Inherently High-dimensional Analysis with Indicator Saturation*

Neil R. Ericsson†

Abstract
Indicator saturation is a generic approach to robust estimation, and to the detection and quantification of structural breaks. Saturation techniques are inherently high-dimensional and require automated model selection with non-standard inference. This paper characterizes several roles for saturation techniques and proposes extensions of *impulse* indicator saturation that have greater power to detect empirically common structural breaks. A model for the Brazilian inflation rate illustrates saturation techniques.

Key Words: Autometrics, Brazil, high-dimensional analysis, impulse indicator saturation, inflation, model selection, robust estimation, structural breaks.

1. Introduction

Impulse indicator saturation is a generic approach to robust estimation; and it provides a coherent framework for detecting and quantifying crises, jumps, and changes in regime. Such structural breaks can be persistent, time-dependent, and difficult to detect, yet have substantive implications for policy analysis. Impulse indicator saturation helps address these issues, albeit being inherently high-dimensional and requiring automated model selection with non-standard inference. This paper characterizes several roles for impulse indicator saturation and proposes extensions with greater power to detect empirically common structural breaks. These saturation techniques can demonstrate the robustness of a model to a wide range of feasible alternatives. They can also yield statistical and economic improvements to a model and thereby provide insights into the practical justification of empirical evidence. Additionally, saturation techniques provide a framework for creating near-realtime early-warning and rapid-detection devices, such as of financial market anomalies.

This paper is organized as follows. Section 2 describes the computer-automated model selection algorithm in Autometrics. That algorithm is central to implementing impulse indicator saturation. Section 3 discusses impulse impulse indicator saturation and proposes several extensions that are potentially useful for detecting structural breaks such as crises, jumps, and changes in regime. Section 4 illustrates impulse indicator saturation with a model for the Brazilian inflation rate from Dées, di Mauro, Pesaran, and Smith (2007), using the methods detailed in Sections 2 and 3 and drawing on results in Ericsson and Reisman (2012). Section 5 concludes.

*The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. The author is grateful to David Hendry and Erica Reisman for helpful discussions and comments. All numerical results were obtained using PcGive Version 14.1, Autometrics Version 1.5g, and Ox Professional Version 7.10 in 64-bit OxMetrics Version 7.01: see Doornik and Hendry (2013) and Doornik(2009).

†Division of International Finance, Board of Governors of the Federal Reserve System, Washington, DC 20551 USA (ericsson@frb.gov), and Research Program on Forecasting, Department of Economics, The George Washington University, Washington, DC 20052 USA (ericsson@gwu.edu)

2. Computer-automated Model Selection

In economics, data mining has traditionally been viewed as a pejorative activity—a necessary evil, with nominal critical values affected; see Lovell (1983), Denton (1985), and the extensive literature on pre-testing. By contrast, data mining can be constructive rather than pejorative; see Campos and Ericsson (1999). This section thus summarizes an approach to constructive data mining as implemented in the model selection algorithm Autometrics. Section 3 discusses indicator saturation techniques, which utilize that algorithm in Autometrics. Together, that algorithm and indicator saturation techniques can aid in detecting crises, jumps, and changes in regime, as Section 4 shows.

Hoover and Perez (1999) proposed an automated general-to-specific model-selection algorithm that incorporated many of the features of the "Hendry" or LSE methodology. Hendry and Krolzig (2001) developed a second-generation algorithm called PcGets, which extended and improved upon Hoover and Perez's algorithm; see also Hendry and Krolzig (1999, 2003, 2005) and Krolzig and Hendry (2001). Doornik and Hendry (2009) implement a third-generation algorithm called Autometrics, which is part of PcGive version 13; see also Hendry and Doornik (2014). Autometrics utilizes one-step and multi-step simplifications along multiple paths following a tree search method. Diagnostic tests serve as additional checks on the simplified models, and encompassing tests resolve multiple terminal models. Both analytical and Monte Carlo evidence show that the resulting model selection is relatively non-distortionary for Type I errors. At an intuitive level, Autometrics functions as a series of sieves that aim to retain parsimonious congruent models while discarding both noncongruent models and over-parameterized congruent models. This feature of the algorithm is eminently sensible, noting that the data generation process itself is congruent and is as parsimonious as feasible. At a more formal level, White (1990, 2009) provides a statistical theoretic basis for this model-building strategy in which diagnostic test statistics (including encompassing statistics) are treated as explicit design criteria. See also Hansen, Lunde, and Nason (2011) for a related discussion on model confidence sets.

The current section summarizes Autometrics as an automated model-selection algorithm, thereby providing the necessary background for interpreting subsequent empirical applications. For ease of reference, the algorithm is divided into three "stages", denoted Stage 0, Stage 1, and Stage 2. For full details of Autometrics's algorithm, see Doornik (2008, 2009) and Doornik and Hendry (2009). Hendry and Krolzig (2003) and Campos, Ericsson, and Hendry (2005) describe the relationship of the general-to-specific approach to other modeling approaches in the literature, and Hoover and Perez (2004) extend the general-to-specific approach to cross-section regressions.

