boldsymbol{\gamma}}_{(1)} &= \lambda \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\boldsymbol{\iota}_{T_1} + \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\boldsymbol{\epsilon} \\
&= \lambda \mathbf{r} + \nabla \boldsymbol{\epsilon}_1
\end{aligned}$$

where $\mathbf{r}$ is a $T/2 \times 1$ selection vector with unity in the T_1th position and zeroes elsewhere, so:

$$\left(\widehat{\gamma}_{(1)} - \lambda \mathbf{r}\right) = \nabla \boldsymbol{\epsilon}_1 \tag{16}$$

where the $(T/2 \times 1)$ vector $\nabla \boldsymbol{\epsilon}_1 = (\nabla \epsilon_1, \nabla \epsilon_2, ..., \nabla \epsilon_{T/2}, \epsilon_{T/2})'$. All the elements of $\widehat{\gamma}_{(1)}$ up to the T_1th should be zero and only the T_1th reflect λ, corresponding to the location shift, with the others being distributed around zero as $\Delta \boldsymbol{\epsilon}_1$. Effectively, (16) shows that only the value of λ at the shift is being picked up, so the incremental information is equivalent to an impulse indicator for T_1:

$$\widehat{\gamma}_{T_1} = \lambda + \nabla \epsilon_{T_1} \tag{17}$$

Also:

$$\mathsf{E}\left[\nabla \boldsymbol{\epsilon}_1 \nabla \boldsymbol{\epsilon}_1'\right] = \sigma_\epsilon^2 \left(\mathbf{D}_1' \mathbf{D}_1\right)^{-1}$$

so:

$$\left(\widehat{\boldsymbol{\gamma}}_{(1)} - \lambda \mathbf{r}\right) \underset{app}{\sim} \mathsf{N}\left[\mathbf{0}, \sigma_\epsilon^2 \left(\mathbf{D}_1' \mathbf{D}_1\right)^{-1}\right] \tag{18}$$

Further:

$$\widehat{\mathbf{y}} = \mathbf{D}_1 \widehat{\boldsymbol{\gamma}}_{(1)} = \lambda \mathbf{D}_1 \mathbf{r} + \mathbf{D}_1 \nabla \boldsymbol{\epsilon}_1 = \lambda \boldsymbol{\iota}_{T_1} + \mathbf{D}_1 \nabla \boldsymbol{\epsilon}_1 \tag{19}$$

as $\mathbf{D}_1 \mathbf{r} = \boldsymbol{\iota}_{T_1}$ and $\mathbf{D}_1 \nabla \boldsymbol{\epsilon}_1 = \boldsymbol{\epsilon}_1$ so:

$$\mathbf{y} - \widehat{\mathbf{y}} = \lambda \boldsymbol{\iota}_{T_1} - \lambda \boldsymbol{\iota}_{T_1} + \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_1 = \boldsymbol{\epsilon}_2 \tag{20}$$

Thus, the estimated error variance adjusted for degrees of freedom:

$$\widehat{\sigma}_\epsilon^2 = \frac{2}{T} \sum_{t=T/2+1}^{T} \left(y_t - \widehat{y}_t\right)^2$$

will be an unbiased estimator of σ_ϵ^2. However, for IID errors:

$$\mathsf{V}\left[\widehat{\gamma}_{T_1}\right] = 2\sigma_\epsilon^2 \tag{21}$$

so that:

$$\mathsf{t}_{\widehat{\gamma}_{T_1}} \approx \frac{\widehat{\gamma}_{T_1}}{\sqrt{2}\sigma_\epsilon} = \frac{\left(\widehat{\gamma}_{T_1} - \lambda\right)}{\sqrt{2}\sigma_\epsilon} + \frac{\lambda}{\sqrt{2}\sigma_\epsilon} \sim \mathsf{N}\left[\frac{\psi_\lambda}{\sqrt{2}}, 1\right] \tag{22}$$

where $\psi_\lambda / \sqrt{2}$ is the non-centrality. In IIS, 1-cut selection was feasible given the orthogonality of the impulse indicators. Here, the intense collinearity between the step indicators entails that there is no information accrual at the level of (22), so sequential selection, commencing by eliminating the least significant, is essential; and it still requires the significance of $\widehat{\gamma}_{T_1}$ in (22). At 1%, $c_\alpha \approx 2.7$, so normalizing on $\sigma_\epsilon = 1$, requires $\lambda > 3.8$ for even a 50% chance of being significant before simplification. However, it is most unlikely that the smallest $\mathsf{t}_{\widehat{\gamma}_j}$ occurs at T_1, and as insignificant irrelevant indicators are deleted from the model, $\mathsf{V}[\widehat{\gamma}_{T_1}]$ will fall rapidly from (21). Because:

$$\mathsf{t}_{\widehat{\gamma}_{j \neq T_1}} \approx \frac{\widehat{\gamma}_{j \neq T_1}}{\sqrt{2}\sigma_\epsilon} \sim \mathsf{N}\left[0, 1\right] \tag{23}$$

on average, $\alpha/2$ of the irrelevant step indicators will be adventitiously significant during selection. If all irrelevant step indicators were eliminated correctly, just $\boldsymbol{\iota}_{T_1}$ remains and the non-centrality would become $\psi = \sqrt{T_1}\lambda/\sigma_\epsilon$ which is $\sqrt{2T_1}$ larger than the non-centrality before selection.

Having completed selection of indicators from the first half, we then add the second half of the step indicators, $\mathbf{D}_2 = (\boldsymbol{\iota}_{T/2+1} \dots \boldsymbol{\iota}_T)$. When the step indicators that were significant in (16) are retained, approximately $\alpha/2$ of the coefficients of $\mathbf{D}_2$ should be significant because the estimated error variance will be close to σ_ϵ^2 from the removal of the location shift. Finally, combine the selected step indicators in a model like (6):

