Abstract

We model expenditure on food in the USA, using an extended time series. Even when a theory is essentially ‘correct’, it can manifest serious mis-specification if just fitted to data, ignoring its observed characteristics and major external events such as wars, recessions and policy changes. When the same theory is embedded in a general framework embracing dynamics and structural breaks, it performs well even over an extended data period, as shown using Autometrics with impulse-indicator saturation. Although this particular illustration involves a simple theory, the point made is generic, and applies no matter how sophisticated the theory.

JEL classifications: C51, C22.
KEYWORDS: Econometric modelling; Food Expenditure; Structural breaks; Impulse-indicator saturation; Autometrics.

1 Introduction

The recent financial crisis and resulting global recession is a stark reminder that economies experience unexpected changes which can have significant effects on their state and operation (see, inter alia, Hendry and Mizon, 2009a, 2009b, for our analyses of the financial crisis). Such changes—or more precisely, structural breaks—lead to difficulties in economic forecasting (and later empirical modelling), as well as raising doubts about the value of economic models and the theories underlying them. The fact that some structural changes are unexpected certainly means that economists must assess the applicability of the underpinning economic theory, the appropriateness and quality of the empirical observations, and the performance of the modelling methodology used to develop their models. For models to be relevant, reliable and robust they must exploit as fully as possible all the available information from economic theory, empirical observations, computer technology (hardware and software), historical and institutional knowledge, as well as being able to allow for, or adapt to, shifts in the economy.

Unanticipated structural changes and regime shifts may initially cause forecasting difficulties and raise general doubts about the value of the economics profession, but they can also help to distinguish structural from non-structural models—see, e.g., Hendry and Mizon (1993). Structure here is defined as the set of basic features of the economy which are invariant to changes in that economy. Hence a set of necessary conditions for structure in a model is that its parameters are invariant to: (a) an extension of

*Contributing to this volume in honour of Svend Hylleberg is a pleasure for both of us. We have known Svend since the late 1960s, and thoroughly enjoyed our many and varied academic and social interactions with him. GEM will always remember with surprise and admiration Svend’s early starts to the day when they were working together as co-authors. We are grateful to Jennifer L. Castle for helpful comments, and to James J. Reade for updating the data. All the calculations reported below were based on PcGive (see Hendry and Doornik, 2009).
the sample; (b) interventions in the economy such as regime shifts; and (c) extensions of the information set (see Hendry, 1995, for more details).

Although the ideal for reliable analysis may be the development of a structural model ab initio, it is more realistic to seek an ability to quickly: (i) identify a new regime’s characteristics, and (ii) develop a model of that regime. It is unclear precisely how this can be done within a framework of theory-led models developed using a simple-to-general strategy that leads to a tightly-specified empirical implementation of that theory, and only introduces modifications of a limited nature, such as ad hoc stickiness to deal with mis-specified dynamics. Such simple-to-general modelling is fraught with difficulties as has been explained by numerous authors (for recent contributions see inter alia, Hendry, 1995, Johansen, 2006, Juselius and Johansen, 2006, Mizon, 1995, and Spanos, 1995). The modelling strategy adopted in this paper is designed to function in a world of intermittent unanticipated location shifts. By modelling from general-to-simple and focusing on selecting variables rather than models, recent developments in the automation of model selection have produced remarkable results. For discussions of this achievement, see Castle, Doornik and Hendry (2009a, 2009b) and Doornik, Hendry and Nielsen (2009). The results of this large body of research are embodied in the software package Autometrics (see Doornik, 2008, 2009, Hendry and Doornik, 2009).

In this paper, we model expenditure on food in the USA, using time series extended to 2006 by Reade (2008). We illustrate that even though a theory is essentially ‘correct’, it can exhibit serious mis-specification if it is just fitted without paying attention to the observed characteristics of the data, especially substantive external events such as wars, major recessions, and policy regime changes. However, if the same theory is embedded in a more general framework embracing dynamics and the possibility of structural changes, then that original theory can perform well even over an extended data period, as shown here using the automatic model selection approach in Autometrics applying impulse-indicator saturation. Although this particular illustration involves a simple theory, the point being made is generic, and so applies no matter how sophisticated the theory.

