Mis-specification Testing: Non-Invariance of Expectations Models of Inflation

Jennifer L. Castle†, Jurgen A. Doornik†, David F. Hendry† and Ragnar Nymoen*

†Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, UK

⋆Economics Department, Oslo University, Norway

Abstract

Many economic models (such as the new-Keynesian Phillips curve, NKPC) include expected future values, often estimated after replacing the expected value by the actual future outcome, using Instrumental Variables or Generalized Method of Moments. Although crises, breaks and regime shifts are relatively common, the underlying theory does not allow for their occurrence. We show the consequences for such models of breaks in data processes, and propose an impulse-indicator saturation test of such specifications, applied to USA and Euro-area NKPCs.

JEL classifications: C5, E3.
KEYWORDS: Testing invariance; Structural breaks; Expectations; Impulse-indicator saturation; New-Keynesian Phillips curve.

1 Introduction

Expectations play an important role in many economic theories as well as most financial markets. Central Banks use interest rates for inflation 'targets' based on expected, or forecast, inflation one or two years ahead. Nevertheless, it is unclear how accurate agents' expectations of future variables are, even considering sophisticated agents: see e.g., Falch and Nymoen (2011) for one evaluation. Although exchange rates are a key financial price, Nickell (2008) shows the 2-year ahead consensus for £ exchange rate index (ERI) systematically mis-forecasting by a large margin over the extensive time period 1996–2002: Chart 1.16 in Bank of England (2009) shows similar mis-forecasting from 2008. Moreover, the recent collapse of many of the world's largest financial institutions reveals how inaccurate their expectations of asset values were. It is difficult to form accurate expectations because future distributions often differ in unanticipated ways from the present one. Indeed, 'crises' occur with impressive frequency, but unimpressive anticipation: see e.g., Barrell (2001). Consequently, forecast failures are common as e.g., Stock and Watson (1996), and Clements and Hendry (1998) document, and the primary causes of such failures seem to be location shifts, namely shifts in previous unconditional means: see e.g., Hendry (2000, 2006).

*We are grateful to Guillaume Chevillon, Neil R. Ericsson, Katarina Juselius, Søren Johansen, and Bent Nielsen for helpful comments and suggestions. This research was supported in part by grants from the Open Society Institute and the Oxford Martin School.

The currently dominant model of agents' expectations assumes that they are rational and coincide with the conditional expectation, $\mathsf{E}\left[y_{t+1}|\mathcal{I}_t\right]$, of the unknown future value, y_{t+1}, given all relevant information, denoted $\mathcal{I}_t$. Most dynamic stochastic general equilibrium models (DSGEs) impose rational expectations: see e.g., Smets and Wouters (2002). In econometric models of inflation, $\mathsf{E}\left[y_{t+1}|\mathcal{I}_t\right]$ is often replaced by the later outcome:

$$\mathsf{E}\left[y_{t+1} \mid \mathcal{I}_t\right] = y_{t+1} + v_{t+1} \tag{1}$$

with an error that is unpredictable from present information:

$$\mathsf{E}[v_{t+1} \mid \mathcal{I}_t] = 0 \tag{2}$$

Then equations of the form, where z_t is assumed 'exogenous':

$$y_t = \beta_1 \mathsf{E}\left[y_{t+1} \mid \mathcal{I}_t\right] + \beta_2 y_{t-1} + \beta_3 z_t + u_t \tag{3}$$

are re-written substituting from (1), as:

$$y_t = \beta_1 y_{t+1} + \beta_2 y_{t-1} + \beta_3 z_t + \epsilon_t \tag{4}$$

usually with the auxiliary assumption that $\epsilon_t \sim \mathsf{D}\left[0, \sigma_\epsilon^2\right]$ (v_{t+1} in (1) is not independent of y_{t+1}).

The formulation in (4) is almost invariably used in new-Keynesian Phillips curve (NKPC) models of inflation. Estimating the parameters of such equations by Instrumental Variables (IV) or Generalized Method of Moments (GMM) methods usually reveals high inflation persistence (i.e., $\beta_1 + \beta_2$ close to unity), implying large costs of reducing inflation once it rises, and consequently entailing 'tough' interest rate policies to avoid such a scenario. ECB and Bank of England policy during much of 2008 reflected that belief, although other empirical models of inflation suggest persistence is partly an artifact of the model specification: see e.g., Castle (2008).

A parameter is invariant if it is unchanged by extensions of the information set over time, variables, and regimes. We develop tests of the invariance of parameters in expectations models like (4) when there are location shifts in the underlying processes. Previous tests of such feedforward models (see Hendry, 1988, and Engle and Hendry, 1993) used parameter non-constancy to differentiate between models. Here we propose tests based on impulse-indicator saturation (IIS, see Hendry, Johansen and Santos, 2008, Johansen and Nielsen, 2009, and Castle, Doornik and Hendry, 2011b). The impulse indicators are selected in the 'forecasting' (reduced form) equation derived from (3) using the automatic model selection procedure *Autometrics* (see Doornik, 2009, and Castle, Doornik and Hendry, 2011a), then tested for significance in (4). The 2-stage form is related to the test for super exogeneity in Hendry and Santos (2010), but applied to the same variable, albeit time shifted: Hendry (2011) analyzes adding instruments to structural equations. Under the null of invariance, impulse indicators from the reduced form should not be significant when added to (4). Under the alternative of non-invariance to breaks, significant impulse indicators in the reduced form will remain significant when added to (4). Such an analysis is important as many Central Banks and policy agencies use models of the form (4), so a rigorous evaluation is needed to discriminate cases where expectations matter from when they are spuriously significant due to unmodeled breaks. The proposed tests have relevance to all empirical equations with leads, which permeate empirical macromodels in monetary policy.

The structure of the paper is as follows. Section 2 reconsiders the properties of conditional expectations of future values given all relevant information. Section 3 describes impulse-indicator saturation, which will provide the tool for investigating reduced-form location shifts. Section 4 develops the test for invariance in expectations models, then section 5 analyzes the impacts of ignoring breaks on estimates of expectations models. Section 6 provides simulation findings on the application of IIS to such models for testing invariance. Section 7 discusses the formulation of new-Keynesian Phillips curve models that embody (1), then section 7.1 considers previous empirical evidence on NKPC estimation. Sections 8 and 9 respectively report new Euro-area and US NKPC estimates with and without IIS. Section 10 concludes.

2 Models of expectations

A 'rational' expectation (denoted RE, following Muth, 1961) is the conditional expectation of a variable, y_{t+1}, given available information $\mathcal{I}_t$:

$$y^{re}_{t+1} = \mathsf{E}\left[y_{t+1} \mid \mathcal{I}_t\right] = \int y_{t+1}\mathsf{f}\left(y_{t+1} \mid \mathcal{I}_t\right) \mathrm{d}y_{t+1} \tag{5}$$

where $\mathsf{f}(\cdot|\mathcal{I}_t)$ is the relevant conditional distribution. Agents are assumed to use RE as it avoids arbitrage, and hence unnecessary losses, and conditional expectations are believed to be minimum mean square error predictors. Since RE requires free information, unlimited computing power, and free discovery of the form of $\mathsf{E}[y_{t+1}|\mathcal{I}_t]$, such an approach has many critics (see e.g., Kirman, 1989, Frydman and Goldberg, 2007, and Juselius, 2006). Nevertheless, in stationary processes (including difference stationary and trend stationary), assuming that agents use $\mathsf{E}[y_{t+1}|\mathcal{I}_t]$ is not unreasonable, perhaps with learning (see e.g., Evans and Honkapohja, 2001).