$$y_t = \gamma_{T_1} 1_{\{t \leq T_1\}} + v_t \tag{24}$$

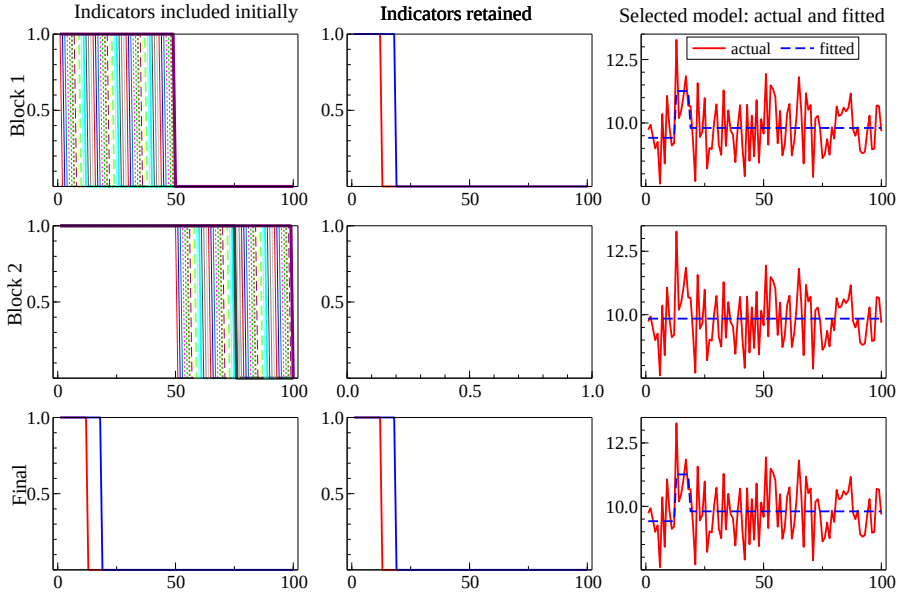
The distribution resulting in the case of a perfect selection must coincide with (8); any irrelevant indicators retained by chance would remove degrees of freedom, and potentially increase variances from collinearity. When the significant indicators from the first half are not carried forward, but a step break occurs in the first half, then the second half may seem to have a break that needs to be modelled, but its retained indicators should vanish on the combined final selection, as occurs with IIS.

Next, Figure 3 illustrates SIS for a location shift over the last 25 observations:

$$y_t = 10 - 10 \times 1_{\{t \geq 76\}} + \varepsilon_t \tag{25}$$

where $\varepsilon_t \sim \mathsf{IN}[0, 1]$. The optimal test for a break is a t-test in (25) at $t = 76$ onwards, but that requires knowledge of the location-shift timing, knowing that it is the only break, and remains the same magnitude break thereafter.

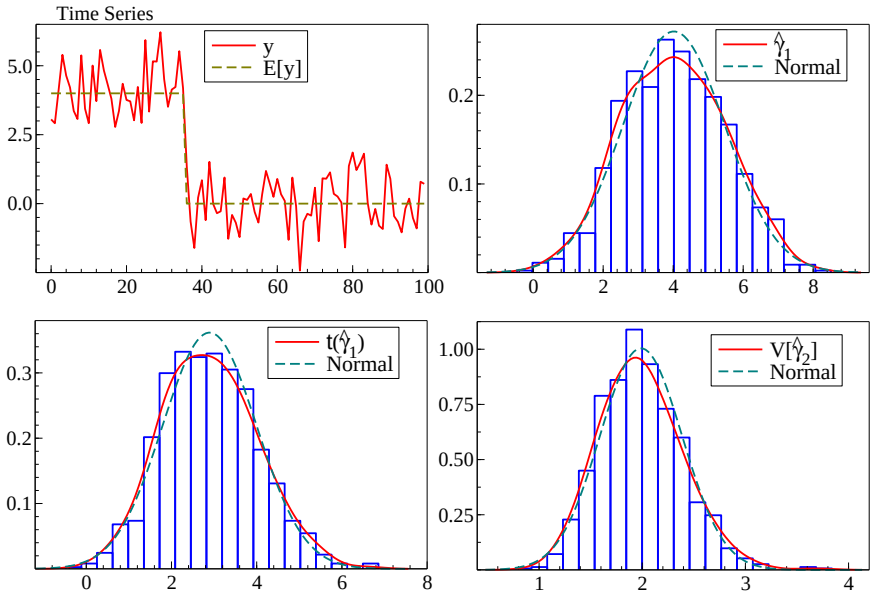
Figure 3: Illustrating SIS for the shift in (25)



Initially the last step indicator captures the mean shift down (row 1), then the location shift is found in row 2, so the now redundant indicator is eliminated in row 3.

As we now discuss, the simulations in §5.2 are consistent with the above analysis: the variance for $\widehat{\gamma}_{T_1}$ from SIS is close to twice that from split-half IIS, but rapidly converges to half that with sequential simplification–compare Figure 4 with Figure 5.

Figure 4: Split half: Single break without sequential selection, $\lambda_1 = 4$SD



5.2 Simulating an unknown break period matched by a single indicator

We now turn to the properties of SIS in the case of a single location shift during the first half of the sample.[1] As before let $T = 100$, for the DGP given in equation (11), and we set $\epsilon \sim \mathsf{IN}[0, \sigma_\epsilon^2]$ where $\sigma_\epsilon^2 = 1$. The timing of the location shift is initially set to $T_1 = 35$: varying break lengths are investigated in Table 3. The break magnitude λ_1 is set equal to 2 and 4 times the standard deviation of the error term ϵ, while selection takes place at 1% significance ($\alpha = 0.01$). The first panel in Figure 4 shows a sample series for $\lambda_1 = 4$SD, where the dotted line graphing $y_t - \epsilon_t$ marks the true underlying step shift.

5.3 Single indicator: Split half without sequential selection

For the split-half approach outlined in section 5.1, Figure 4 shows that, while the density of $\widehat{\gamma}_{T1}$ is centered around the true value of λ_1, in the case of IID error terms the variance of the estimator is twice that of the error term. In line with equation (22), the associated t-statistic is shifted towards zero. The retention frequencies of the step indicator (T_1 Step) in the first $T/2$ for varying levels of λ_1 without sequential selection of indicators are provided in the first section of Table 2. Given the relatively low retention frequency in the simple split-half approach, sequential selection of step indicators is essential.

5.4 Single indicator: Split half with sequential selection

Sequential selection through iterative elimination of the least significant indicators leads to a rapid fall in the variance of the estimator, $V[\widehat{\gamma}_{T1}]$, an increase in the retention frequency of the correct step indicators, and a reduction in the number of incorrectly retained variables. As Figure 5 and the lower section of Table 2 show, sequential selection dramatically improves the outcomes of SIS in the split-half single

[1] In simulation analysis, mean-zero step indicators perform nearly identically to the standard 0/1 step indicators, so we only report the results for the $\mathcal{S}_1$ case.