The paper is organized as follows. Section 2 outlines the theory behind impulse-indicator saturation (IIS), following Hendry, Johansen and Santos (2008) and Johansen and Nielsen (2009), and then in §2.1 provides an overview of the analysis, approach, and results. Some previous analyses of food expenditure in the USA from shorter data sets are summarized in Section 3. Estimates of the static theory model are presented in Section 4, which is then re-estimated using IIS, still imposing that theory. Next, a dynamic specification embedding the imposed theory in a more general model is explored in Section 4.2, but without IIS cointegration fails. It is the combination of dynamics and IIS that dramatically improves the results, once shifts from food programs, World War II, and the Korean war are detected. Applying Autometrics next to select unrestrictedly from a general model including the theory variables, 2 lags and IIS yields equally good results, eliminates insignificant variables, and has a cointegrating vector corresponding to the theory. Section 4.3 shows that without the indicator variables, cointegration does not occur. The conditional forecasting performance of the model, both from 1952 and from 2003, is evaluated in Section 5. Two solutions to the poor forecasting performance over the final period are explored in Section 5.1, and conclusions are presented in Section 6.

2 Background: dynamics and breaks

When modelling aspects of sophisticated modern economies subject to adjustment (rather than instantaneous responses), habit, expectations formation, and external surprises, then dynamics clearly matter. However, there are also instances when including lags in the specification of an econometric model designed to embody a particular theory will not be sufficient to arrive at a well-specified outcome. Specifically, breaks matter greatly for food demand in a world subject to unanticipated shifts due to health
scares, major wars, and the introduction of policy programs like food relief. Even when a theory is explicit about which variables are relevant, it is rarely exact about dynamics, since the unit of time for agents’ decisions has never been derived theoretically. Here, the data are annual, but even for the demand for food, two lags will transpire to matter empirically. Theory is even less explicit about structural breaks, and if not modelled, these can have a pernicious effect on a model’s fit, parameter estimates, and constancy even when the underlying theory is ‘valid’. To jointly tackle both issues requires either an assumption of omniscience or empirical model discovery (see Doornik et al., 2009). It is essential to model major interventions, such as food programs and switches in rationing, otherwise they can distort the performance of the model. Indeed not modelling them can result in estimated models exhibiting unit roots when none are there (see inter alia, Perron, 1989 and Hendry and Neale, 1991). However, few theory models allow for such shifts and so these other non-modelled effects must be dealt with by including indicators that remove location shifts and outliers.

The inclusion of deterministic time trends, seasonal dummy variables, and event-specific dummy variables is a well established practice in empirical econometric modelling, but the recent developments in allowing an impulse indicator for every observation (leading to more candidate variables than the sample size when other explanatory variables are included) are less well known. Since there are likely to be multiple breaks, impulse-indicator saturation (IIS) offers an attractive approach. Indeed, the concept originated in an earlier re-analysis of the model in Tobin (1950), by including indicators for the interwar period to investigate the heterogeneity that had led other investigators to eschew that sample (see Hendry, 1999). Hendry et al. (2008) and Johansen and Nielsen (2009) analyze the theory of IIS under the null that there are no breaks or outliers, and the latter relate it to robust statistical methods which seek to eliminate possible data contamination. Since IIS is a form of robust estimation, it can jointly tackle data contamination and ‘fat-tailed’ distributions (by removing extreme observations), as well as location shifts and innovation outliers, at the same time as selecting over potential regressors.

In the ‘split sample’ theory, a set of T (the sample size) impulse indicators is created, half are included in the candidate regressor set, and the significant indicators are recorded. Then that half is removed and the remainder added, again recording significant outcomes. Finally, the two sets of significant indicators are included and the subset of significant ones selected. Under the null of no outliers, breaks, or contaminants, the probability that the \(|t|\)-value for an estimated impulse indicator exceeds the corresponding critical value \(c_{\alpha}\) by chance is the nominal significance level \(\alpha\). Consequently, it is almost costless to add \(T\) impulse indicators to the candidate regressor set in a model selection exercise when the significance level for retaining an individual indicator is set at \(\alpha = 1/T\), since \(P(\{|t| > c_{\alpha}\}) = \alpha = 1/T\), so the cost is just one observation ‘dummied out’. The proof in Johansen and Nielsen (2009), covers a wide range of models including regressions and autoregressions (possibly with unit roots). Although their analysis does not jointly select over regressors as well as indicators, as we do below, simulation studies show that does not affect their basic conclusions.