From (2), $\mathsf{E}[v_{t+1}|\mathcal{I}_t] = 0$, may be thought to imply that $\mathsf{E}\left[y_{t+1}|\mathcal{I}_t\right]$ is an unbiased predictor of y_{t+1}. However, expectations need to be subscripted by the distribution over which they are calculated, since they are implicitly conditional on that, as well as on $\mathcal{I}_t$ (see Hendry and Mizon, 2010). As economic processes lack time invariance, without a 'crystal ball' assumption that agents know the future distribution in advance, (5) should be written formally as:

$$y^{re}_{t+1} = \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} \mid \mathcal{I}_t\right] = \int y_{t+1}\mathsf{f}_t\left(y_{t+1} \mid \mathcal{I}_t\right) \mathrm{d}y_{t+1} = \mu_t \tag{6}$$

in which case y^{re}_{t+1} will be unbiased for y_{t+1} only if $\mathsf{f}_{t+1}(\cdot) = \mathsf{f}_t(\cdot)$, so $\mathsf{E}_{\mathsf{f}_{t+1}} = \mathsf{E}_{\mathsf{f}_t}$, as:

$$\mathsf{E}_{\mathsf{f}_{t+1}}\left[y_{t+1} \mid \mathcal{I}_t\right] = \int y_{t+1}\mathsf{f}_{t+1}\left(y_{t+1} \mid \mathcal{I}_t\right) \mathrm{d}y_{t+1} = \mu_{t+1} \tag{7}$$

When $\mu_t \neq \mu_{t+1}$, (6) is integrating over a distribution that is not relevant for $t+1$, so the zero conditional expectation of v_{t+1} over f_t in (2) does not entail an unbiased outcome over f_{t+1} as then $\mathsf{E}_{\mathsf{f}_{t+1}}[v_{t+1}|\mathcal{I}_t] = \mu_{t+1} - \mu_t \neq 0$.

Explicit subscripting of the expectations operator, $\mathsf{E}_{\mathsf{f}_{t+1}}$, is crucial for a valid analysis. That formalization was deliberately omitted in both (1) and (5) to reflect widely-used conventions. In practice, the best any agent can do is to form a 'sensible expectation', y^{se}_{t+1}, which involves 'forecasting' $\mathsf{f}_{t+1}(\cdot)$ by $\widehat{\mathsf{f}}_{t+1}(\cdot)$:

$$y^{se}_{t+1} = \int y_{t+1}\widehat{\mathsf{f}}_{t+1}\left(y_{t+1} \mid \mathcal{I}_t\right) \mathrm{d}y_{t+1}. \tag{8}$$

However, if the moments of $\mathsf{f}_{t+1}(y_{t+1}|\mathcal{I}_t)$ alter unexpectedly, there are no good rules for $\widehat{f}_{t+1}(\cdot)$, except that $\mathsf{f}_t(\cdot)$ is rarely a good choice when there are location shifts. Agents cannot know $\mathsf{f}_{t+1}(\cdot)$ at time t when there is a failure of time invariance. Since RE therefore requires $\mathsf{f}_{t+1}(\cdot) = \mathsf{f}_t(\cdot)$ to be even unbiased, its viability depends on the extent and magnitude of distributional shifts in the underlying processes, so we now turn to their detection.

3 Impulse-indicator saturation

Impulse-indicator saturation (IIS) adds to the set of candidate regressors an indicator for every observation. The theory of IIS is derived under the null of no breaks or outliers, but with the aim of detecting and removing outliers and location shifts when they are present. We first describe the simplest form of 'split half' IIS, the case for which Hendry *et al.* (2008) and Johansen and Nielsen (2009) develop an analytic theory and derive the resulting distributions of estimators, then consider the more sophisticated algorithm used by *Autometrics*, an Ox Package implementing automatic model selection: see Doornik (2007, 2009).

First, add half the impulse indicators to the model (i.e., $T/2$ for T observations when there are fewer than $T/2$ other regressors), record the significant ones, then drop that first set of impulse indicators. Now add the other half, recording again. These first two steps correspond to 'dummying out' $T/2$ observations for estimation, noting that impulse indicators are mutually orthogonal. Finally combine the recorded indicators and select the significant subset. Under the null of no outliers or location shifts, Johansen and Nielsen (2009) derive the distribution of IIS for dynamic models with possibly unit roots, and show that on average, when α is the nominal significance level, then αT indicators will be retained adventitiously: the actual null retention rate is called the gauge. Moreover, Johansen and Nielsen (2009) generalize the theory to more, and unequal, splits, and prove that under the null of no outliers or shifts, there is almost no loss of efficiency in testing for T impulse indicators when setting $\alpha \leq 1/T$, even in dynamic models. While such high efficiency despite having more candidate regressors than observations is surprising at first sight, retaining an impulse indicator when it is not needed merely 'removes' one observation, so the loss of efficiency is just of the order of $100\alpha\%$.

The algorithm used by *Autometrics* has several block divisions and does not rely on the impulse indicators being orthogonal. Monte Carlo experiments of IIS in Hendry and Santos (2010) and Castle *et al.* (2011b) at the recommended tight significance levels, have confirmed the null distribution. Hendry and Santos (2010) analyze the ability of IIS to detect a single location shift (called the potency), and Castle *et al.* (2011b) show in simulations that IIS is capable of detecting multiple shifts, including breaks close to the start and end of the sample. This procedure is used to select the significant indicators in the reduced form, to be tested for significance in the expectations equation.

4 Testing expectations models for invariance

We test the invariance of expectations and feedback mechanisms when there are location shifts by comparing selection and estimation with and without IIS under both null and alternative. The general form of DGP for a single shift over $t = 1, \ldots, T$ is given by:

$$y_t = \beta_1 y^e_{t+1} + \beta_2 y_{t-1} + \beta_3 z_t + \psi d_{(T_1,T_2),t} + \epsilon_t, \quad \epsilon_t \sim \mathsf{IN}\left[0, \sigma^2_\epsilon\right] \tag{9}$$

$$z_t = \lambda_0 + \lambda_1 z_{t-1} + \lambda_2 z_{t-2} + \eta_t, \quad \eta_t \sim \mathsf{IN}\left[0, \sigma^2_\eta\right], \tag{10}$$

where $y^e_{t+1} = \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1}|z_t, \mathcal{I}_t\right]$ is the conditional expectation when $\mathcal{I}_t$ is the additional information set available at t. Then $\mathsf{E}_{\mathsf{f}_t}\left[\epsilon_t|z_t, \mathcal{I}_t\right] = 0$ when z_t is exogenous and observed. An AR(2) process (at least) for the exogenous variable is required for identification: see Pesaran (1981).

When $\psi \neq 0$, there is an unanticipated location shift in the equation for y_t of ψ/σ_ϵ standard deviations from T_1 to T_2, denoted $d_{(T_1,T_2),t} = (1_{T_1,t} + \cdots + 1_{T_2,t})$ when $1_{T_1,t}$ is an indicator equal to unity only when $t = T_1$. When $\psi = 0$ but $d_{(T_1,T_2),t}$ is part of $\mathcal{I}_t$, there is an anticipated location shift in the forcing, or reduced form, equation for y_t, so the reduced form shifts but the structural equation (9) is constant. In that case, $d_{(T_1,T_2),t}$ should be significant in the reduced form, but not in (9). The optimal test would include the relevant dummy $d_{(T_1,T_2),t}$ in (9) and conduct a t-test on its significance, although when $\psi = 0$, that procedure, and the test proposed below, will have no power to detect a lack of invariance in (9).