Table 2: Split half: Gauge and retention frequencies for a single break

		Gauge		**Retention Freq.**
	λ_1	D_1	D_2	T_1 Step
No Sequential:	2 SD	0.007	0.002	0.15
	4 SD	0.006	0.0002	0.59
Sequential:	2 SD	0.017	0.023	0.56
	4 SD	0.003	0.02	0.93

Table 3: Split half: Varying break lengths

		Retention freq. of T_1 for break length l				
	λ_1	$l = 1$	$l = 5$	$l = 10$	$l = 20$	$l = 35$
No sequential:	2SD	0.14	0.12	0.12	0.13	0.15
	4SD	0.61	0.6	0.62	0.58	0.59
Sequential:	2SD	0.18	0.51	0.57	0.55	0.56
	4SD	0.9	0.93	0.93	0.92	0.93

break experiment. In the case of a step shift of 4 standard deviations, sequential selection increases the retention frequency on average from 0.59 to 0.93.

Varying the break length at the start of the sample appears to have little impact on the retention frequencies of the break indicator (see Table 3), except in the case of a single impulse. Using split-half sequential selection, a break of 2SD is retained on average around 50% of the time, with an increase to around 90% for a break of 4SD.

5.5 Single indicator: Effect on D2 indicators

In the case of a single break in the first half of the sample, if the correct indicators from D_1 are not carried forward into D_2 there is a chance that the first step indicators in D_2 are spuriously retained. Based on the simulation evidence, this effect only occurs under sequential selection. The retention frequency of the first i irrelevant indicators in D_2 are given in Table 4 for the cases where the indicators from D_1 are included as well as omitted, where $D_2[i]$ refers to the ith column in the matrix D_2.

There is no notable difference in the retention frequency of the first step indicator in D_2 in a simple split-half approach when the correct indicators from D_1 are included or omitted. However, once sequential selection is applied and the correct indicators are omitted the retention frequency of the first indicator $D_2[1]$ increases from 0.032 to 0.92. Given the importance of sequential selection in increasing the power of step indicators, it remains crucial to carry forward the retained indicators into D_2 since otherwise the first indicator in D_2 is likely to be spuriously retained.

Figure 5: Split-half: Single break with sequential selection, $\lambda_1 = 4$SD

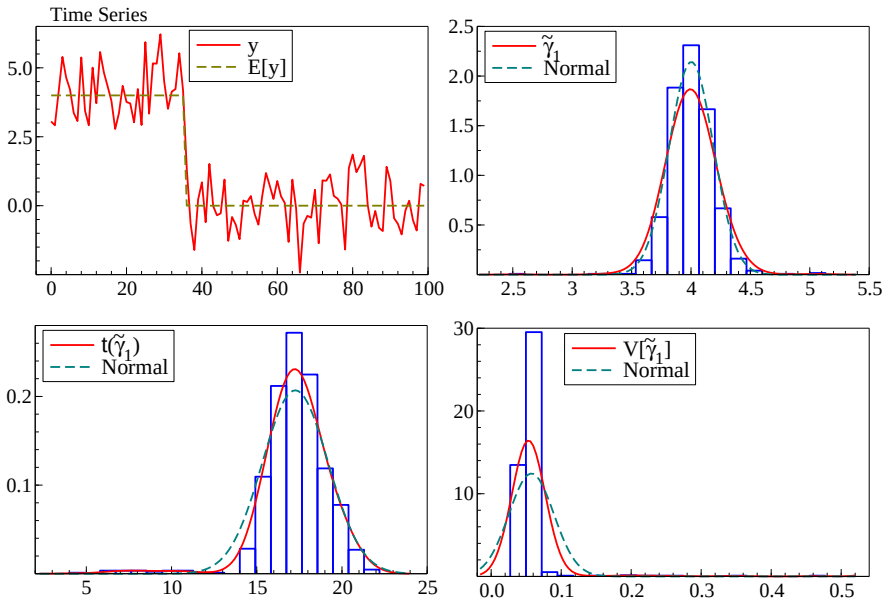


Table 4: Gauge and retention frequencies in D_2 when correct indicators are included/omitted

	Gauge	**Retention freq.**		
		$D_2[1]$	$D_2[2]$	$D_2[3]$
No sequential & Correct step included:	0.015	0.016	0.02	0.013
No sequential & Correct step omitted:	0.002	0.0	0.0	0.0
Sequential & Correct step included:	0.032	0.032	0.031	0.027
Sequential & Correct step omitted:	0.02	0.92	0.068	0.013

Simulations for $\lambda = 4$SD

6 Unknown break period requiring two step indicators in one sample split

Suppose an unknown break that requires two step indicators occurs in one of the sample splits, such as given by the following DGP:

$$y_t = \lambda \left(1_{\{t \leq T_2\}} - 1_{\{t \leq T_1\}}\right) + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{26}$$

where $\lambda \neq 0$, and $T_1 < T_2 < T/2$, as in section 5.1, the shift is entirely within one half of the sample. The model for the first-half split remains (12):

$$\mathbf{y} = \mathbf{D}_1 \boldsymbol{\gamma}_{(1)} + \mathbf{v} \tag{27}$$

where $\boldsymbol{\gamma}_{(1)} = (\gamma_1 \dots \gamma_{T/2})'$ and $\mathbf{D}_1 = (\boldsymbol{\iota}_1 \dots \boldsymbol{\iota}_{T/2})$. For model (27) estimated on data from (26),

$$\begin{aligned} \widetilde{\boldsymbol{\gamma}}_{(1)} &= \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\mathbf{y} \\ &= \lambda \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1' \left(\boldsymbol{\iota}_{T_2} - \boldsymbol{\iota}_{T_1}\right) + \left(\mathbf{D}_1'\mathbf{D}_1\right)^{-1} \mathbf{D}_1'\boldsymbol{\epsilon} \\ &= \lambda \mathbf{s} + \nabla \boldsymbol{\epsilon}_1 \end{aligned} \tag{28}$$

where $\mathbf{s}$ is a $T/2 \times 1$ selection vector with unity in the T_2th position, -1 in the T_1th position and zeroes elsewhere. Thus, a similar analysis holds, with two relevant indicators, so (18) still appears to apply with $\mathbf{r}$ replaced by $\mathbf{s}$. The second split selecting indicators in the latter half of the sample should remain as before in section 5.1.