Given the trivial cost of IIS under the null, the next issue is its ability to detect outliers, breaks, or contamination. This is difficult to ascertain theoretically for multiple sources of discrepancies, but Hendry and Santos (2009) and Castle et al. (2009b) provide analyses and simulations under the alternatives of location shifts and outliers, and the latter demonstrate that it performs favourably compared to procedures like Bai and Perron (1998). Moreover, since only \(\alpha T\) indicators would be retained under the null on average, when many are then the procedure must have detected substantive departures from the null. Applying IIS in econometric modelling both assesses the adequacy of a model and allows for external events that have a significant effect on the phenomena being analyzed. Since there will always be more candidate variables, \(N\), than observations, \(T\), when IIS is applied to a model, the use of an automatic procedure to select variables is essential. Autometrics includes IIS amongst the options available to a modeller, and has been shown by Castle et al. (2009a) to perform well when selecting over all candidate variables as well as breaks, so we apply it below to investigate models of food demand.
2.1 Overview

After summarizing some previous models of US food expenditure and analyses of it using shorter time series, Section 4 first illustrates that an economic theory that is fundamentally correct is not coherent with the data set extended to 2002 (§5 considers the discrepant sub-sample to 2006). Indeed, the estimated static model has wrong coefficient signs and is seriously mis-specified, each of which could lead to the (false) rejection of the theory. The model is then re-estimated with 2 lags on each of the theory variables and though the fit is better, the results would also lead to the rejection of the theory. Since the sample period 1931–2002 is long and includes many changes, the alternative of the static model with IIS was estimated to allow for discrepant observations, possibly due to early measurement errors. The results are closer to theory, but almost all the inter war and war years are dummied out, so remain unexplained.

Consequently, Section 4.2 explores the possibility of the theory performing better with a dynamic specification. When an equation is estimated with the theory model imposed, but allowing 2 lags on each variable, the fit is greatly improved, but no cointegration is found and the coefficient on the lagged dependent variable is close to unity. The latter finding is consistent with breaks being captured by unit roots when they are not dealt with directly (see e.g., Perron, 1989, and Hendry and Neale, 1991). Next applying Autometrics with each of the theory variables forced to be included and selecting the lags, yields similar results: there is still no cointegration and the estimated coefficient on the lagged dependent variable remains close to unity. In the light of these results, it seems essential to consider selecting both the lags and indicators in a model with the theory variables imposed (analyses in levels, despite the unit-root–I(1)–non-stationarity of the data, are valid given the results in Sims, Stock and Watson, 1990, since the I(1) to I(0) reduction is not being tested). As a result there is substantial improvement in the fit, strong cointegration now occurs, and there are significant indicators for a food program in the depression, World War II, the Korean war, as well as smaller effects in the 1970s. Now allowing Autometrics to select over the theory variables as well as lags and IIS barely alters the findings, yielding excellent results with sensible coefficient magnitudes and signs, and powerful cointegration.1 The theory variables did not need to be forced into the analysis once a well-specified empirical model was developed: far from being problematic, model selection has little impact once the basic regressor set can account for the data, and indeed greatly improves theory consistency relative to the initial specification.

Given that automatic selection over the combination of the theory variables, lags and IIS has produced an economically meaningful cointegration vector, Section 4.3 adopts the re-parameterization of an equilibrium correction model (EqCM). Nevertheless, it is easy to establish that cointegration alone is an inadequate representation, whereas an EqCM with IIS performs sufficiently well to merit assessing its forecasting ability. Section 5 uses the data up to 1952 to estimate the EqCM with IIS and then generates the conditional forecasts of the change in US per capita expenditure on food ($\Delta e_f$) up to 2002 which produces very good forecasts, particularly impressive given the difference in behaviour of both $e_f$ and $\Delta e_f$ pre and post 1952 – see Figure 1. Indeed, that model can forecast the rest of the sample from pre-1952 estimates, using the very data other investigators omitted as discrepant!

However, estimating the EqCM with IIS up to 2002 and then forecasting the remaining 4 observations delivers poor forecasts, illustrating that unanticipated changes continue to happen. When the underlying distributions shift, the pre-existing conditional expectation is no longer the minimum mean squared error predictor–see Hendry and Mizon (2009b) for details. The last four observations 2003–2006 are in fact very discrepant. This might be the result of data errors–see the large earlier revisions discussed in Hendry (2009). The data for population ($n_t$) and family size ($a_t$) are most likely to suffer in this regard, and could be the culprits, especially as the former is needed in all models despite being theoretically redundant

---

1 The critical values of cointegration tests should allow for dummy variables, as in (e.g.) Nielsen, 1996, Johansen, Mosconi and Nielsen, 2000, and Nielsen, 2003, and those used here for the PcGive unit-root test are from Ericsson and MacKinnon, 2002, which take account of the number of regressors but not their specific form.
given the per capita formulation and the inclusion of $a_t$. Alternatively, there could have been a location shift, and though this could be dealt with by estimating the model up to 2006 with IIS (as is done for the results shown in Figure 13) that is only an ex post solution. Robust forecasting methods (see Hendry, 2006, Castle and Hendry, 2008 and Castle, Fawcett and Hendry, 2009c for explanations) offer a practical alternative in such situations, as seen from Figure 11, and although far from being even an approximation to the conditional expectation, considerably outperforms the congruent, theory-consistent model which had been constant over more than 70 years.