Knowledge of the form and timing of such location shifts is rarely available to the econometrician, so we propose approximating $d_{(T_1,T_2),t}$ by selecting indicators in the reduced form by IIS and using those for the test. Testing is undertaken in two stages. First, the reduced form is estimated with IIS at significance level α_1, to obtain the set of indicators:

$$y_t = \rho y_{t-1} + \gamma_0 z_t + \gamma_1 z_{t-1} + \sum_{i=1}^{T} \delta_i 1_i + u_t \tag{11}$$

where y_{t-1}, z_t and z_{t-1} are forced to enter the regression (so only the indicators are selected over: see Hendry and Johansen, 2011). When there are no breaks, $\alpha_1 T$ indicators will be retained by chance in (11), hence usually setting $\alpha_1 \leq 1/T$. For $T = 100$ and $\alpha_1 = 0.01$, say, because $T\alpha_1 = 1$, the probability of retaining more than two irrelevant indicators is:

$$\mathsf{p}_1 = 1 - \sum_{i=0}^{2} \frac{(1)^i}{i!} e^{-1} \simeq 8\%.$$

However, under normality, let $h > 2/c_{\alpha_1}$ then when more than one indicator is retained, the probability any one has a t-value exceeding hc_{α_1} is:

$$\Pr\left(|\mathsf{t}| \geq hc_{\alpha_1} \mid \mathsf{H}_0\right) \leq \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{h^2}{2} c^2_{\alpha_1}\right)$$

which is 0.01% at $h = 1.5$ and $c_{0.01} = 2.65$. Thus, one null-rejection decision rule before proceeding to the second stage is that more than one indicator is retained, and the larger $|\mathsf{t}|$-value exceeds $1.5 \times c_{\alpha_1}$. A stringent test is justified here both by the form of the 2-stage test and to accord the 'benefit of doubt' to the incumbent. The potency at the second stage should remain high for substantial breaks (e.g., larger than 5σ in the reduced form).

In empirical applications, an easier decision rule to use before conducting the test at stage two is that the F-test probability for all the retained indicators at the first stage is less than 0.1%. Since indicators are orthogonal, this uses the approximation for τ retained indicators:

$$\mathsf{F}_{\mathsf{IIS}}\left(\tau, T-\tau\right) \simeq \frac{1}{\tau} \sum_{i=1}^{\tau} \mathsf{t}^2_i$$

Any retained indicators, denoted $\mathbf{d}_t$, are then included in the structural equation (12), which is estimated by 2SLS with the set of instruments including the constant, z_{t-1} and z_{t-2}:

$$y_t = \beta_1 y_{t+1} + \beta_2 y_{t-1} + \beta_3 z_t + \boldsymbol{\rho}'\mathbf{d}_t + \nu_t \tag{12}$$

The test of their relevance is an approximate F-test, denoted $\mathsf{F}_{\mathsf{inv}}$, of H_0:$\boldsymbol{\rho} = \mathbf{0}$ at α_2, where $\alpha_1 > \alpha_2$ (0.1% versus 0.05%, say) to avoid rejecting on chance retained indicators from (11). Here there are two nulls of interest:
(a) the DGP is the expectations process (9) with $\psi = 0$, but there are shifts in the reduced form (11), so invariance holds; and
(b) $\beta_1 = 0$, so expectations do not matter in the DGP but future values are included in the model, and $\boldsymbol{\rho} \neq \mathbf{0}$ so breaks occur in (12), which is the case considered in the next section.

4.1 Anticipated breaks

When the reduced form has a single shift over $t = 1, \ldots, T$ given by:

$$y_t = \beta_1 y^e_{t+1} + \beta_2 y_{t-1} + \beta_3 z_t + \epsilon_t, \qquad \epsilon_t \sim \mathsf{IN}\left[0, \sigma^2_\epsilon\right] \tag{13}$$

$$z_t = \lambda_0 + \lambda_1 z_{t-1} + \lambda_2 z_{t-2} + \phi d_{(T_1,T_2),t} + \eta_t, \qquad \eta_t \sim \mathsf{IN}\left[0, \sigma^2_\eta\right], \tag{14}$$

whereas the structural equation is constant, then:

$$\begin{aligned}\mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} \mid z_t, \mathcal{I}_t\right] &= \beta_1 \mathsf{E}_{\mathsf{f}_t}\left[y^e_{t+2} \mid z_t, \mathcal{I}_t\right] + \beta_2 \mathsf{E}_{\mathsf{f}_t}\left[y_t \mid z_t, \mathcal{I}_t\right] + \beta_3 \mathsf{E}_{\mathsf{f}_t}\left[z_{t+1} \mid z_t, \mathcal{I}_t\right] \\ &= \beta_1 \mathsf{E}_{\mathsf{f}_t}\left[y^e_{t+2} \mid z_t, \mathcal{I}_t\right] + \beta_1\beta_2 \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} \mid z_t, \mathcal{I}_t\right] + \beta_3\lambda_0 \\ &\quad + \beta_2^2 y_{t-1} + \beta_3\left(\beta_2 + \lambda_1\right) z_t + \beta_3\lambda_2 z_{t-1} + \beta_3\phi \mathsf{E}_{\mathsf{f}_t}\left[d_{(T_1,T_2),t}\right]\end{aligned}$$

where:

$$\mathsf{E}_{\mathsf{f}_t}\left[y_t \mid z_t, \mathcal{I}_t\right] = \beta_1 \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} \mid z_t, \mathcal{I}_t\right] + \beta_2 y_{t-1} + \beta_3 z_t$$

and:

$$\mathsf{E}_{\mathsf{f}_t}\left[z_{t+1} \mid z_t, \mathcal{I}_t\right] = \lambda_0 + \lambda_1 z_t + \lambda_2 z_{t-1} + \phi \mathsf{E}_{\mathsf{f}_t}\left[d_{(T_1,T_2),t}\right]$$

so when $\mathsf{E}_{\mathsf{f}_t}\left[d_{(T_1,T_2),t}\right] = d_{(T_1,T_2),t}$

$$\begin{aligned}y^e_{t+1} &= \left\{\beta_1 \mathsf{E}_{\mathsf{f}_t}\left[y^e_{t+2} \mid z_t, \mathcal{I}_t\right] + \beta_3\lambda_0 + \beta_2^2 y_{t-1} + \beta_3\left(\beta_2 + \lambda_1\right) z_t \right. \\ &\quad \left. + \beta_3\lambda_2 z_{t-1} + \beta_3\phi d_{(T_1,T_2),t}\right\} / \left(1 - \beta_1\beta_2\right) \\ &= g\left(y_{t-1}, z_t, z_{t-1}, d_{(T_1,T_2),t}\right)\end{aligned} \tag{15}$$

is the conditional expectation. When $\phi \neq 0$, there is an anticipated location shift in the reduced form equation for y_t, so $d_{(T_1,T_2),t}$ should be significant there, but not in (13) as the shift is fully incorporated in y^e_{t+1}. The optimal test would include the relevant dummy $d_{(T_1,T_2),t}$ in (13) and conduct a t-test on its significance. Here we use IIS to find any shifts in the reduced form to approximate $d_{(T_1,T_2),t}$ and add the resulting indicators to (13) where they should be insignificant.