Stage 0: the general model, impulse indicator saturation, and F pre-search tests. Stage 0 involves three parts: the estimation and evaluation of the general model, inclusion of impulse indicator dummies for all observations, and some pre-search tests aimed at simplifying the general model before instigating formal multi-path searches.

First, the general model is estimated, and diagnostic statistics are calculated for it. If any of those diagnostic statistics is unsatisfactory, the modeler must decide what to do next—whether to "go back to the drawing board" and develop another general model, or to continue with the simplification procedure, perhaps ignoring the offending diagnostic statistic or statistics.

Second, and optionally, Autometrics performs block additions and searches of impulse indicator dummies for all observations in a process known as impulse indicator saturation. Doing so generates a robust regression estimator, and it provides a check for parameter constancy. Section 3 discusses impulse indicator saturation further.

Third, and also optionally, Autometrics attempts to drop various sets of potentially insignificant variables. Autometrics does so by dropping all variables at a given lag, starting with the longest lag. Autometrics also does so by ordering the variables by the magnitude of their t-ratios and either dropping a group of individually insignificant variables or (alternatively) retaining only a group of individually statistically significant variables. In effect, an F pre-search test for a group of variables is a single test for multiple simplification paths, a characteristic that helps control the costs of search. If these tests result in a statistically satisfactory reduction of the general model, then that new model is the starting point for Stage 1. Otherwise, the general model itself is the starting point for Stage 1.

Stage 1: a multi-path encompassing search. Stage 1 tries to simplify the model from Stage 0 by searching along multiple paths, ensuring that the diagnostic tests are not rejected. If all variables are individually statistically significant, then the initial model in Stage 1 is the final model. If some variables are statistically insignificant, then Autometrics tries deleting those variables to obtain a simpler model. If a simplification is rejected, Autometrics backtracks along that simplification path to the most recent previous acceptable model and then tries a different simplification path. A terminal model results if the model's diagnostic statistics are satisfactory and if no remaining regressors can be deleted.

If Autometrics obtains only one terminal model, then that model is the final model. However, because Autometrics pursues multiple simplification paths in Stage 1, Autometrics may obtain multiple terminal models. To resolve such a situation, Autometrics creates a union model from those terminal models and tests each terminal model against that union model. Autometrics then creates a new union model, which nests all of the surviving terminal models; and that union model is passed on to Stage 2.

Stage 2: another multi-path encompassing search. Stage 2 in effect repeats Stage 1 (possibly iteratively) by applying the simplification procedures from Stage 1 to the union model obtained at the end of Stage 1. The resulting model is the final model. If Stage 2 obtains more than one terminal model after applying encompassing tests, then the final model is selected by using the Akaike, Schwarz, and Hannan–Quinn information criteria. See Akaike (1973, 1981), Schwarz (1978), and Hannan and Quinn (1979) for the design of these information criteria, and Atkinson (1981) for the relationships between them.

In addition to optional pre-search simplification and impulse indicator saturation, the algorithm requires several other choices: in particular, target size and fixity of regressors.

Target size. Autometrics requires the modeler to choose which tests are calculated and to specify the critical values for those tests. In principle, the modeler can choose the test statistics and their critical values directly, although doing so is tedious because of the number of statistics involved. To simplify matters, Autometrics offers several options for the "target size", which incorporates pre-designated selections of test statistics and critical values.

The "target size" is meant to equal "the proportion of irrelevant variables that survives the [simplification] process" (Doornik, 2009, p. 100). In empirical analyses, Autometrics's target size is often either 5% or 1%—which are values that appear to approximate the liberal and conservative strategies in PcGets. The liberal strategy errs on the side of keeping some variables, even although they may not actually matter. The conservative strategy keeps only variables that are clearly significant statistically, erring in the direction of excluding some variables, even although those variables may matter. Which strategy is preferable depends in part on the data themselves, in part on the class of regressors examined (and, in particular, whether indicator saturation is considered), and in part on the objectives of the modeling exercise. Also, the two approaches *may* generate similar or identical results. In the empirical analysis below, the target size is 0.1% (very stringent)

Fixity of regressors. Autometrics offers the option of designating certain variables as "fixed". Fixed variables are forced to always be included in regression, whereas free variables (variables that are not fixed) may be deleted by the algorithm.

In short, Autometrics is general-to-specific, multi-path, iterative, and encompassing, with diagnostic tests providing additional assessments of statistical adequacy, and with options for pre-search simplification and robustification. Autometrics can be characterized as having two components:

1. Estimation and diagnostic testing of the general unrestricted model (Stage 0); and
2. Selection of the final model by
    (a) pre-search simplification of the general unrestricted model (Stage 0),
    (b) impulse indicator saturation (Stage 0), and
    (c) multi-path (and possibly iterative) selection of the final model (Stages 1 and 2).

3. Indicator Saturation Techniques

Impulse indicator saturation (IIS) is a general procedure for model evaluation, and in particular for testing parameter constancy. Section 4 applies IIS to analyze potential structural breaks in a model of Brazilian inflation. IIS is beneficial more generally; and it has been employed for model evaluation, model design, and robust estimation. Moreover, Ericsson (2015) re-interprets existing tests of forecast bias as special cases of IIS and shows how IIS can be used to detect arbitrarily time-varying forecast bias.