Though we have long argued in favour of general-to-simple modelling, it might be thought that we have abandoned this view in the modelling undertaken in this paper. However, that is not the case: rather we have illustrated what might happen with a simple-to-general approach, namely rejection of the theory, or maintaining a theory model that performs poorly in forecasting and policy analysis. Further, had the modelling from the outset been from a general unrestricted model (GUM) including all the theory variables, lags and IIS, with a general-to-simple strategy adopted via Autometrics, the approach would have been efficient and delivered our finally selected equation in one step.

3 Modelling food demand

An early time-series analysis of USA food expenditure was Tobin (1950), followed by a number of studies reported in Magnus and Morgan (1999). Building on Hendry (2009), we re-examine US food expenditure over 1931–2002, based on the update of Tobin (1950) by Magnus and Morgan (1999) to 1989 and extended as described in Reade (2008). Most contributors to Magnus and Morgan (1999) had found their models were non-constant over the combined inter-war and post-war samples, so did not model the data before 1950. Hendry (1999) found that impulse indicators for a food programme and post-war de-rationing allowed a constant equation to be developed over the (then) whole sample 1931–1989.

The data variables are defined as follows (lower case denotes logs):

- $e_f$: constant-price per capita expenditure on food;
- $p_f - p$: real price of food;
- $e$: constant-price per capita total expenditure;
- $p$: deflator of total expenditure;
- $y$: constant-price per capita income;
- $s = y - e$: an approximation to the savings ratio;
- $a$: average family size;
- $n$: total population of the USA.

Here the basic theory is that:

$$ e_f = f(e, p_f - p, s, a) \tag{1} $$

The earlier study of a cointegrated VAR for the system in Hendry (2009) established that $e$, $s$, and $p_f - p$ were weakly exogenous in the food demand equation, justifying the analysis of models conditional on these variables. Correction to per capita should mean that $n$ is irrelevant, and demographic effects are captured by $a$. According to conventional theory, the anticipated signs are:

$$ \frac{\partial e_f}{\partial e} > 0, \quad \frac{\partial e_f}{\partial (p_f - p)} < 0, \quad \frac{\partial e_f}{\partial s} > 0, \quad \frac{\partial e_f}{\partial a} < 0, \quad \frac{\partial e_f}{\partial n} = 0 \tag{2} $$

Because his estimates of quantities of food supply suggested these had not changed greatly over the period 1912–1948, Tobin (1950) ‘inverted’ (1) to use food price as the dependent variable in his time-series model (he also estimated a cross-section model). Here, we take food prices as being determined in the world market, so model per capita expenditure. Like Tobin, we take $f(\cdot)$ to be multiplicative, and hence linear after logs.
Figure 1 shows the times series, which reveal considerable changes over the period. After falling sharply at the commencement of the Great Depression, both \( e_f \) and \( e \) rise substantially till World War II, fall after, then resume a gentle rise (panels a and c), so \( \Delta e_f \) is much more volatile pre-war (panel e: \( \Delta e_f \) has a similar but less pronounced pattern). Next, \( p_f - p \) is quite volatile till after the war, then is relatively stable (panel b), whereas the dramatic rise in \( s \) from ‘forced saving’ during the war is manifest (panel d). Average family size has fallen considerably.

4 Static models

In this section, the simple theoretical framework of equation (1) with no lagged values of variables is estimated, and following its poor performance, is re-estimated with impulse-indicator saturation (IIS), to investigate whether ‘removing’ major breaks and outliers is sufficient by itself.

The results obtained for the imposed static theory model are:

\[
e_{f,t} = 5.30 + 0.77 e_t + 0.11 (p_f - p)_t + 0.72 s_t - 0.36 a_t - 0.73 n_t \tag{3}
\]

\[
R^2 = 0.94 \quad \hat{\sigma} = 0.055 \quad \chi^2(2) = 19.5^{**} \quad F_{ar}(2, 66) = 44.3^{**}
\]

\[
F_{arch}(1, 72) = 216.8^{**} \quad F_{reset}(2, 66) = 18.1^{**} \quad F_{het}(10, 63) = 23.2^{**}
\]

In (3), \( R^2 \) is the squared multiple correlation, and \( \hat{\sigma} \) is the residual standard deviation, with coefficient standard errors shown in parentheses. The diagnostic tests are of the form \( F_j(k, T - \ell) \) which denotes an approximate F-test against the alternative hypothesis \( j \) for: \( k^{th}\)-order serial correlation (\( F_{ar} \): see Godfrey, 1978), \( k^{th}\)-order autoregressive conditional heteroscedasticity (\( F_{arch} \): see Engle, 1982), heteroscedasticity (\( F_{het} \): see White, 1980); the RESET test (\( F_{reset} \): see Ramsey, 1969); and a chi-square test for
normality ($\chi^2_{nd}(2)$: see Doornik and Hansen, 2008). In the later sections of the paper we also report a parameter constancy test ($F_{Chow}$: see Chow, 1960) over $k$ periods.