6.1 Simulating an unknown break period requiring two step indicators

To investigate the case of an unknown break period requiring two step indicators in the first half we set $T_1 = 25$ and $T_2 = 35$. Since sequential selection dominates the standard split-half approach we only consider sequentially selected indicators here. Further, as the above analysis shows it is important to carry forward the selected indicators from D_1 into D_2, we continue under this specification. Table 5 shows the simulation results for sequential selection in the split-half case of two breaks. Retention frequencies in the two break case are close to the case of a single break.

Table 5: Split half: Sequential gauge and retention frequencies for a break with two indicators

	Gauge		**Retention freq.**	
λ_1	D_1	D_2	T_1 Step	T_2 Step
2 SD	0.020	0.011	0.52	0.55
4 SD	0.004	0.017	0.91	0.94

6.1.1 Unknown opposite-signed shifts in each split half

If shifts in each half of the sample have opposite signs, or perhaps very different magnitudes, then both can be detected even in a split-half approach. Consider a DGP of the form:

$$y_t = \lambda_1 \left(1_{\{t \leq T_2\}} - 1_{\{t \leq T_1\}}\right) + \lambda_2 \left(1_{\{t \leq T_4\}} - 1_{\{t \leq T_3\}}\right) + \epsilon_t \tag{29}$$

where $\epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right]$ as before, and $T_1 < T_2 \leq T/2$ whereas $T/2 \leq T_3 < T_4$ with $\lambda_1 \lambda_2 < 0$. Since (e.g.):

$$\frac{1}{T}\sum_{t=1}^{T}\left(\lambda_2\left(1_{\{t\leq T_4\}} - 1_{\{t\leq T_3\}}\right)\right) = \lambda_2 \frac{(T_4 - T_3)}{T} = \phi_2$$

(say) an intercept retained without selection is needed to remove the mean effect of the other location shift.

The formula in (28) still applies, with appropriate adjustments for estimating the intercepts, but even if the first shift is correctly modelled, the equation in (20) for the residuals becomes:

$$\mathbf{y} - \widehat{\mathbf{y}} = \lambda_2\left(\boldsymbol{\iota}_{T_4} - \boldsymbol{\iota}_{T_3}\right) - \phi_2\boldsymbol{\iota} + \boldsymbol{\epsilon}_2 = \widehat{\mathbf{v}}_2 \tag{30}$$

which has a larger estimated error variance than in the previous cases because:

$$\begin{aligned}\frac{2}{T}\mathsf{E}\left[\widehat{\mathbf{v}}_2'\widehat{\mathbf{v}}_2\right] &= \frac{2}{T}\mathsf{E}\left[\boldsymbol{\epsilon}_2'\boldsymbol{\epsilon}_2\right] + \frac{2}{T}\left(\lambda_2\left(\boldsymbol{\iota}_{T_4} - \boldsymbol{\iota}_{T_3}\right) - \phi_2\boldsymbol{\iota}\right)'\left(\lambda_2\left(\boldsymbol{\iota}_{T_4} - \boldsymbol{\iota}_{T_3}\right) - \phi_2\boldsymbol{\iota}\right) \\ &= \sigma_\epsilon^2 + \lambda_2^2\frac{2\left(T_4 - T_3\right)}{T} - 2\lambda_2\phi_2\frac{2\left(T_4 - T_3\right)}{T} + 2\phi_2^2 \\ &= \sigma_\epsilon^2 + 2\lambda_2^2\frac{(T_4 - T_3)}{T}\left(1 - \frac{(T_4 - T_3)}{T}\right)\end{aligned} \tag{31}$$

To compensate for the equivalent effect to (31) when searching for a second break, step indicators found in the first half should be be included in the second half selection.

6.2 Simulating unknown opposite-signed shifts in each split half

To consider unknown opposite-signed shifts in each half λ_1 and λ_2 are chosen such that $\lambda_1\lambda_2 < 0$. For simulation we set $\lambda_1 = -\lambda_2$ with the break timing given by: $T_1 = 25$ to $T_2 = 35$ and $T_3 = 75$ to $T_4 = 85$. Table 6 shows that even with shifts falling in the middle of each half, SIS can be successful in identifying the breaking points. Sequential selection, as is to be expected, dominates the simple split-half approach.

Table 6: Split half: Opposite-signed shifts in each half

		Gauge		**Retention frequency**			
	$\lambda_{1,2}$	D_1	D_2	T_1 Step	T_2 Step	T_3 Step	T_4 Step
No Sequential:	2 SD	0.016	0.016	0.13	0.17	0.13	0.13
	4 SD	0.016	0.016	0.61	0.61	0.61	0.6
Sequential:	2 SD	0.021	0.045	0.52	0.55	0.57	0.56
	4 SD	0.004	0.03	0.91	0.94	0.93	0.93

6.2.1 Unknown equal breaks in each split half

Shifts with relatively equal magnitudes, durations, and the same signs in each half, so they are roughly evenly distributed between the two halves, could well appear as just a larger error variance, rendering the simplest split-half approach ineffective. Nevertheless, when T is sufficiently large, both shifts can be

detected using a modified split-half approach. First saturate the second half by impulse indicators, then the first half can be tackled by a split-half approach, so quarters are examined, without any additional cost under the null. That procedure is then reversed for the first half. This is a variant of super saturation, where IIS is also undertaken, but limiting IIS to the alternate half and not using the information it reveals about outliers and shifts.

Under the alternative, by eliminating the shift in the second half, the first half comes under the above analysis for a single shift, which is then detectable provided it is not evenly split between the quarters. In practice, *Autometrics* uses multiple block searches, and for IIS, this has proved effective in detecting multiple breaks. Blocks would need to span most of the length of a location shift to detect it using SIS, but that may be less essential for super saturation.

6.3 Simulating unknown equal breaks in each half

Unknown breaks of equal magnitude are assessed by setting $\lambda_1 = \lambda_2$. The shift timing is set to: $T_1 = 25$ to $T_2 = 35$ and $T_3 = 75$ to $T_4 = 85$, there are two step shifts of a length of 10 periods in each half. We first consider a simple split-half approach with sequential selection, followed by the same experiment run under multiple splits using *Autometrics*. Due to the nature of the multiple-split division only a single gauge value is reported for both D_1 and D_2 together. Table 7 provides summary results. There appears to be no major difference in retention frequencies, both split-half and a multiple-split approach exhibit similar results.