5 Estimating expectations models when unanticipated breaks occur

Once breaks occur, and under the null that $\beta_1 = 0$ in (4), the DGP at time t becomes:

$$y_t = \beta_2 y_{t-1} + \beta_3 z_t + \boldsymbol{\rho}'\mathbf{d}_t + \eta_t \tag{16}$$

where $\mathbf{d}_t$ denotes a vector of dummies for location shifts, and η_t is the resulting constant-distribution error. The process at $t+1$ is:

$$y_{t+1} = \beta_2 y_t + \beta_3 z_{t+1} + \boldsymbol{\rho}'\mathbf{d}_{t+1} + \eta_{t+1} \tag{17}$$

Subtracting (17) from (16) to difference the location shifts to impulses, and renormalizing by $(1+\beta_2)$, denoted by ^{a} where $\beta_1^a = 1/(1+\beta_2)$:

$$\begin{aligned} y_t &= \beta_1^a y_{t+1} + \beta_2^a y_{t-1} - \boldsymbol{\rho}^{\mathbf{a}\prime}\Delta\mathbf{d}_{t+1} - \beta_3^a \Delta z_{t+1} - \Delta\eta_{t+1}^a \\ &= \beta_1^a y_{t+1} + \beta_2^a y_{t-1} + \beta_3^a z_t + u_t \end{aligned} \tag{18}$$

This transformation introduces the future value, even though expectations are not part of the DGP. The differenced dummies become 'blips' rather than impulses, or impulses rather than step shifts, so if not directly tested for, will be treated as part of the error in a formulation like (18). Finally, only including z_t as a regressor will lead to a small coefficient given the downward bias due to omitting the opposite signed future value. Thus, a 'hybrid' equation is artificially created, where even $\beta_1^a + \beta_2^a \simeq 1$. We now investigate the consequences of estimating models like (4), when $\beta_1 = 0$, but are mimicked by (18).

5.1 Static DGP

We consider the simplest case where $\beta_2 = \beta_3 = 0$, perhaps after implicit application of the Frisch and Waugh (1933) theorem to remove any exogenous regressors, so that (18) becomes:

$$y_t = y_{t+1} - \boldsymbol{\rho}'\Delta\mathbf{d}_{t+1} - \Delta\eta_{t+1} \tag{19}$$

This suggests that a coefficient near unity may be obtained for β_1 when estimating (19) using instrumental variables (IVs) that are correlated with y_{t+1} and orthogonal to $\Delta\eta_{t+1}$. Since the break in (16) is also partly proxied by the lagged dependent variable, providing lagged values of y_t are used as instruments, even after instrumenting, y_{t+1} will 'pick up' a spurious effect and lead to a large coefficient in (19).

To illustrate, when y_{t-1} is the only IV used in estimation for the model:

$$y_t = \theta y_{t+1} + e_t \tag{20}$$

then from (16) for a sample $t = 1, \ldots, T$ (assuming the moments exist):

$$\begin{aligned} \mathsf{E}\left[\widehat{\theta}\right] &= \mathsf{E}\left[\frac{\sum_{t=2}^{T-1} y_t y_{t-1}}{\sum_{t=2}^{T-1} y_{t+1} y_{t-1}}\right] = \mathsf{E}\left[\frac{\sum\left(\boldsymbol{\rho}'\mathbf{d}_{t-1} + \eta_{t-1}\right)\left(\boldsymbol{\rho}'\mathbf{d}_t + \eta_t\right)}{\sum\left(\boldsymbol{\rho}'\mathbf{d}_{t-1} + \eta_{t-1}\right)\left(\boldsymbol{\rho}'\mathbf{d}_{t+1} + \eta_{t+1}\right)}\right] \\ &\simeq \frac{\boldsymbol{\rho}'\left(\sum \mathbf{d}_{t-1}\mathbf{d}_t'\right)\boldsymbol{\rho}}{\boldsymbol{\rho}'\left(\sum \mathbf{d}_{t-1}\mathbf{d}_{t+1}'\right)\boldsymbol{\rho}}. \end{aligned}$$

Even if there is just a single location shift of size δ from T_1 to $T_2 > T_1 + 2$ so:

$$\mathbf{d}_t' = \left(1_{\{T_1\}}\ 1_{\{T_1+1\}} \cdots 1_{\{T_2\}}\right)$$

where $1_{\{t\}}$ is an indicator for observation t, as $\sum_{j=T_1}^{T_2} 1_{\{j\}} = 1_{\{T_1 \le t \le T_2\}}$, then $\boldsymbol{\rho}'\mathbf{d}_t = \delta 1_{\{T_1 \le t \le T_2\}}$ and hence:

$$\boldsymbol{\rho}' \left(\sum_{t=2}^{T-1} \mathbf{d}_{t-1}\mathbf{d}_t' \right) \boldsymbol{\rho} = \delta^2 \sum 1_{\{T_1 \le t-1 \le T_2\}} 1_{\{T_1 \le t \le T_2\}} = \delta^2 \left(T_2 - T_1\right)$$
$$\boldsymbol{\rho}' \left(\sum_{t=2}^{T-1} \mathbf{d}_{t-1}\mathbf{d}_{t+1}' \right) \boldsymbol{\rho} = \delta^2 \sum 1_{\{T_1 \le t-1 \le T_2\}} 1_{\{T_1 \le t+1 \le T_2\}} = \delta^2 \left(T_2 - T_1 - 1\right)$$

leading to the estimate in (20) having the approximate expected value:

$$\mathsf{E}\left[\widehat{\theta}\right] \simeq \frac{\left(T_2 - T_1\right)\delta^2}{\left(T_2 - T_1 - 1\right)\delta^2} = \frac{\left(T_2 - T_1\right)}{\left(T_2 - T_1 - 1\right)} \simeq 1. \tag{21}$$

Consequently, despite the complete irrelevance of y_{t+1} in the DGP, and the apparently valid use of the lagged value y_{t-1} as an instrument, the estimated coefficient of β_1 will be near unity when there are unmodeled location shifts. Even if $T_2 - T_1 - 1$ is as small as 3, a notable coefficient will be obtained. The estimated standard error of $\widehat{\theta}$ will be approximately:

$$\mathsf{SE}\left[\widehat{\theta}\right] \simeq \frac{\sqrt{2\sigma_\eta^2}}{\sqrt{\left(T_2 - T_1 - 1\right)\delta^2}}$$

as:

$$e_t = \delta\left(1_{\{T_2\}} - 1_{\{T_1\}}\right) - \Delta\eta_{t+1}$$

so:

$$\mathsf{E}\left[\widehat{\sigma}_e^2\right] \simeq 2\left(\sigma_\eta^2 + T^{-1}\delta^2\right)$$

If there is a single location shift of $\delta = r\sigma_\eta$:

$$\mathsf{SE}\left[\widehat{\theta}\right] \simeq \frac{\sqrt{2\left(1 + r^2\right)}\sigma_\eta}{r\sigma_\eta\sqrt{\left(T_2 - T_1 - 1\right)}} = \frac{\sqrt{2\left(1 + r^{-2}\right)}}{\sqrt{\left(T_2 - T_1 - 1\right)}} \tag{22}$$

which will be less than $1/2$ for even small and relatively short breaks (e.g., $r = 3$ and $T_2 - T_1 = 7$) leading to a 'significant' $\widehat{\theta}$.