The static economic-theory model has a very poor fit, and does not capture well the behaviour of the observed data. The price elasticity has the ‘wrong sign’, contradicting (2), but is insignificant. Also, although it is theoretically irrelevant, population $n_t$ is significant. Finally, every mis-specification test strongly rejects. Figure 2 shows that the estimated model fails to describe the 1930s.

![Figure 2: Static unselected ‘theory’ equation](image)

### 4.1 Static ‘theory’ equation with IIS

Given its poor performance, but the different behaviour of the pre and post war times series, IIS is next introduced into the specification. The resulting selected model is:

$$ e_{f,t} = 0.64 \hat{e}_t - 0.15 (p_f - p)_t + 1.0 s_t - 0.01 a_t - 0.47 n_t + 0.24 - 0.30 I_{29} - 0.27 I_{30} - 0.26 I_{31} - 0.21 I_{32}$$

$$ - 0.16 I_{33} - 0.12 I_{34} - 0.11 I_{35} - 0.08 I_{36} - 0.06 I_{37}$$

$$ - 0.08 I_{41} - 0.17 I_{42} - 0.16 I_{43} - 0.10 I_{44} - 0.10 I_{46}$$

$$ + 0.09 I_{47} + 0.03 I_{70} + 0.03 I_{72} - 0.034 I_{73} - 0.03 I_{98}$$

$$ R^2 = 0.997 \quad \hat{\sigma} = 0.015 \quad F_{ar}(2, 47) = 4.9 \quad \chi^2(2) = 2.3 \quad F_{arch}(1, 72) = 6.7^* \quad F_{reset}(2, 47) = 3.1 \quad F_{het}(10, 44) = 2.3 $$
The impulse-indicators dummy out almost all interwar and war data, and as a result, few of the mis-specification test statistics are significant. Further, the estimated model is much closer to (1), though \( n \) remains significant, a feature of all subsequent results. It is tempting to conclude that the theory in (1) is valid–but only after the war for some reason, although Tobin originally estimated his time-series model over the very period now excluded. We will see that it is features of the data ‘outside’ of the theory that induces that failure (mainly location shifts from policy interventions and wars), and the final model we obtain satisfies most of the theory for most of the sample.

### 4.2 Dynamic models

In this section, two lagged values of each variable are introduced to check whether the absence of a dynamic specification is the reason for the previous ‘rejection’ of the theory. The dynamic model is first estimated without selection for all variables (including \( n \) as a check) and 2 lags:

\[
\begin{align*}
\hat{e}_{f,t} &= 0.71 e_t - 0.20 \ (p_f - p)_t + 0.26 s_t + 0.05 a_t + 0.13 n_t \\
&+ 0.98 \ e_{f,t-1} - 0.14 \ e_{f,t-2} - 0.73 \ e_{t-1} + 0.01 \ e_{t-2} \\
&+ 0.25 \ (p_f - p)_{t-1} - 0.04 \ (p_f - p)_{t-2} + 0.01 \ s_{t-1} - 0.04 \ s_{t-2} \\
&- 0.10 \ a_{t-1} - 1.45 \ n_{t-1} + 1.41 \ n_{t-2} - 3.1 \\
\end{align*}
\]

(5)

\[
R^2 = 0.996 \quad \hat{\sigma} = 0.015 \quad F_{ar}(2, 53) = 0.56 \\
\chi^2(2) = 0.14 \quad F_{arch}(1, 70) = 2.60 \quad F_{reset}(2, 53) = 1.79 \quad F_{het}(32, 39) = 3.78^{**}
\]
The *PcGive* unit-root test does not reject the null of no cointegration ($t_{ur} = -2.59$: see Banerjee and Hendry, 1992, and Ericsson and MacKinnon, 2002) and delivers the pseudo long-run solution:

$$c_1 = e_f + 0.06 e - 0.09(p_f - p) - 1.42 s + 0.31 a - 0.59 n + 19.6$$

Thus, although the fit is hugely improved (see Figure 4), adding lags does not solve the mis-match between the theory and evidence: both key long-run price and expenditure elasticities are ‘wrong signed’.