This section summarizes how impulse indicator saturation provides a general procedure for analyzing a model's constancy. Specifically, IIS is a generic test for an unknown number of breaks, occurring at unknown times anywhere in the sample, with unknown duration, magnitude, and functional form. IIS is a powerful empirical tool for both evaluating and improving existing empirical models. Hendry (1999) proposes IIS as a procedure for testing parameter constancy. See Hendry, Johansen, and Santos (2008), Doornik (2009), Johansen and Nielsen (2009, 2013), Hendry and Santos (2010), Ericsson (2011a, 2011b, 2012, 2016), Ericsson and Reisman (2012), Bergamelli and Urga (2014), Hendry and Pretis (2013), Hendry and Doornik (2014), Castle, Doornik, Hendry, and Pretis (2015), and Marczak and Proietti (2016) for further discussion and recent developments.

Impulse indicator saturation uses the zero-one impulse indicator dummies $\{I_{it}\}$ to analyze properties of a model. For a sample of T observations, there are T such dummies, so unrestricted inclusion of all T dummies in an estimated model (thereby "saturating" the sample) is infeasible. However, blocks of dummies *can* be included, and that insight provides the basis for IIS. To motivate how IIS is implemented in practice, this subsection employs a bare-bones version of IIS in two simple Monte Carlo examples.

Example 1. This example illustrates the behavior of IIS when the model is correctly specified. Suppose that the data generation process (DGP) for the variable w_t is:

$$w_t = \mu_0 + \varepsilon_t \quad \varepsilon_t \sim NID(0, \sigma^2), \qquad t = 1, \ldots, T, \tag{1}$$

where w_t is normally and independently distributed with mean μ_0 and variance σ^2. Furthermore, suppose that the model estimated is a regression of w_t on an intercept, i.e., the model is correctly specified. Figure 1a plots Monte Carlo data from the DGP in equation (1) with $\mu_0 = 20$, $\sigma^2 = 1$, and $T = 100$. Figure 1b plots the estimated model's residuals, scaled by that model's residual standard error.

The bare-bones version of IIS is as follows.

1. Estimate the model, including impulse indicator dummies for the first half of the sample, as represented by Figure 2a. That estimation is equivalent to estimating the model over the second half of the sample, ignoring the first half. Drop all statistically insignificant impulse indicator dummies and retain the statistically significant dummies (Figure 2b).
2. Repeat this process, but start by including impulse indicator dummies for the *second* half of the sample (Figure 2d), and retain the significant ones (Figure 2e).
3. Re-estimate the original model, including all dummies retained in the two block searches (Figure 2g), and select the statistically significant dummies from that combined set (Figure 2h).

Hendry, Johansen, and Santos (2008) and Johansen and Nielsen (2009) have shown that, under the null hypothesis of correct specification, the expected number of impulse indicator dummies retained is roughly αT, where α is the target size. In Figure 2h, five dummies are retained; $\alpha = 5\%$; and $\alpha T = (5\% \cdot 100) = 5$, an exact match.

Example 2. This example illustrates the behavior of IIS when there is an unmodeled break. Suppose that the DGP for the variable w_t is:

$$w_t = \mu_0 + \mu_1 S_{64t} + \varepsilon_t \quad \varepsilon_t \sim NID(0, \sigma^2), \qquad t = 1, \ldots, T, \tag{2}$$

where S_{64t} is a one-off step dummy that equals 0 ($t = 1, \ldots, 63$) or 1 ($t = 64, \ldots, 100$), and μ_1 is its coefficient in the DGP. The model estimated is a regression of w_t on an intercept alone, ignoring the break induced by the step dummy S_{64t}. As in Example 1, w_t is normally and independently distributed with a nonzero mean. However, that mean alters at $t = 64$. The model ignores that change in mean (aka a "location shift") and hence is mis-specified. Figure 3a plots Monte Carlo data from the DGP in equation (2) with $\mu_0 = 20$, $\mu_1 = -10$, $\sigma^2 = 1$, and $T = 100$. Figure 3b plots the estimated model's residuals. Interestingly, no residuals lie outside the estimated 95% confidence region, even though the break is -10σ. The model has no "outliers".