5.2 Dynamic DGP, dynamic model

We now allow for a dynamic DGP:

$$y_t = \kappa y_{t-1} + \boldsymbol{\rho}'\mathbf{d}_t + \eta_t \tag{23}$$

The model is:

$$y_t = \theta_1 y_{t+1} + \theta_2 y_{t-1} + e_t \tag{24}$$

For the same shift, $\boldsymbol{\rho}'\mathbf{d}_t = \delta 1_{\{T_1 \leq t \leq T_2\}}$, estimation of (24) using y_{t-2} as the identifying instrument yields:[1]

$$\mathsf{E}\left[\left(\begin{array}{c}\widetilde{\theta}_1\\ \widetilde{\theta}_2\end{array}\right)\right] \simeq \frac{1}{(1-\kappa-\kappa^2)}\left(\begin{array}{c}(1-\kappa)^2\\ \kappa\,(1-2\kappa+\kappa^2)\end{array}\right) \tag{25}$$

where $\delta \neq 0$. Because of an approximation that $\kappa^3 \simeq 0$, values of κ have to be less than about 0.5 in (25). For example, when $\kappa = 0.35$, (25) delivers:

$$\mathsf{E}\left[\left(\begin{array}{c}\widetilde{\theta}_1\\ \widetilde{\theta}_2\end{array}\right)\right] \simeq \frac{1}{(1-0.35-0.35^2)}\left(\begin{array}{c}(1-0.35)^2\\ 0.35\times(1-2\times 0.35+0.35^2)\end{array}\right) = \left(\begin{array}{c}0.81\\ 0.28\end{array}\right)$$

so there would be a root just outside the unit circle. If $\kappa = 0$, then:

$$\mathsf{E}\left[\left(\begin{array}{c}\widetilde{\theta}_1\\ \widetilde{\theta}_2\end{array}\right)\right] \simeq \left(\begin{array}{c}1\\ 0\end{array}\right).$$

matching (21). Consequently, expectations are estimated to be important, even when they are in fact irrelevant, and persistence is thought to be high even though κ is also zero.

6 Simulation of expectations versus feedback models

The Monte Carlo simulations assess the properties of testing the invariance of expectations models when there are location shifts by selecting indicators with IIS by *Autometrics* under both null and alternative. The experimental design covers four cases: a backward-looking DGP with no future expectations term (the null), and a DGP with forward- and backward-looking mechanisms (the alternative), both with and without breaks. Parameter values in (9) and (10) are $\psi = 5$, $\beta_3 = 1$, $\lambda_0 = 0$, $\lambda_1 = 1.5$, $\lambda_2 = -0.7$, $\sigma^2_\epsilon = 1$, $\sigma^2_\eta = 1$, $T = 100$, and $M = 1000$ replications are undertaken. We examine a location shift in the equation for y_t of $\psi = 5$ error standard deviations from $T_1 = 81$, denoted $d_{(81,100),t} = (1_{81,t} + \cdots + 1_{100,t})$.[2]

Six models are considered, namely (i) expectations: $\beta_1 \neq 0$, $\beta_2 = 0$; (ii) hybrid: $\beta_1 \neq 0$, $\beta_2 \neq 0$; and (iii) feedback: $\beta_1 = 0$, $\beta_2 \neq 0$; and these three models augmented by IIS with selection of the indicators at $\alpha = 0.01$.

RESULTS TO COME

7 The New-Keynesian Phillips curve

The 'hybrid' new-Keynesian Phillips curve (NKPC) is usually given by a model of the form:

$$\Delta p_t = \underset{\geq 0}{\gamma_\ell}\mathsf{E}_t\left[\Delta p_{t+1}\right] + \underset{\geq 0}{\gamma_b}\Delta p_{t-1} + \underset{\geq 0}{\lambda}\, s_t + u_t \tag{26}$$

where Δp_t is the rate of inflation, $\mathsf{E}_t\left[\Delta p_{t+1}\right]$ is expected inflation one-period ahead conditional on information available today, using the conventions of the literature, and s_t denotes firms'

[1] Detailed calculations are available on request.

[2] Diagnostic tests are switched off during selection as the reduced-form errors are necessarily autocorrelated, but are reported for the final models.

real marginal costs. For estimation, $\mathsf{E}_t\left[\Delta p_{t+1}\right]$ in (26) is usually replaced by Δp_{t+1} as in (1), leading to:

$$\Delta p_t = \gamma_\ell \Delta p_{t+1} + \boldsymbol{\beta}' \mathbf{x}_t + \epsilon_t \text{ where } \epsilon_t \sim \mathsf{D}\left[0, \sigma_\epsilon^2\right] \tag{27}$$

which includes Δp_{t+1} as a feedforward variable, where all other variables (including lags) are components of $\mathbf{x}_t$. Generally, Δp_{t+1} in (27) is instrumented by k variables $\mathbf{z}_t = (\mathbf{x}_t' : \mathbf{w}_t')'$ using whole-sample estimates based on GMM, thereby implicitly postulating relationships of the form:

$$\Delta p_t = \boldsymbol{\kappa}' \mathbf{z}_t + v_t \tag{28}$$

as in Galí and Gertler (1999) and Galí, Gertler and Lopez-Salido (2001): compare Bjørnstad and Nymoen (2008). Mavroeidis (2004) discusses the potential problems of weak identification in such forward-looking models.

To test the invariance of γ_ℓ, IIS is applied to the marginal model (28) for Δp_t ('the forecasting equation') to check for location shifts, then the retained impulses are added to the structural equation (27) and tested for significance. Impulses that matter in the marginal model for Δp_t should nevertheless be insignificant in models like (27) when that is correctly specified, as they should enter through $\mathsf{E}_t\left[\Delta p_{t+1}\right]$. Consequently, the significance of added dummies refutes invariance of the equation. If estimates of γ_ℓ also cease to be significant, that entails the potential spurious significance of the feedforward terms (27) as proxies for the unmodeled location shifts, as simulated in the previous section.

Intuitively, because Δp_{t+1} reflects breaks before they occur, as seen from time t, even instrumenting Δp_{t+1} could let it act as a proxy for those breaks, leading to γ_ℓ being 'spuriously significant' in (27). As breaks are generally unanticipated, even by sophisticated economic agents, precisely in a setting where (27) is an invalid representation, we have shown above that one would find $\widehat{\gamma}_\ell \neq 0$.

7.1 Empirical evidence on NKPC estimation

The basis for the NKPC as a model of inflation is mainly due to results like those of Galí and Gertler (1999) (henceforth GG) on US data, and Galí *et al.* (2001) (henceforth GGL1) on Euro-area data, who claim the following three results as established characteristics of NKPC:

1. The two null hypotheses $\gamma_\ell = 0$ and $\gamma_b = 0$ are rejected both individually and jointly.
2. The coefficient γ_ℓ on expected inflation exceeds the coefficient γ_b on lagged inflation substantially. The hypothesis of $\gamma_\ell + \gamma_b = 1$ is typically not rejected at conventional levels of significance, although the estimated sum is usually a little less than unity numerically.
3. When real marginal costs are proxied by the log of the wage-share, the coefficient λ in (26) is positive and significantly different from zero at conventional levels of significance.

All three claims, as well as identification, have been challenged. The inference procedures and estimation techniques used by GG and GGL1 have been criticized by Rudd and Whelan (2005, 2007), and Bårdsen, Jansen and Nymoen (2004) showed that the significance of the wage share in the GGL1 model is fragile, depending on the precise implementation of the GMM estimation used. Galí, Gertler and Lopez-Salido (2005) (GGL5) claim that their initial results remain robust to these objections, and in particular that the dominance of γ_ℓ characterizing forward-looking behavior is robust to the choice of estimation procedure and specification bias. However, on US data, Mavroeidis (2005) shows that real marginal costs appear to be irrelevant, confirming the view in Fuhrer (2006) about the difficulty of developing a sizeable coefficient on the forcing variable in the US NKPC, and for a Euro-area data set, Fanelli (2008) finds that the NKPC is a poor explanatory model. Bårdsen *et al.* (2004) (Euro area), and Bjørnstad and Nymoen (2008) (OECD panel data) demonstrate that the significance of γ_ℓ depends on omitting forcing variables that are a subset of the over-identifying instruments, revealed by the significance of χ^2_S, so the NKPC fails to parsimoniously encompass these models.