![Figure 4: Estimated theory model with dynamics](image)

Applying *Autometrics* with the theory variables forced to be retained, but selecting the dynamic reactions, has almost no impact on the match of the theory and evidence, the fit of the model, the coefficient magnitudes and their standard errors, or the lack of cointegration. Selection simply eliminates insignificant lagged variables:

$$e_{f,t} = 0.74 e_t - 0.17 (p_f - p) + 0.27 s_t - 0.05 a_t - 1.21 n_t$$

$$0.84 e_{f,t-1} - 0.75 e_{t-1} + 0.17 (p_f - p)_{t-1} + 1.30 n_{t-1} - 3.01 (6)$$

$$R^2 = 0.995 \quad \hat{\sigma} = 0.015 \quad F_{ar}(2,60) = 0.80$$

$$\chi^2(2) = 0.90 \quad F_{arch}(1,70) = 0.01 \quad F_{reset}(2,60) = 1.60 \quad F_{het}(18,53) = 2.24^*$$

where:

$$c_2 = e_f + 0.045 e - 0.011(p_f - p) - 0.56 n + 19.1 - 1.72 s + 0.33 a \quad t_{ur} = -3.32$$

The finding of no cointegration with ‘wrong signs’, and an estimated lagged dependent variable coefficient close to unity, are consistent with unit roots ‘capturing’ breaks that are not dealt with directly.
Hence, we next estimated a dynamic model including IIS, with automatic selection applied to the dynamics and impulse indicators, but retaining all theory variables.

\[
\begin{align*}
\hat{e}_{f,t} &= 0.61 \hat{e}_t - 0.20 (p_f - p)_t + 0.21 s_t + 0.07 a_t - 0.34 n_t \\
&\quad + 0.49 \hat{e}_{f,t-1} - 0.21 e_{t-1} + 0.14 (p_f - p)_{t-2} + 0.27 s_{t-1} \\
&\quad - 0.15 I_{31} - 0.16 I_{32} - 0.06 I_{33} - 0.03 I_{43} \\
&\quad - 0.03 I_{50} + 0.03 I_{70} - 0.03 I_{73} + 2.12 \\
R^2 &= 0.999 \quad \hat{\sigma} = 0.0078 \quad F_{ar}(2, 53) = 0.16 \\
\chi^2(2) &= 0.07 \quad F_{arch}(1, 70) = 0.01 \quad F_{reset}(2, 53) = 1.26 \quad F_{het}(18, 46) = 0.95 \\
c_3 &= e_f - 0.78c + 0.12(p_f - p) - 0.94s - 4.19 - 0.14a + 0.68\alpha - 14.5**
\end{align*}
\]

There is a vast improvement in fit relative to the static or dynamic models without IIS, in the coherence of the theory and evidence, with the anticipated signs on long-run elasticities, and importantly, cointegration is clearly indicated. The main impulses in (7) are for the food program in the depression, World War II, the Korean War, and smaller impacts in the early 1970s.

We now select over all candidate variables from a GUM with 2 lags and IIS at \(\alpha = 0.01\). The results are almost identical to (7), changed merely by the insignificant variables having been eliminated, with no substantive effect on the theory-model relationship:
4.3 Cointegration

Since an EqCM has a much clearer economic interpretation than a levels autoregression (see e.g., Hendry, 1995, 2008), and the previous results produced a strong and economically meaningful cointegration relation:

\[ c_4 = e_f - 0.63 e + 0.13 (p_f - p) - 1.12 s + 0.45 n \quad \text{where} \quad t_{ur} = -12.1^{**} \]

Again, there are no substantive changes from selecting over all variables, merely the elimination of insignificant regressors. Figure 8 plots the resulting \( c_{4,t} \), which shows a steady return from the disequilibrium following 1929, a further dip at the onset of World War II, peaking around the time of the Korean War, then benign behaviour thereafter (dummies are excluded from the calculation of \( c_{4,t} \)).
relationship, we next estimated this re-parameterization. Nevertheless, using the solved cointegrating relation $c_{4,t}$ as the feedback in an EqCM, but without IIS, again results in the failure of several mis-specification tests ($(R^*)^2$ is computed including an intercept):

$$
\Delta e_{f,t} = 0.85 \Delta e_t - 0.20 \Delta (p_f - p)_t - 0.60 \Delta p_{t-1} + 0.28 \Delta s_t - 0.22 c_{4,t-1}
$$