Table 7: Split half and multiple split: Unknown equal breaks in each half

	Gauge		**Retention Frequency**			
Split half: $\lambda_{1,2}$	D_1	D_2	T_1 Step	T_2 Step	T_3 Step	T_4 Step
(sequential) 2 SD	0.021	0.044	0.52	0.55	0.59	0.60
4 SD	0.005	0.03	0.91	0.94	0.94	0.94
Multiple split: $\lambda_{1,2}$	D_1 & D_2		T_1 Step	T_2 Step	T_3 Step	T_4 Step
2 SD	0.038		0.53	0.48	0.55	0.55
4 SD	0.018		0.87	0.91	0.94	0.92

6.3.1 Unknown break period spanning both splits

The analysis in §6.2.1 should be even more effective in capturing a location shift spanning the initial halves, as then the break will almost always lie entirely within a quarter of the sample. This follows since within the first half of $T/2$ where the break lies towards the end by necessity of spanning into the second half, if it were longer than $T/4$, SIS would find the shorter as if it were the break, and similarly for the second half.

6.4 Simulating an unknown break period spanning both splits

To simulate a break period spanning both splits, the break timing is set such that the end of the first break occurs just as the second break starts, i.e. $T_2 = T/2$ and $T_3 = T/2 + 1$. Additionally the present simulation sets $T_1 = 35$ and $T_4 = 65$ leading to a single step shift of a length of 30 periods spanning both halves. Table 8 presents the results when using split-half with sequential selection as well as multiple

splits. As before, there is little difference between split-half and multiple splits, both exhibit retention frequencies around 0.9 for a step shift of 4 standard deviations.

Table 8: Split half and multiple split: Break spanning both splits

	Gauge		**Retention Frequency**			
Split half: λ_1	D_1	D_2	T_1 Step	T_2 Step	T_3 Step	T_4 Step
(sequential) 2 SD	0.011	0.039	0.58	0.001	0.0	0.56
4 SD	0.002	0.02	0.94	0.0	0.0	0.93
Multiple splits: λ_1	D_1 & D_2		T_1 Step	T_2 Step	T_3 Step	T_4 Step
2 SD	0.029		0.57	0.01	0.01	0.55
4 SD	0.019		0.94	0.02	0.02	0.96

6.5 Summary of the simulation results

The Monte Carlo experiments provide evidence for the feasibility of detecting location shifts using SIS. In the case of static DGPs with specific location shifts, the step indicators exhibit high retention frequencies – around 50% in the case of a shift equal to 2 standard deviations and around 90% for shifts of 4 standard deviations. These results appear constant across single shifts, multiple shifts, spanning both halves and even including additional regressors and impulses. Sequential selection of step shifts is crucial to ensure high power, which in turn requires the selected indicators to be carried forward into the second half of indicators when using split half. Multiple splits using *Autometrics* without the condition of carrying relevant indicators forward yield equal results to the split-half method.

7 Comparison with IIS

A group of impulses of the same magnitude is equivalent to a step shift and thus IIS can in theory be used to detect step changes. However, SIS exhibits a higher power in detection of these shifts as we demonstrate in this section. The corresponding IIS-based test would initially specify:

$$y_t = \sum_{j=1}^{T} \gamma_j 1_{\{t=j\}} + v_t \tag{32}$$

The potency to retain a single step indicator in (32) depends on the probability of rejecting the null for the associated estimated γ_j:

$$\widehat{\gamma}_j = \lambda + \epsilon_j \tag{33}$$

The properties of such tests on impulse indicators are discussed in Hendry and Santos (2005). Let ψ_λ denote the non-centrality. Because $\mathsf{V}[\widehat{\gamma}_{,j}] = \sigma_\epsilon^2$, then:

$$\mathsf{E}\left[\mathsf{t}_{\gamma_j=0}\left(\psi_\lambda\right)\right] = \mathsf{E}\left[\frac{\widehat{\gamma}_j}{\widehat{\sigma}_\epsilon}\right] \approx \frac{\lambda}{\sigma_\epsilon} = \psi_\lambda \tag{34}$$

When ϵ_t is independent normal, the potency could be computed directly from the t-distribution. More generally, since most outliers will have been removed, normality should be a reasonable approximation.

However, the denominator approximation requires that almost all other shifts and outliers have been detected. An F-test of all $\boldsymbol{\gamma}' = (\gamma_1 \dots \gamma_{T_1}) = \mathbf{0}$ would have T_1 degrees of freedom in the numerator, compared to one for the step indicator test. The potency depends on the signal-noise ratio and the length of break, T_1, which determines how many indicators need to be located, and the break magnitude. Under orthogonality of indicators, the F-test can be related to the individual $\mathsf{t}_{\gamma_j=0}$ tests using the approximation:

$$\mathsf{F}^{T_1}_{T-T_1}(\boldsymbol{\gamma} = \mathbf{0}) \approx \frac{1}{T_1} \sum_{j=1}^{T_1} \mathsf{t}^2_{\gamma_j=0} \tag{35}$$

so:

$$\mathsf{E}\left[\mathsf{F}^{T_1}_{T-T_1}(\boldsymbol{\gamma} = \mathbf{0})\right] \approx \frac{1}{T_1} \mathsf{E}\left[\sum_{j=1}^{T_1} \mathsf{t}^2_{\gamma_j=0}\right] = \psi^2_\lambda \tag{36}$$

as against $T_1\psi^2_\lambda$ for the known step indicator F-test. Thus, this approach is far better for a location shift within a block. As IIS can handle other break types (blips and impulses, and even trending shifts), empirical modelling may need super saturation rather than SIS alone, although the example in §9 shows how well SIS can do in capturing outliers.

7.1 Unknown break period

When the break period is not known, the number of impulse indicators retained will vary between samples. In the example below, the test is an F-test in an IID DGP, so the particular relevant and irrelevant impulse indicators retained does not matter, merely their total numbers. As in Hendry and Santos (2005), for a single location shift, the potency therefore depends on:

(a) the length of the break, T_1,

(b) the number of relevant retained elements in the index, which on average will be $\mathsf{p}_\lambda T_1$, where p_λ is the probability of retaining any given relevant impulse when testing at α_1, and

(c) the number of irrelevant impulse indicators retained in the model, which on average will be $T\alpha_1$.