Figure 4 plots the corresponding graphs for the bare-bones implementation of IIS described in Example 1, as applied to the Monte Carlo data in Example 2. As

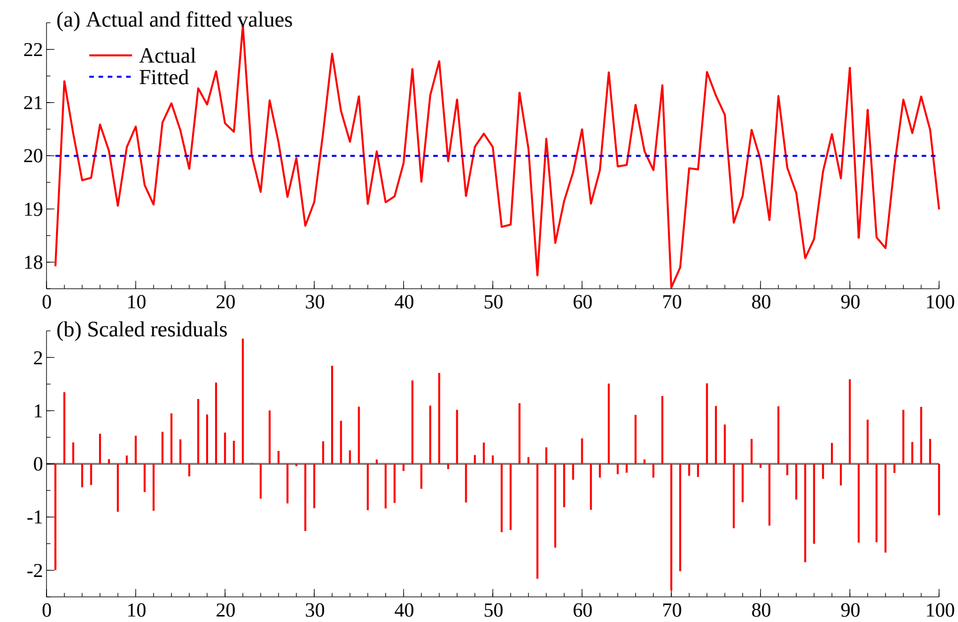

Figure 1: Actual and fitted values and the corresponding scaled residuals for the estimated model when the DGP does not have a break.

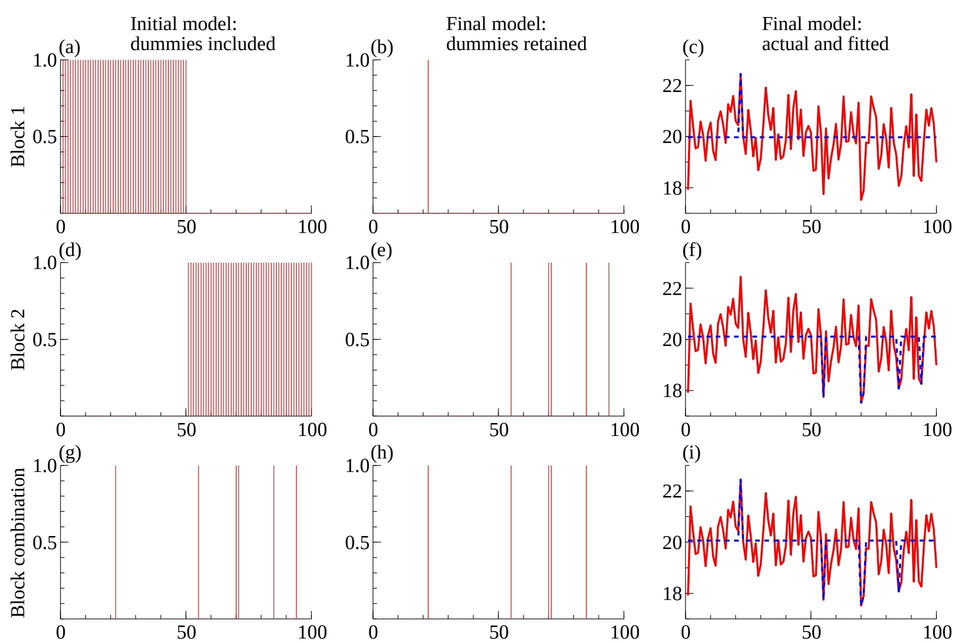

Figure 2: A characterization of bare-bones impulse indicator saturation with a target size of 5% when the DGP does not have a break.