$$(R^*)^2 = 0.82 \ \hat{\sigma} = 0.015 \ \text{F}_{ar}(2, 65) = 0.50$$

$$\chi^2(2) = 3.08 \ \text{F}_{arch}(1, 70) = 0.28 \ \text{F}_{reset}(2, 65) = 5.46^{**} \ \text{F}_{het}(10, 61) = 8.09^{**}$$
However, introducing IIS into the EqCM with the cointegrating relation $c_4$ delivers an appropriate l(0) form:

$$
\Delta e_{f,t} = 0.13 \Delta e_{f,t-1} + 0.58 \Delta e_t - 0.32 \Delta(p_f - p)_t + 0.23 \Delta s_t - 0.36 c_{4,t-1}
$$

$$
- 0.11 I_{31} - 0.11 I_{32} + 0.03 I_{34} - 0.03 I_{43} + 0.03 I_{70}
$$

(9)

$$(R^*)^2 = 0.95 \, \hat{\sigma} = 0.0083 \, F_{ar}(2, 60) = 0.60
$$

$$
\chi^2(2) = 1.51 \, F_{arch}(1, 70) = 0.06 \, F_{reset}(2, 59) = 0.44 \, F_{het}(10, 56) = 1.08
$$

where $F_{Chow}(21, 41) = 0.62$ for a break after 1981. This model satisfies all the theory and statistical criteria—and would have been the outcome from a one-off general-to-simple selection with IIS after an l(0) reduction.

5 Conditional forecasting

Given that a model with good economic and econometric properties has been estimated, we now assess its ex post ‘forecasting’ ability. Using data up to 1952 to estimate the EqCM with IIS, forecasts of $\Delta e_f$ were generated up to 2002, conditional on observed values of the explanatory variables. The estimated sub-sample model was:

$$
\Delta e_{f,t} = 0.11 \Delta e_{f,t-1} + 0.62 \Delta e_t - 0.28 \Delta(p_f - p)_t + 0.24 \Delta s_t
$$

$$
- 0.35 c_{4,t-1} - 0.11 I_{31} - 0.109 I_{32} + 0.02 I_{34} - 0.03 I_{43}
$$

(10)

$$(R^*)^2 = 0.98 \, \hat{\sigma} = 0.0098 \, F_{ar}(2, 11) = 0.37
$$

$$
\chi^2(2) = 2.42 \, F_{arch}(1, 20) = 0.24 \, F_{reset}(2, 10) = 0.70 \, F_{het}(10, 7) = 0.22
$$

Apart from two observations in the early 1970s (reflected in the retention of an indicator in (9)), parameter constancy is clear—see Figure 9—albeit taking $c_4$ as given. The 1-step (ex post) ‘forecasts’ over the period 1953–2002 perform well, as indicated by the corresponding Chow forecast statistic $F_{Chow}(50, 13) = 0.81$. This result is particularly impressive given that many other investigators of these data had to omit the inter-war period as being discrepant, and the data variation then is manifestly much larger.

5.1 ‘Forecasting’ 2003–2006

However, the model is less impressive in generating conditional forecasts from 2002 of the 4 observations 2003–2006, as is evidenced by Figure 10, and the large Chow statistic $F_{Chow}(4, 62) = 10.95^{**}$. This inability of the model to produce reasonable forecasts of the data for 2003–2006 illustrates that change is ever with us, so apparently good models can break down at any time. The fact that in the presence of unanticipated changes, the pre-existing conditional expectation is no longer the minimum mean squared error predictor is proved in Hendry and Mizon (2009b). This not only entails it is imperative that modellers quickly identify a new regime’s presence and characteristics, and develop a model thereof, but also that rational expectations theory-led models cannot be invariant when the underlying process changes.

That economies are subject to unanticipated change, and that this can make forecasting difficult, is well documented in academic journals and the media. Equally, a literature has developed in recent years
that has explained the ability of robust forecasting methods to out-perform econometric models, and can be used to improve their performance (see e.g., Clements and Hendry, 1998, 1999, Hendry, 2006,
and Castle and Hendry, 2008). Drawing on this literature, we generate robust forecasts by imposing coefficient values from (9) on the differences of all variables:

\[
\Delta e_{f,t} = \Delta e_{f,t-1} + 0.13\Delta^2 e_{f,t-1} + 0.59\Delta^2 e_t + 0.23\Delta^2 s_t - 0.36\Delta c_{4,t-1} - 0.32\Delta^2 (p_f - p)_t
\]

\[
F_{\text{Chow}}(4, 71) = 0.24
\]

(11)

The results recorded in Figure 11 show no evidence of forecast failure.