Cogley and Sbordone (2008) formulate a model where price setting firms take into account time-varying mean inflation at the macro level. The estimation results for US inflation show $\widehat{\gamma}_\ell > 1$ and $\widehat{\gamma}_b \simeq 0$. They estimate the coefficient of the wage-share to be larger than zero, but lower than in the constant-parameter version of their model. Like Cogley and Sbordone (2008), we are concerned with the consequences of non-constancies in the mean of the inflation process for the estimation of NKPC parameters, but as in Russell, Banerjee, Malki and Ponomareva (2010) (US panel data), model time-variation as intermittent unanticipated location shifts, investigating the consequences thereof for the rational (rather than subjective) expectations version of the NKPC, to which issue we now turn.

8 Euro-area NKPC estimation with IIS

As a reference, we estimate the 'pure' NKPC similar to equation (13) in GGL1, but with IV estimation instead of GMM.[3] Both Δp_{t+1} and s_t are treated as endogenous. We use the same set of instruments: five lags of inflation, two lags of s_t, detrended output and wage inflation, for $T = 102$ (1972(2) to 1998(1)):

$$\begin{array}{rcl} \widehat{\Delta p}_t & = & \underset{(0.083)}{0.925}\ \widehat{\Delta p}_{t+1} + \underset{(0.016)}{0.0142}\ s_t + \underset{(0.011)}{0.010} \\ & & \widehat{\sigma} = 0.32\% \quad \chi^2_S(9) = 14.57 \end{array} \tag{29}$$

Let $\mathsf{F}_{\mathsf{name}}$ denote an approximate F-test. Then F_{ar} tests are Lagrange-multiplier tests for autocorrelation of order k: see Godfrey (1978), and Pagan (1984) for an exposition. The heteroscedasticity test, $\mathsf{F}_{\mathsf{het}}$, computed only for OLS estimation, uses squares of the original regressors: see White (1980). Engle (1982) provides the $\mathsf{F}_{\mathsf{arch}}$ test for k^{th}-order autoregressive conditional heteroskedasticity (ARCH); $\chi^2_{\mathsf{nd}}(2)$ is the normality test in Doornik and Hansen (2008); and the χ^2_S test for the validity of instruments is from Sargan (1958). Significance of mis-specification tests at 5% and 1% are respectively denoted by * and **. The estimated γ_ℓ is less than unity, so formally a stable rational expectations solution applies for strongly exogenous s_t. But $\gamma_\ell = 1$ cannot be rejected, so stability of the rational expectations solution hinges on stationarity of s_t.

The hybrid NKPC over the same Euro-area sample is:

$$\begin{array}{rcl} \widehat{\Delta p}_t & = & \underset{(0.135)}{0.655}\ \widehat{\Delta p}_{t+1} + \underset{(0.117)}{0.280}\ \Delta p_{t-1} + \underset{(0.014)}{0.012}\ s_t + \underset{(0.010)}{0.009} \\ & & \widehat{\sigma} = 0.28\% \quad \chi^2_S(6) = 11.88 \end{array} \tag{30}$$

[3] Bårdsen *et al.* (2004) use GMM with results similar to the IV estimates here. Changes in the GMM estimation method affect the point estimates as much as the change to IV does. For example, there is a sign change in the estimated coefficient of the wage-share coefficient as a result of a change in the pre-whitening method.

The dominance of Δp_{t+1} over Δp_{t-1} is confirmed (#2 above), and the elasticities sum to 0.94. The 0.66 estimate of γ_ℓ is comparable to, and only a little lower than, the GMM estimates in Table 2 in GGL1 who report four estimates: $0.77, 0.69, 0.87$, and 0.60.

We next investigate the reduced form ('the forecasting equation'). We model Δp_t by the variables that are in the instrument set for the NKPC estimation, and then investigate structural breaks using impulse-indicator saturation in *Autometrics*. With the significance level set at 0.01, *Autometrics* finds 5 dummies with $\mathsf{F}_{\mathsf{IIS}}(5,94) = 9.53^{***}$, where *** denotes significance at 0.1%. When the hybrid NKPC is augmented by these dummies, the model is not congruent, with $\mathsf{F}_{\mathsf{ar}}(5,90) = 5.84^{***}$, $\mathsf{F}_{\mathsf{het}}(9,89) = 3.08^{**}$ and $\chi^2(8) = 18.9^{*}$. Following previous analyses (in Bårdsen *et al.*, 2004), an interpretation is that some of the variables in the instrument set have separate explanatory power for Δp_t, consistent with (earlier) standard models of inflation. Adding gap_{t-1} to the equation as an explanatory variable makes the dummy-augmented NKPC more congruent, and re-estimating yields (coefficients of dummies are multiplied by 100):

$$
\begin{aligned}
\widehat{\Delta p}_t \;=\;& \underset{(0.187)}{0.272}\ \widehat{\Delta p}_{t+1} + \underset{(0.019)}{0.048}\ s_t + \underset{(0.105)}{0.384}\ \Delta p_{t-1} + \underset{(0.014)}{0.036} + \underset{(0.0004)}{0.0009}\ gap_{t-1} \\
&+ \underset{(0.24)}{0.80}\ I_{73(1),t} + \underset{(0.29)}{0.43}\ I_{76(2),t} + \underset{(0.23)}{0.63}\ I_{76(3),t} - \underset{(0.23)}{0.66}\ I_{78(4),t} + \underset{(0.23)}{0.75}\ I_{83(1),t} \\
&\widehat{\sigma} = 0.23\% \;\; \chi^2_S(7) = 15.96^{*} \;\; \mathsf{F}_{\mathsf{ar}}(5,89) = 1.54 \\
&\mathsf{F}_{\mathsf{arch}}(4,96) = 0.14 \;\; \mathsf{F}_{\mathsf{het}}(14,84) = 1.61 \;\; \chi^2_{\mathsf{nd}}(2) = 0.94
\end{aligned}
\tag{31}
$$

The new test for invariance on adding the indicators yields $\mathsf{F}_{\mathsf{inv}}(5,94) = 7.97^{***}$, strongly rejecting. The significant dummies from the reduced form are strong evidence for lack of invariance in the expectations NKPC, and the estimated coefficient of the forward term is no longer significantly different from zero.

To check robustness, we also selected dummies at 2.5% in the reduced form, and 11 are retained. Equation (31) compares the estimates, now retaining only coefficients significant at 1%:

$$
\begin{aligned}
\widehat{\Delta p}_t \;=\;& -\underset{(0.264)}{0.298}\ \widehat{\Delta p}_{t+1} + \underset{(0.029)}{0.115}\ s_t + \underset{(0.126)}{0.505}\ \Delta p_{t-1} + \underset{(0.021)}{0.086} + \underset{(0.0004}{0.0015}\ gap_{t-1} \\
&+ \underset{(0.30)}{1.09}\ I_{73(1),t} + \underset{(0.38)}{1.09}\ I_{73(3),t} + \underset{(0.31)}{0.73}\ I_{73(4),t} + \underset{(0.34)}{0.85}\ I_{74(2),t} \\
&+ \underset{(0.33)}{0.80}\ I_{74(3),t} + \underset{(0.38)}{0.98}\ I_{76(2),t} + \underset{(0.29)}{0.57}\ I_{76(3),t} - \underset{(0.28)}{0.66}\ I_{78(4),t} + \underset{(0.28)}{0.69}\ I_{83(1),t} \\
&\widehat{\sigma} = 0.27\% \;\; \chi^2_S(5) = 5.06 \;\; \mathsf{F}_{\mathsf{ar}}(5,85) = 1.55 \\
&\mathsf{F}_{\mathsf{arch}}(4,96) = 1.49 \;\; \mathsf{F}_{\mathsf{het}}(14,80) = 1.41 \;\; \chi^2_{\mathsf{nd}}(2) = 1.04
\end{aligned}
\tag{32}
$$

$\mathsf{F}_{\mathsf{inv}}(11,88) = 4.87^{***}$, and nine of the 11 reduced form dummies are retained, with the estimated coefficient of the forward term now negative, and estimated persistence is substantively lower. The coefficient of the wage-share has become sizeable, allowing the wage-share to serve as a statistically and substantively important equilibrating mechanism.