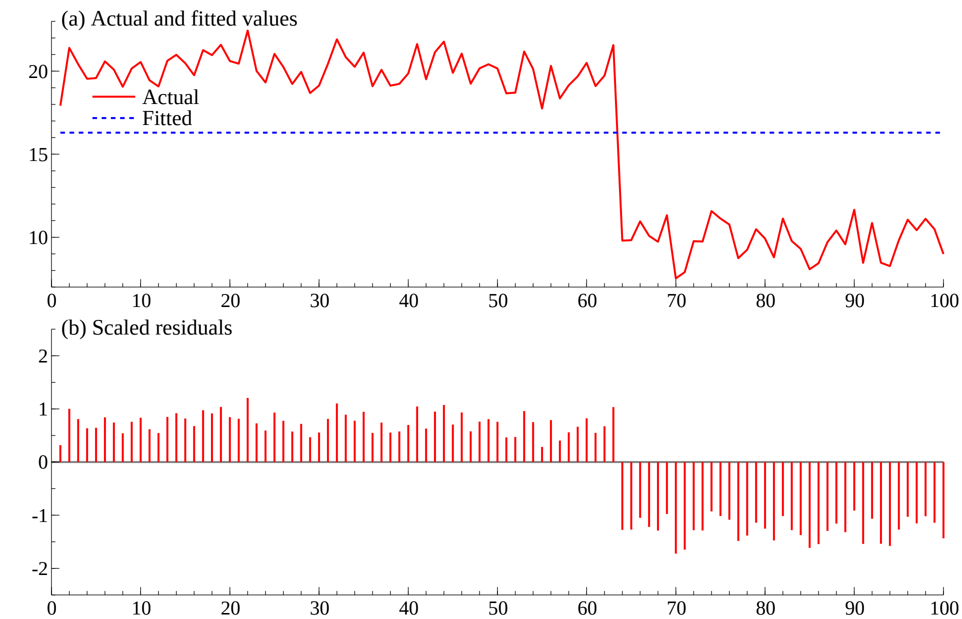

Figure 3: Actual and fitted values and the corresponding scaled residuals for the estimated model when the DGP has a break and the model ignores that break.

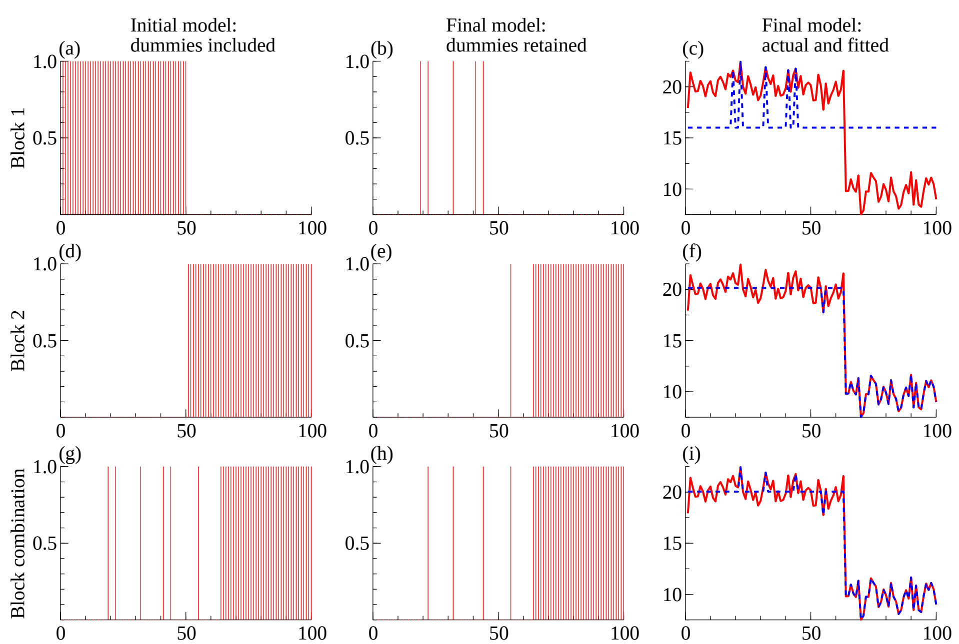

Figure 4: A characterization of bare-bones impulse indicator saturation with a target size of 5% when the DGP has a break and the model ignores that break.

Table 1: Impulse indicator saturation and two extensions, as characterized by the variables involved.

Name	Description	Variables	Definition
Impulse indicator saturation	Zero-one dummies	$\{I_{it}\}$	$I_{it} = 1$ for $t = i$, zero otherwise
Super saturation	Step functions	$\{I_{it}, S_{it}\}$	$S_{it} = 1$ for $t \geq i$, zero otherwise
Ultra saturation	Broken linear trends	$\{I_{it}, S_{it}, T_{it}\}$	$T_{it} = t - i + 1$ for $t \geq i$, zero otherwise

the penultimate graph (Figure 4h) shows, the procedure has high power to detect the break, even although the nature of the break is not utilized in the procedure itself.

In practice, IIS as an algorithm may be more complicated than this bare-bones version, which employs two equally sized blocks, selects dummies by t-tests, and is non-iterative. In Doornik and Hendry's (2013) Autometrics econometrics software, IIS utilizes many possibly unequally sized blocks, rather than just two blocks; the partitioning of the sample into blocks may vary over iterations of searches; dummy selection includes F-tests against a general model; and residual diagnostics help guide model selection. Notably, the specific algorithm for IIS can make or break IIS's usefulness; cf. Doornik (2009), Castle, Fawcett, and Hendry (2010), and Hendry and Doornik (2014). IIS is a statistically valid procedure for integrated, cointegrated data; see Johansen and Nielsen (2009). IIS can serve as a diagnostic statistic, and it can aid in model development, as discussed in Ericsson (2011a).

Many existing procedures can be interpreted as special cases of IIS in that they represent particular algorithmic implementations of IIS. Such special cases include recursive estimation, rolling regression, the Chow (1960) predictive failure statistic (including the 1-step, breakpoint, and forecast versions implemented in OxMetrics), the Andrews (1993) unknown breakpoint test, the Bai and Perron (1998) multiple breakpoint test, tests of extended constancy in Ericsson, Hendry, and Prestwich (1998, pp. 305ff), tests of nonlinearity, intercept correction (in forecasting), and robust estimation. IIS thus provides a general and generic procedure for analyzing a model's constancy. Algorithmically, IIS also solves the problem of having more potential regressors than observations by testing and selecting over blocks of variables.