Letting $r = (\mathsf{p}_\lambda \tau + \alpha_1)$, then on average the F-test will have $k = Tr$ numerator degrees-of-freedom and (allowing for n regressors) $T(1-r) - n$ denominator. Under orthogonality, the F-test can be related to the individual $\mathsf{t}_{\gamma_j=0}$ tests again using the approximation:

$$\mathsf{F}^{k}_{(T-k-n)} \approx \frac{1}{k} \sum_{j=1}^{k} \mathsf{t}^2_{\gamma_j=0} \tag{37}$$

so now for the indicators alone, since $\mathsf{E}[\mathsf{t}^2_{\gamma_j=0}] = 1$ for the α_1 irrelevant impulses retained and ψ^2_λ for the $\mathsf{p}_\lambda\tau$ relevant:

$$\mathsf{E}\left[\mathsf{F}^{k}_{(T-k-n)}\right] \approx \frac{1}{k} \mathsf{E}\left[\sum_{j=1}^{k} \mathsf{t}^2_{\gamma_j=0}\right] = \frac{\mathsf{p}_\lambda \tau \psi^2_\lambda + \alpha_1}{(\mathsf{p}_\lambda \tau + \alpha_1)} = 1 + \frac{\mathsf{p}_\lambda \tau \left(\psi^2_\lambda - 1\right)}{(\mathsf{p}_\lambda \tau + \alpha_1)} \tag{38}$$

noting that p_λ depends on α_1. Smaller α_1 will lower p_λ (in a complicated tradeoff), but anyway need $\alpha_1 \leq 1/T$, so main loss is probably stringent significance level rather than spurious breaks.

7.2 Simulation comparisons with IIS

Here we first compare the success of identifying step shifts via SIS against using IIS. Second, we assess the situation when IIS and SIS are combined and selection takes place over both impulses and step indicators. To evaluate the success of identifying step shifts two measures are used. As before, the retention frequency of the step indicator and impulse indicator at the specified break timing is measured. Further, we use proportional potency as a measure which is defined as the average percentage of the step shift covered by the step and impulse indicators respectively.

IIS vs SIS

In presence of a step shift: The DGP here is identical to the simulations above, we consider a single step shift falling in the first half of the sample, with $T_1 = 25$ and $T_2 = 35$. *Autometrics* is used in the simulations to select over multiple splits. Tables 9 and 10 summarize the simulation outcomes showing the much higher retention frequencies and proportional potency of using step indicators over impulse indicators in the presence of a small step shift.

Table 9: Multiple split: IIS vs SIS gauge and potency

	IIS Gauge	**SIS Gauge**	**IIS Potency**	**SIS Potency**
λ_1				
2 SD	0.007	0.032	0.25	0.98
4 SD	0.01	0.022	0.91	0.99

Table 10: Multiple split: IIS vs SIS retention frequency

	Retention freq.			
λ_1	IIS T_1 Step	SIS T_1 Step	IIS T_2 Step	SIS T_2 Step
2 SD	0.25	0.51	0.22	0.56
4 SD	0.90	0.93	0.90	0.93

In presence of an impulse: To consider the retention frequency of a single impulse, the DGP is set such that there is a single period break of magnitude λ_1 at time $T_1 = 25$. In terms of step shift variables this is equivalent to a shift upwards at T_1 and an immediate shift downwards at $T_2 = T_1 + 1$. Table 11 shows the simulation outcomes. For a single impulse in the middle of the sample IIS exhibits higher power than SIS.

Table 11: Multiple split: IIS vs SIS single impulse

	IIS Gauge	**SIS Gauge**	**Retention Freq.**		
λ_1			IIS T_1 Impulse	SIS T_1 Step	SIS T_2 Step
2 SD	0.013	0.02	0.3	0.17	0.19
4 SD	0.014	0.02	0.91	0.75	0.75

IIS and SIS

In presence of a step shift: Super saturation (see Ericsson and Reisman, 2012) combines IIS with SIS and allows for selection over both. IIS in theory can recover step shifts by selecting a group of impulses and vice versa a linear combination of step indicators in SIS can identify a single impulse. Here we consider the case of a single step shift as before ($T_1 = 25$ and $T_2 = 35$). Tables 12 and 13 show the simulation outcomes.

Comparing the results to using SIS only as in the previous section, it appears that the presence of impulses does not have a detrimental effect on the high power of the step indicators. For a step shift of 4 standard deviations using SIS alone (see table 10) yields average retention frequencies of 0.95 (for T_1) and 0.96 (for T_2), while combined with IIS (see table 13) the average retention frequencies are 0.94 (for T_1) and 0.93 (for T_2).

Table 12: Super saturation: Step shift gauge and potency

	IIS Gauge	**SIS Gauge**	**IIS Potency**	**SIS Potency**
λ_1				
2 SD	0.008	0.025	0.014	0.97
4 SD	0.007	0.013	0.012	0.99

Table 13: Super saturation: Step shift retention frequency

	Retention Freq.			
	IIS T_1 Step	SIS T_1 Step	IIS T_2 Step	SIS T_2 Step
λ_1				
2 SD	0.014	0.51	0.014	0.53
4 SD	0.021	0.91	0.018	0.91

In presence of a single impulse: In the case of a single impulse $T_1 = 25$, IIS combined with SIS leads to a reduction in power of the impulse indicators in detecting the shift (for $\lambda_1 = 4$SD the change is from 0.91 to 0.55). This result stems from the fact that an impulse indicator would not be selected when two step indicators already identified the single shift. The joint rejection frequency, a measure of both IIS and SIS missing the impulse, accounts for this effect. An impulse of 4 standard deviations is missed by SIS and IIS jointly on average 15% of the time.