Re-estimation of the augmented model on the shorter sample commencing after the breaks in (31), starting in 1983(2) yields consistent outcomes, with the estimated coefficient of the expectations term of 0.082, confirming that its significance in (29) and (30) is as a proxy for unmodeled shifts. All these findings match the theoretical and simulation results above.

9 US NKPC estimation with IIS

The pure NKPC on the same sample period 1960(1)–1997(4) used by GG, with their instruments, but using IV instead of GMM gives for $T = 152$:

$$\widehat{\Delta p}_t = \underset{(0.048)}{0.992}\ \Delta p_{t+1} + \underset{(0.018)}{0.011}\ s_t + \underset{(0.00052)}{5.12e^{-5}}$$
$$\widehat{\sigma} = 0.20\%\ \ \chi^2_S(8) = 17.7^{*} \tag{33}$$

comparable to the GMM estimates at the top of page 207 in GG, which are $0.95(0.045)$ and $0.023(0.012)$. GG's equation is without an intercept, and that is also near zero in (33). Without the intercept, the standard error of the wage-share is reduced to 0.011, the point estimates being unaffected.[4] GGL5 compare Euro and US results, where the pure NKPC is reported as having $\widehat{\lambda} = 0.25$, so the significance and magnitude of the wage-share coefficient depends on 'technicalities' in NKPC estimation. The hybrid NKPC for the US is:

$$\widehat{\Delta p}_t = \underset{(0.092)}{0.623}\ \Delta p_{t+1} + \underset{(0.081)}{0.357}\ \Delta p_{t-1} + \underset{(0.014)}{0.014}\ s_t + \underset{(0.00041)}{0.00016}$$
$$\widehat{\sigma} = 0.23\%\ \ \chi^2_S(7) = 7.60 \tag{34}$$

The estimates of γ_b and γ_ℓ are similar to the Euro-area hybrid in equation (30), and they are representative of the GMM estimates found in Table 2 in GG. $\widehat{\gamma}_\ell$ dominates, and they sum almost to unity, so #1 and #2 are confirmed by the estimation. Sargan's χ^2_S test which is significant in (33), is insignificant in (34), evidence that Δp_{t-1} is misplaced as an instrument and belongs to the category of explanatory variables.

Autometrics finds 9 impulse dummies in the reduced form at a 0.01 significance level. When added to (34), those indicators are significant with $\mathsf{F}_{\mathsf{inv}}(9, 137) = 7.61^{***}$, again strongly rejecting, and the diagnostics improve, except for the residual autocorrelation, which is still highly significant. It was not straightforward to find a congruent model from this information set, but moving Δp_{t-2} and gap_{t-1} from being instruments to explanatory variable helps (the autocorrelation and heteroskedasticity statistics now have significance levels of 0.023 and 0.013). Estimation of the augmented hybrid US NKPC yields (coefficients of dummies are multiplied by 100):

$$\widehat{\Delta p}_t = \underset{(0.168)}{0.253}\ \Delta p_{t+1} + \underset{(0.083)}{0.502}\ \Delta p_{t-1} + \underset{(0.085)}{0.196}\ \Delta p_{t-3} + \underset{(0.013)}{0.022}\ s_t + \underset{(0.018)}{0.028}\ gap_{t-1}$$
$$+ \underset{(0.00041)}{0.00032} + \underset{(0.18)}{0.51}\ I_{63(4),t} + \underset{(0.19)}{0.66}\ I_{72(1),t} - \underset{(0.19)}{0.62}\ I_{72(2),t} + \underset{(0.24)}{0.73}\ I_{74(3),t}$$
$$- \underset{(0.20)}{0.63}\ I_{75(2),t} + \underset{(0.21)}{0.44}\ I_{76(4),t} + \underset{(0.18)}{0.59}\ I_{77(4),t} + \underset{(0.19)}{0.46}\ I_{78(2),t} - \underset{(0.20)}{0.44}\ I_{81(2),t}$$
$$\widehat{\sigma} = 0.18\%\ \ \chi^2_S(5) = 3.87\ \ \mathsf{F}_{\mathsf{ar}}(5, 132) = 2.71^{*}$$
$$\mathsf{F}_{\mathsf{arch}}(4, 144) = 0.79\ \ \mathsf{F}_{\mathsf{het}}(20, 122) = 1.93^{*}\ \ \chi^2_{\mathsf{nd}}(2) = 1.04 \tag{35}$$

All the shift-dummies from the reduced form are statistically significant at the 5% level (and most at lower levels). The estimate of the feed-forward term has been reduced from 0.62 to

[4] The mean of s_t is not exactly zero over the sample period, despite being described as 'a deviation from steady-state' in the text.

0.25, so the t-value is just 1.5. However, considerable persistence remains. The coefficient of the wage-share improves: compared to (34), the point estimate has increased somewhat, and the standard error has been reduced in (35). When the 'post-break' sample 1981(3) to 1997(4) is used to estimate the augmented model, the results are similar to (35): $\widehat{\gamma}_\ell = 0.28$, t-value 0.94.

Estimating the hybrid form over the whole available sample 1948(2)–1998(1) yields:

$$\widehat{\Delta p_t} = \underset{(0.149)}{0.854}\ \Delta p_{t+1} + \underset{(0.101)}{0.175}\ \Delta p_{t-1} + \underset{(0.023)}{0.021}\ s_t - \underset{(0.0007)}{0.0003}$$
$$\widehat{\sigma} = 0.46\% \quad \chi^2_S(8) = 14.0$$

but every mis-specification test is significant at 0.1% level. IIS in the reduced form at 1% delivered 18 indicators (the additional ones being mainly 1940s & 1950s) with $\widehat{\sigma} = 0.22\%$ and no significant mis-specification tests. Then the augmented feed-forward model yielded:

$$\begin{aligned}\widehat{\Delta p_t} = {} & \underset{(0.071)}{0.373}\ \Delta p_{t+1} + \underset{(0.062)}{0.306}\ \Delta p_{t-1} + \underset{(0.035)}{0.224}\ \Delta p_{t-3} + \underset{(0.012)}{0.017}\ s_t \\ & + \underset{(0.010)}{0.029}\ gap_{t-1} + \underset{(0.0006)}{0.0006} + \underset{(0.031)}{0.059}\ \Delta w_{t-3} + \text{indicators} \\ & \widehat{\sigma} = 0.22\% \quad \chi^2_S(7) = 11.7 \quad \mathsf{F}_{\mathsf{ar}}(5, 171) = 2.27^{*} \\ & \mathsf{F}_{\mathsf{arch}}(4, 192) = 2.98^{*} \quad \mathsf{F}_{\mathsf{het}}(12, 169) = 1.65 \quad \chi^2_{\mathsf{nd}}(2) = 4.64\end{aligned} \tag{36}$$

These extended results are similar to those above, with $\mathsf{F}_{\mathsf{inv}}(18, 176) = 25.1^{***}$. The fitted and actual values, residuals, residual density and residual correlogram are shown in figure 1.