Table 1 summarizes IIS and two extensions of IIS, drawing on expositions and developments in Ericsson (2011b, 2012) and Ericsson and Reisman (2012). Throughout, T is the sample size, t is the index for time, i and j are the indexes for indicators, k is the index for economic variables (denoted x_{kt}), and K is the total number of potential regressors considered. A few remarks may be helpful for interpreting the entries in Table 1.

Impulse indicator saturation. This is the standard IIS procedure proposed by Hendry (1999), with selection among the T zero-one impulse indicators $\{I_{it}\}$.

Super saturation. Super saturation searches across all possible one-off step functions $\{S_{it}\}$, in addition to $\{I_{it}\}$. Step functions are of economic interest because they may capture permanent or long-lasting changes that are not otherwise incorporated into a specific empirical model. A step function is a partial sum of impulse indicators. Equivalently, a step function is a parsimonious representation of a sequential subset of impulse indicators that have equal coefficients. Castle, Doornik, Hendry, and Pretis (2015) investigate the statistical properties of a closely related saturation estimator—step indicator saturation (SIS)—which searches among only the step indicator variables $\{S_{it}\}$. Autometrics now includes IIS, SIS, super saturation (IIS+SIS), and zero-sum pairwise IIS (mentioned below); see Doornik and Hendry (2013).

Ultra saturation. Ultra saturation (earlier, sometimes called "super duper" saturation) searches across $\{I_{it}, S_{it}, T_{it}\}$, where the $\{T_{it}\}$ are broken linear trends. Broken linear trends may be of economic interest. Mathematically, the $\{T_{it}\}$ are partial sums of the partial sums of impulse indicators. Broken quadratic trends, broken cubic trends, and higher-order broken trends are also feasible.

Table 1 is by no means an exhaustive list of extensions to IIS. Other extensions include sequential ($j = 1$) and non-sequential ($j > 1$) pairwise impulse indicator saturation for an indicator P_{it}, defined as $I_{it} + I_{i+j,t}$; zero-sum pairwise IIS for an indicator Z_{it}, defined as ΔI_{it}; many many variables for a set of K potential regressors $\{x_{kt},\ k = 1, \ldots, K\}$ for $K > T$; factors; principal components; and multiplicative indicator saturation for the set of $S_{it}x_{kt}$. See Ericsson (2011b, 2012) and Castle, Clements, and Hendry (2013) for details, discussion, and examples in the literature. Also, the saturation procedure chosen may itself be a combination of extensions; and that choice may affect the power of the procedure to detect specific alternatives. For instance, in Example 2 above, the 37 impulse indicators $\{I_{it},\ i = 64, \ldots, 100\}$ are not a particularly parsimonious way of expressing the step shift that occurs two thirds of the way through the sample, whereas the single one-off step dummy S_{64t} is.

4. An Empirical Illustration with Brazilian Inflation

Dées, di Mauro, Pesaran, and Smith (2007) develop a global vector autoregression (GVAR), and Ericsson and Reisman (2012) evaluate that GVAR with impulse indicator saturation. Drawing on Ericsson and Reisman (2012), this section illustrates IIS with that GVAR's equation for the Brazilian inflation rate.

IIS for the equation of Brazilian inflation is particularly revealing. IIS detects 14 dummies—all prior to 1995—with an F-statistic of 75.1 [$p = 0.0000$] for the significance of those dummies. As is evident from Figure 5, Brazilian inflation changed its behavior radically after 1994, resulting in historically very low inflation rates and much less variability, and in line with marked changes in government policies. Economically and statistically, IIS for this equation implies that the GVAR model is inadequate to capture the essential features of Brazilian inflation.

Figure 5: The quarterly inflation rate for Brazil (Δp_{Brazil}), as a fraction.

5. Conclusions

Indicator saturation defines a generic procedure for examining empirical model properties, and it provides a potential mechanism for improving such models. In particular, indicator saturation provides a mechanism for systematically detecting and quantifying crises, jumps, and changes in regime. Such structural breaks can be persistent yet time-varying; they can be difficult to detect in a timely fashion; and they may have substantive implications for policy analysis. The IIS approach aims to address these issues, with extensions of IIS aiming to characterize systematic properties across retained impulses. The IIS approach also links directly to techniques for robustifying forecasts, noting that intercept correction is a variant of super saturation; see Clements and Hendry (1996, 1999, 2002), Hendry (2006), and Castle, Fawcett, and Hendry (2010). Computer-automated model selection is central to implementing saturation techniques.

IIS-based tests may detect not only "location shifts", but also outliers and other forms of mis-specification, such as changes in the forecast error variance. For outliers and changes in error variance in particular, retained impulses would not typically be sequential or (for adjacent impulses) of the same sign and similar magnitude. These caveats emphasize the usefulness of super saturation and other extended forms of IIS in characterizing the persistence of structural breaks.