Differencing reduces step shifts to impulses, and hence offsets systematic mis-forecasting following a location shift. However, note that the forecasts, based on identical parameter values but applied to the differences of the original variables, now show a rise, whereas previously (e.g., Figure 10), they showed a fall. Hendry (2006) explains why differencing is so successful, not only here but in many other settings (see e.g., http://www.ofcom.org.uk/research/tv/reports/tvadvmarket.pdf): after any unmodelled parameter change, an estimated EqCM will reflect the previous parameter values, and so mis-forecast, but the data generation process must embody the parameter shifts, and hence so must the change in the dependent variable, making it an excellent robust forecasting device, albeit one period later.

In algebraic terms for an \( n \)-dimensional vector of variables \( x_t \) determined by the EqCM:

\[
\Delta x_t = \gamma + \lambda (\beta' x_{t-1} - \mu) + \epsilon_t
\]

when the parameters shift over a forecast period from \( T \) to (say):

\[
\Delta x_{T+1} = \gamma^* + \lambda^* (\beta^*)' x_T - \mu^* + \epsilon_{T+1}
\]

but stay constant at these new values over \( T+2 \):

\[
\Delta x_{T+2} = \gamma^* + \lambda^* (\beta^*)' x_{T+1} - \mu^* + \epsilon_{T+2} = \Delta x_{T+1} + [\lambda^* ((\beta^*)' \Delta x_{T+1}) + \Delta \epsilon_{T+2}]
\]

(12)

then \( \Delta x_{T+1} \) ‘contains’ all the information about the shifts—without any modelling or estimation—other than I(-1) noise, shown in brackets. This analysis suggests that (11) should also be able to ‘forecast’ the entire post-war period, and that is indeed the case as Figure 12 shows.

As an alternative, we also selected the model with IIS over the whole sample 1931–2006:
\[ \Delta e_{f,t} = 0.14 \Delta e_{f,t-1} + 0.59 \Delta e_t - 0.32 \Delta (p_f - p)_t + 0.23 \Delta s_t - 0.35 e_{4,t-1} - 0.11 I_{31} - 0.11 I_{32} + 0.03 I_{34} \\
- 0.03 I_{43} + 0.03 I_{70} + 0.03 I_{05} + 0.04 I_{06} \]

\[ (R^*)^2 = 0.94 \quad \hat{\sigma} = 0.0086 \quad F_{ar}(2, 62) = 0.05 \]

\[ \chi^2(2) = 0.64 \quad F_{arch}(1, 74) = 0.21 \quad F_{reset}(2, 62) = 0.43 \quad F_{het}(10, 58) = 0.96 \]

Hence, the results are almost the same as the estimates in (10), but with two extra large indicators in 2005 & 2006 relative to that model. These results show that the final 4 observations require careful analysis in order to understand the changes that have taken place.

6 Conclusion

The focus of this paper has been the difficulties for economic modelling that arise when the underlying economic generation processes change. The recent financial crisis and resulting global recession is a vivid reminder that unexpected changes can have significant effects on both economies and econometric models thereof. Unanticipated changes cause problems for forecasting, and result in estimated models that cast serious doubt on the economic theories on which they are based, when in fact this doubt might be removed by appropriate treatment of discrepant observations. It is also possible that an economic theory might have to be modified in the light of new and important information.

We have revisited a model of expenditure on food in the USA using an extended dataset covering the period 1931–2002, and illustrated the fact that even though a theory is essentially ‘correct’, it can exhibit serious signs of mis-specification if it is just fitted to the data without paying attention to the observed characteristics of the data and significant external events such as food programs, wars, major recessions and policy changes. However, when the theory is embedded in a more general framework embracing
dynamics, outliers and structural change, using automatic model selection as implemented in software like Autometrics, then the original theory may perform well even over an extended data period during which there has been substantial change. Although this particular illustration involves a simple theory, the point being made is generic and applies no matter how sophisticated the theory.

We also conclude that model selection was highly beneficial in detecting and removing structural shifts, and innocuous in eliminating insignificant variables, such that the theory implications were always improved relative to any other specification by commencing from very general models with IIS. Thus, this approach allows all available theory information to be incorporated, evaluated, and retained where it is helpful empirically, while revealing what aspects need modification or even elimination. Despite some claims that there is an inherent ‘trade-off’ between economic theory consistency and data congruence (see e.g., Pagan, 2003), we believe there need be no such conflict once a general modelling methodology is adopted, within which the best available theory is embedded: the above analysis is one illustration thereof.

Finally, robust forecasts can function well even in the face of further unexplained location shifts, although the causes of such shifts merit deeper investigation.

References