Table 14: Super–Saturation: Single impulse gauge, potency and retention frequency

	IIS Gauge	**SIS Gauge**	**Retention Freq.**			**Joint Rejection**
λ_1			IIS T_1 Imp.	SIS T_1 Step	SIS T_2 Step	Frequency
2 SD	0.009	0.018	0.16	0.1	0.11	0.74
4 SD	0.009	0.018	0.55	0.3	0.31	0.15

8 Generalization to retained regressors

In line with the theoretical work of Johansen and Nielsen (2009) on IIS, we assess by means of simulations the use of SIS with $n < T$ general regressors by including the $T \times n$ matrix $\mathbf{Z}$ as independent variables. For a single step shift with unknown timing requiring two indicators the DGP is then given by:

$$y_t = \boldsymbol{\beta}_1' \mathbf{z}_t + \lambda \left(1_{\{t \leq T_2\}} - 1_{\{t \leq T_1\}}\right) + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{IN}\left[0, \sigma_\epsilon^2\right] \tag{39}$$

For the present simulation we set $\sigma_\epsilon^2 = 1$ and $n = 10$. For each of the $i = 1, ..., n$ regressors the associated non-centralities are set to $\mathsf{E}[t_i] = \psi_i = 4$. The individual z_i are orthogonal in expectation and not selected over, thus present in every selection iteration of the step indicators. The break timing is set as before to $T_1 = 25$ and $T_2 = 35$. Table 15 displays the simulation outcomes and properties of the step indicators. With the inclusion of 10 relevant independent variables the densities of the two break estimators are centered around the true value of $\lambda_1 = 4SD$. The high power of SIS seems unaffected by the presence of additional regressors – the retention frequencies are close to values in experiments without regressors.

Table 15: SIS with General Regressors

	Gauge	**Retention Frequency**	
λ_1		T_1 Step	T_2 Step
2 SD	0.035	0.5	0.62
4 SD	0.024	0.91	0.94

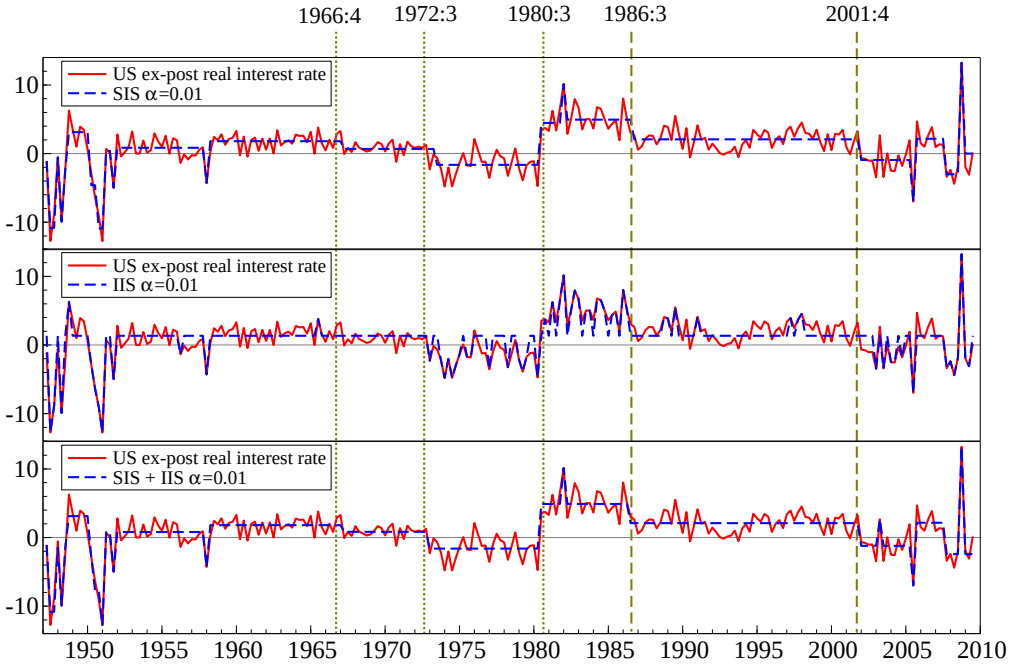
9 Application to U.S. *ex-post* real interest rates

We consider an application of SIS on U.S. *ex-post* real interest rates dating from the second quarter of 1947 (1947:2) until the third quarter of 2009 (2009:3) covering a total of $T = 250$ observations. To detect the presence of shifts, Bai and Perron (2003) initially investigated structural breaks in a shorter subset (1961:1 to 1986:3) of the same series. Using the Bai and Perron (1998) test, they find evidence of three main breaks in the form of mean-shifts in 1966:4, 1972:3 and 1980:3. Castle *et al.* (2012), using IIS, confirm these breaks in the short sample and demonstrate the success of IIS in the extended sample (1947:2 – 2009:3) – in which the original Bai and Perron test fails to detect the above mentioned structural changes in the mean[2]. However, as our above analysis outlines, IIS exhibits lower power than SIS in detecting step shifts and requires an additional procedure to identify step changes. In order for a step shift to be detected using IIS there are two options, either a group of adjacent impulses needs to be determined and tested for joint significance, or a step variable constructed for the duration of chosen impulses – both add additional uncertainty in detecting shifts. To assess the performance of SIS, we apply our methodology on the same time series as used by Castle *et al.* (2012), comparing outcomes of using SIS against IIS, and seeing whether the originally detected shifts by Bai and Perron can be found in the longer sample.

Using indicators of the form $\mathcal{S}_1$ we apply SIS, IIS, and SIS+IIS together at a significance level of $\alpha = 0.01$. The estimated models include 250 potential independent variables for SIS and IIS, and

[2]In Castle *et al.* (2012), the Bai and Perron (1998) test is applied for a maximum of 8 breaks at significance of 1%, 2.5% and 5%, with a trimming factor of 0.10 which implies a minimum step-shift or break length of 25 obervations.

Figure 6: U.S. ex-post real interest rate 1947:2–2009:3 with structural breaks determined by SIS and IIS



500 independent variables for SIS+IIS combined. Table 16 and Figure 6 provide detailed results of the estimation, with the top panels showing SIS alone, the middle panels depicting IIS, and the bottom panels combining both SIS and IIS. The breaks found initially by Bai and Perron are indicated with dotted vertical lines, the additional breaks discovered in the longer sample using SIS are marked with dashed vertical lines.