IIS-based tests can serve as diagnostic tools to detect what's wrong with a model, and as developmental tools to suggest how that model may be improved. While rejection of the null doesn't imply the alternative, the *date*-specific nature of IIS-type procedures when applied to time series data can aid in identifying the sources of a model's weaknesses, and hence ways in which a model can be improved.

Use of these tests in model development is consistent with a progressive modeling approach; see White (1990).

The IIS approach has many potential applications, beyond its initial roles in model evaluation and robust estimation. As discussed above, it can be used for detecting crises, jumps, and changes in regime. IIS also provides a framework for creating near-realtime early-warning and rapid-detection devices, e.g., of financial market anomalies. Finally, IIS generalizes to systems. In their examination of computer-automated model selection and saturation techniques, Hendry and Doornik (2014) elucidate how recent econometric tools and software developments have made empirical modeling an exciting and vibrant enterprise in economics. Additional results in the statistical theory of model selection—and corresponding software implementation—will no doubt further aid this endeavor. The current paper serves to exemplify, extend, and illustrate some aspects of how such empirical modeling might proceed.

REFERENCES

Akaike, H. (1973) “Information Theory and an Extension of the Maximum Likelihood Principle”, in B. N. Petrov and F. Csaki (eds.) *Second International Symposium on Information Theory*, Akademiai Kiado, Budapest, 267–281.

Akaike, H. (1981) “Likelihood of a Model and Information Criteria”, *Journal of Econometrics*, 16, 1, 3–14.

Andrews, D. W. K. (1993) “Tests for Parameter Instability and Structural Change with Unknown Change Point”, *Econometrica*, 61, 4, 821–856.

Atkinson, A. C. (1981) “Likelihood Ratios, Posterior Odds and Information Criteria”, *Journal of Econometrics*, 16, 1, 15–20.

Bai, J., and P. Perron (1998) “Estimating and Testing Linear Models with Multiple Structural Changes”, *Econometrica*, 66, 1, 47–78.

Bergamelli, M., and G. Urga (2014) “Detecting Multiple Structural Breaks: Dummy Saturation vs Sequential Bootstrapping. With an Application to the Fisher Relationship for US”, CEA@Cass Working Paper Series No. WP–CEA–03–2014, Cass Business School, London, April.

Campos, J., and N. R. Ericsson (1999) “Constructive Data Mining: Modeling Consumers’ Expenditure in Venezuela”, *Econometrics Journal*, 2, 2, 226–240.

Campos, J., N. R. Ericsson, and D. F. Hendry (2005) “Introduction: *General-to-Specific Modelling*”, in J. Campos, N. R. Ericsson, and D. F. Hendry (eds.) *General-to-Specific Modelling*, Volume I, Edward Elgar, Cheltenham, xi–xci.

Castle, J. L., M. P. Clements, and D. F. Hendry (2013) “Forecasting by Factors, by Variables, by Both or Neither?”, *Journal of Econometrics*, 177, 2, 305–319.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015) “Detecting Location Shifts During Model Selection by Step-indicator Saturation”, *Econometrics*, 3, 2, 240–264.

Castle, J. L., N. W. P. Fawcett, and D. F. Hendry (2010) “Forecasting with Equilibrium-correction Models During Structural Breaks”, *Journal of Econometrics*, 158, 1, 25–36.

Chow, G. C. (1960) “Tests of Equality Between Sets of Coefficients in Two Linear Regressions”, *Econometrica*, 28, 3, 591–605.

Clements, M. P., and D. F. Hendry (1996) “Intercept Corrections and Structural Change”, *Journal of Applied Econometrics*, 11, 5, 475–494.

Clements, M. P., and D. F. Hendry (1999) *Forecasting Non-stationary Economic Time Series*, MIT Press, Cambridge.

Clements, M. P., and D. F. Hendry (2002) “Explaining Forecast Failure in Macroeconomics”, Chapter 23 in M. P. Clements and D. F. Hendry (eds.) *A Companion to Economic Forecasting*, Blackwell Publishers, Oxford, 539–571.

Dées, S., F. di Mauro, M. H. Pesaran, and L. V. Smith (2007) "Exploring the International Linkages of the Euro Area: A Global VAR Analysis", *Journal of Applied Econometrics*, 22, 1, 1–38.

Denton, F. T. (1985) "Data Mining as an Industry", *Review of Economics and Statistics*, 67, 1, 124–127.

Doornik, J. A. (2008) "Encompassing and Automatic Model Selection", *Oxford Bulletin of Economics and Statistics*, 70, supplement, 915–925.

Doornik, J. A. (2009) "Autometrics", Chapter 4 in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, 88–121.

Doornik, J. A., and D. F. Hendry (2009) *PcGive 13*, Timberlake Consultants Press, London (3 volumes).

Doornik, J. A., and D. F. Hendry (2013) *PcGive 14*, Timberlake Consultants Press, London (3 volumes).

Ericsson, N. R. (2011a) "Improving Global Vector Autoregressions", draft, Board of Governors of the Federal Reserve System, Washington, D.C., June.

Ericsson, N. R. (2011b) "Justifying Empirical Macro-econometric Evidence in Practice", invited presentation, online conference *Communications with Economists: Current and Future Trends* commemorating the 25th anniversary of the *Journal of Economic Surveys*, November.