Overall, SIS analysis (top panel in Figure 6) suggests five major structural changes in the mean of U.S. *ex-post* real interest rates around 1967:1, 1973:2, 1980:2, 1986:3, and 2001:4, with periods of high volatility at both the start and end of the sample. Without having to specify a minimum break length or a maximum number of breaks, SIS captures breaks close to the ones which the Bai and Perron (1998) test finds in the short sample but is unable to detect in the large sample. Using SIS, additionally two further step shifts are detected, one at 1986:3 coinciding with the findings of Castle *et al.* (2012), and a shift around 2001:4 just before a period of high volatility (dashed vertical lines). SIS is able to capture variation and changes at both the start and end of the sample in the form of outliers similar to IIS. Single outlying observations are detected using two consecutive step indicators with opposite signs, this is most notable in the periods of high volatility at the beginning and end of the sample in the top panel of Figure 6.

IIS analysis alone (middle panel in Figure 6) has the potential to detect step shifts through adjacent impulses from 1972:3–1986:3, which, however, have to be untangled by selecting groups of outlying observations.

Combining SIS and IIS (bottom panel) leads to little change of results relative to just using SIS. The combination of SIS and IIS results in 500 independent variables to model the sample of 250 observations, but even with this large number of potential regressors only 21 are retained, providing evidence that overfitting is of little concern when using SIS.

Table 16: SIS and IIS Results for U.S. ex-post real interest rate, 1947:2–2009:3

SIS at 1%	24 step indicators retained
Step Indicators:	1947:2, 1947:4, 1948:1, 1948:2, 1948:3, 1950:1, 1950:3, 1951:1, 1951:3, 1951:4, 1957:4, 1958:1, 1967:1, 1973:2, 1980:2, 1981:4, 1982:1, 1986:3, 2001:4, 2005:2, 2005:3, 2007:3, 2008:3, 2008:4
IIS at 1%	63 impulse indicators retained
Impulse Indicators:	1947:4, 1948:2, 1948:4, 1950:2, 1950:3, 1950:4, 1951:1, 1951:4, 1956:2, 1958:1, 1965:3, 1973:1, 1973:4, 1974:1, 1974:2, 1974:3, 1974:4, 1975:1, 1975:3, 1975:4, 1977:1, 1978:2, 1979:1, 1979:2, 1979:3, 1980:2, 1980:4, 1981:2, 1981:4, 1982:1, 1982:3, 1982:4, 1983:1, 1983:3, 1983:4, 1984:2, 1984:3, 1984:4, 1985:2, 1985:3, 1986:1, 1986:2, 1988:4, 1989:1, 1989:3, 1991:1, 1997:2, 1997:4, 1998:1, 2003:1, 2003:3, 2004:1, 2004:2, 2004:4, 2005:3, 2007:4, 2008:1, 2008:2, 2008:3, 2008:4, 2009:1, 2009:2
SIS and IIS at 1%	7 Impulse and 15 step indicators retained
Step indicators:	1947:2, 1947:4, 1948:3, 1950:1, 1951:1, 1957:4, 1958:1, 1967:1, 1973:1, 1980:2, 1986:2, 2001:4, 2005:2, 2005:3, 2007:3
Impulse Indicators:	1948:2, 1950:2, 1951:1, 1951:4, 1982:1, 2003:2, 2008:4

10 Conclusion

A test for location shifts by step-indicator saturation seems feasible. The test has the correct null retention frequency in constant conditional models for a nominal test size of α. The approximate alternative retention-frequency function was derived analytically for simple models, and helps explain the simulation outcomes. Although only one and two shifts were considered in detail, the general nature of the test makes it applicable when there are multiple breaks. While all the derivations and Monte Carlo experiments here have been for simple static equations and specific location shifts, the principles seem general, and should apply to dynamic equations (although with approximate null-rejection frequencies) and to conditional systems. Generalizations to non-stationary settings would need to extend Johansen and Nielsen (2009).

Important applications for SIS are to check for location shifts at the forecast origin, where they would be pernicious, and to testing super exogeneity, where the IIS-based test has relatively low power for long breaks.

References

Bai, J., and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**, 47–78.

Bai, J., and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, **18**, 1–22.

Bergamelli, M., and Urga, G. (2013). Detecting multiple structural breaks: A Monte Carlo Study and an Application to the Fisher Equation for US. Discussion paper, Cass Business School, London.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, **3 (1)**, DOI: 10.2202/1941–1928.1097.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2012). Model selection when there are multiple breaks. *Journal of Econometrics*, **169**, 239–246.

Castle, J. L., and Shephard, N. (eds.)(2009). *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, **28**, 591–605.

Clements, M. P., and Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

De Peretti, C., and Urga, G. (2005). Stopping tests in the sequential estimation of multiple structural breaks. Discussion paper, Cass Business School, London.

Doornik, J. A. (2009a). Autometrics. in Castle, and Shephard (2009), pp. 88–121.

Doornik, J. A. (2009b). *Object-Oriented Matrix Programming using Ox* 7th edn. London: Timberlake Consultants Press.

Doornik, J. A., and Hendry, D. F. (2009). *Empirical Econometric Modelling using PcGive: Volume I*. London: Timberlake Consultants Press.

Ericsson, N. R., and Reisman, E. L. (2012). Evaluating a global vector autoregression for forecasting. *International Advances in Economic Research*, **18**, 247–258.

Frisch, R., and Waugh, F. V. (1933). Partial time regression as compared with individual trends. *Econometrica*, **1**, 221–223.

Hendry, D. F., and Johansen, S. (2013). Model discovery and Trygve Haavelmo's legacy. *Econometric Theory*, forthcoming.

Hendry, D. F., Johansen, S., and Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, **33**, 317–335. Erratum, 337–339.

Hendry, D. F., and Mizon, G. E. (2011). Econometric modelling of time series with outlying observations. *Journal of Time Series Econometrics*, **3 (1)**, DOI: 10.2202/1941–1928.1100.

Hendry, D. F., and Santos, C. (2005). Regression models with data-based indicator variables. *Oxford Bulletin of Economics and Statistics*, **67**, 571–595.

Hendry, D. F., and Santos, C. (2010). An automatic test of super exogeneity. In Watson, M. W., Bollerslev, T., and Russell, J. (eds.), *Volatility and Time Series Econometrics*, pp. 164–193. Oxford: Oxford University Press.

Johansen, S., and Nielsen, B. (2009). An analysis of the indicator saturation estimator as a robust regression estimator. in Castle, and Shephard (2009), pp. 1–36.

Salkever, D. S. (1976). The use of dummy variables to compute predictions, prediction errors and confidence intervals. *Journal of Econometrics*, **4**, 393–397.