Ericsson, N. R. (2012) "Detecting Crises, Jumps, and Changes in Regime", draft, Board of Governors of the Federal Reserve System, Washington, D.C., November.

Ericsson, N. R. (2015) "How Biased Are U.S. Government Forecasts of the Federal Debt?", *International Journal of Forecasting*, forthcoming.

Ericsson, N. R. (2016) "Eliciting GDP Forecasts from the FOMC's Minutes Around the Financial Crisis", *International Journal of Forecasting*, 32, 2, 571–583.

Ericsson, N. R., D. F. Hendry, and K. M. Prestwich (1998) "The Demand for Broad Money in the United Kingdom, 1878–1993", *Scandinavian Journal of Economics*, 100, 1, 289–324 (with discussion).

Ericsson, N. R., and E. L. Reisman (2012) "Evaluating a Global Vector Autoregression for Forecasting", *International Advances in Economic Research*, 18, 3, 247–258.

Hannan, E. J., and B. G. Quinn (1979) "The Determination of the Order of An Autoregression", *Journal of the Royal Statistical Society, Series B*, 41, 2, 190–195.

Hansen, P. R., A. Lunde, and J. M. Nason (2011) "The Model Confidence Set", *Econometrica*, 79, 2, 453–497.

Hendry, D. F. (1999) "An Econometric Analysis of US Food Expenditure, 1931–1989", Chapter 17 in J. R. Magnus and M. S. Morgan (eds.) *Methodology and Tacit Knowledge: Two Experiments in Econometrics*, John Wiley and Sons, Chichester, 341–361.

Hendry, D. F. (2006) "Robustifying Forecasts from Equilibrium-correction Systems", *Journal of Econometrics*, 135, 1–2, 399–426.

Hendry, D. F., and J. A. Doornik (2014) *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*, MIT Press, Cambridge, Massachusetts.

Hendry, D. F., S. Johansen, and C. Santos (2008) "Automatic Selection of Indicators in a Fully Saturated Regression", *Computational Statistics*, 23, 2, 317–335, 337–339.

Hendry, D. F., and H.-M. Krolzig (1999) "Improving on 'Data Mining Reconsidered' by K. D. Hoover and S. J. Perez", *Econometrics Journal*, 2, 2, 202–219.

Hendry, D. F., and H.-M. Krolzig (2001) *Automatic Econometric Model Selection Using PcGets 1.0*, Timberlake Consultants Press, London.

Hendry, D. F., and H.-M. Krolzig (2003) "New Developments in Automatic General-to-Specific Modeling", Chapter 16 in B. P. Stigum (ed.) *Econometrics and the Philosophy of Economics: Theory-Data Confrontations in Economics*, Princeton University Press, Princeton, 379–419.

Hendry, D. F., and H.-M. Krolzig (2005) "The Properties of Automatic *Gets* Modelling", *Economic Journal*, 115, 502, C32–C61.

Hendry, D. F., and F. Pretis (2013) “Anthropogenic Influences on Atmospheric CO_2”, Chapter 12 in R. Fouquet (ed.) *Handbook on Energy and Climate Change*, Edward Elgar, Cheltenham, 287–326.

Hendry, D. F., and C. Santos (2010) “An Automatic Test of Super Exogeneity”, Chapter 12 in T. Bollerslev, J. R. Russell, and M. W. Watson (eds.) *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Oxford University Press, Oxford, 164–193.

Hoover, K. D., and S. J. Perez (1999) “Data Mining Reconsidered: Encompassing and the General-to-specific Approach to Specification Search”, *Econometrics Journal*, 2, 2, 167–191 (with discussion).

Hoover, K. D., and S. J. Perez (2004) “Truth and Robustness in Cross-country Growth Regressions”, *Oxford Bulletin of Economics and Statistics*, 66, 5, 765–798.

Johansen, S., and B. Nielsen (2009) “An Analysis of the Indicator Saturation Estimator as a Robust Regression Estimator”, Chapter 1 in J. L. Castle and N. Shephard (eds.) *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford University Press, Oxford, 1–36.

Johansen, S., and B. Nielsen (2013) “Outlier Detection in Regression Using an Iterated One-step Approximation to the Huber-skip Estimator”, *Econometrics*, 1, 1, 53–70.

Krolzig, H.-M., and D. F. Hendry (2001) “Computer Automation of General-to-specific Model Selection Procedures”, *Journal of Economic Dynamics and Control*, 25, 6–7, 831–866.

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

Marczak, M., and T. Proietti (2016) “Outlier Detection in Structural Time Series Models: The Indicator Saturation Approach”, *International Journal of Forecasting*, 32, 1, 180–202.

Schwarz, G. (1978) “Estimating the Dimension of a Model”, *Annals of Statistics*, 6, 2, 461–464.

White, H. (1990) “A Consistent Model Selection Procedure Based on m-testing”, Chapter 16 in C. W. J. Granger (ed.) *Modelling Economic Series: Readings in Econometric Methodology*, Oxford University Press, Oxford, 369–383.