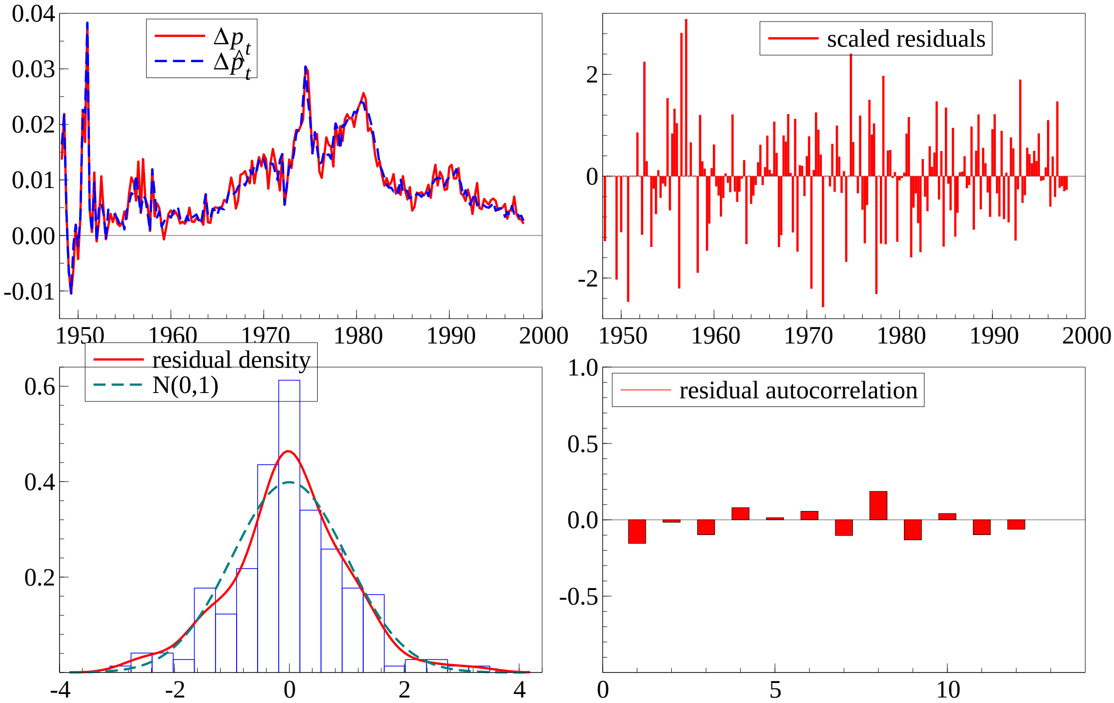

Figure 1: Graphical outcomes

10 Conclusion

Many economic models, such as the new-Keynesian Phillips curve (NKPC), include expected future values to explain current outcomes. Models of this type are regularly estimated by replacing the expected value by the actual future outcome, then using Instrumental Variables (IV) or Generalized Method of Moments (GMM) methods to estimate the parameters. However, the underlying theory does not allow for various forms of non-stationarity–despite the fact that crises, breaks and regimes shifts are relatively common. We demonstrated that potentially spurious outcomes can arise when breaks are not modeled and expectations are in fact irrelevant.

We proposed a test for the invariance of the parameters of such expectations-based formulations using a 2-stage procedure. The first stage applies impulse-indicator saturation (IIS) to the reduced form to detect the presence of any unmodeled outliers or location shifts; and the second is an F-test of their presence in the structural equation. A tight first-stage significance criterion is used to control the second-stage rejection frequency under the null that the structural equation is correctly specified by including the future variable as an approximation to the expected value.

Applying the resulting methods to two salient empirical studies of Euro-area and US NKPCs radically alters previous results. In the former, the future variable had an insignificant coefficient; and in the latter, its value was more than halved. The added indicators were highly significant in both cases and rejected invariance.

References

Bank of England (2009). *February Inflation Report*. London: Bank of England Monetary Policy Committee.

Bårdsen, G., Jansen, E. S., and Nymoen, R. (2004). Econometric evaluation of the new Keynesian Phillips curve. *Oxford Bulletin of Economics and Statistics*, **66**, 671–686.

Barrell, R. (2001). Forecasting the world economy. In Hendry, D. F., and Ericsson, N. R. (eds.), *Understanding Economic Forecasts*, pp. 149–169. Cambridge, Mass.: MIT Press.

Bjørnstad, R., and Nymoen, R. (2008). The New Keynesian Phillips curve tested on OECD panel data. *Economics-The Open-Access, Open-Assessment E-Journal*, **2**, 1–18. http://www.economics-ejournal.org.

Castle, J. L. (2008). Checking the robustness and validity of model selection: An application to UK wages. Mimeo, Economics Department, Oxford University.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2011a). 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. (2011b). Model selection when there are multiple breaks. *Journal of Econometrics*, forthcoming.

Castle, J. L., and Shephard, N. (eds.)(2009). *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. Oxford: Oxford University Press.

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

Cogley, T., and Sbordone, A. M. (2008). Trend inflation, indexation, and inflation persistence in the new Keynesian Phillips curve. *American Economic Review*, **98**, 2101–2126.

Doornik, J. A. (2007). *Object-Oriented Matrix Programming using Ox* 6th edn. London: Timberlake Consultants Press.

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

Doornik, J. A., and Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, **70**, 927–939.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1007.

Engle, R. F., and Hendry, D. F. (1993). Testing super exogeneity and invariance in regression models. *Journal of Econometrics*, **56**, 119–139. Reprinted in Ericsson, N. R. and Irons, J. S. (eds.) *Testing Exogeneity*, Oxford: Oxford University Press, 1994.

Evans, G. W., and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.

Falch, N. S., and Nymoen, R. (2011). The accuracy of a forecast targeting central bank. *Economics-The Open-Access, Open-Assessment E-Journal*, **5**, 2011–15. http://www.economics-ejournal.org.

Fanelli, L. (2008). Testing the new Keynesian Phillips curve through vector autoregressive models: Results from the Euro area. *Oxford Bulletin of Economics and Statistics*, **70**, 53—66.

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

Frydman, R., and Goldberg, M. D. (eds.)(2007). *Imperfect Knowledge Economics: Exchange Rates and Risk*. Princeton, New Jersey: Princeton University Press.

Fuhrer, J. C. (2006). Intrinsic and inherited inflation persistence. *International Journal of Central Banking*, **2**, 49–86.

Galí, J., and Gertler, M. (1999). Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, **44**, 195–222.

Galí, J., Gertler, M., and Lopez-Salido, J. D. (2001). European inflation dynamics. *European Economic Review*, **45**, 1237–1270.

Galí, J., Gertler, M., and Lopez-Salido, J. D. (2005). Robustness of the estimates of the hybrid new Keynesian Phillips curve. *Journal of Monetary Economics*, **21**, 1107–1118.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, **46**, 1303–1313.

Hendry, D. F. (1988). The encompassing implications of feedback versus feedforward mechanisms in econometrics. *Oxford Economic Papers*, **40**, 132–149. Reprinted in Ericsson, N. R. and Irons, J. S. (eds.) *Testing Exogeneity*, Oxford: Oxford University Press, 1994; and in Campos, J., Ericsson, N.R. and Hendry, D.F. (eds.), *General to Specific Modelling*. Edward Elgar, 2005.

Hendry, D. F. (2000). On detectable and non-detectable structural change. *Structural Change and Economic Dynamics*, **11**, 45–65. Reprinted in *The Economics of Structural Change*, Hagemann, H. Landesman, M. and Scazzieri, R. (eds.), Edward Elgar, Cheltenham, 2002.

Hendry, D. F. (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics*, **135**, 399–426.

Hendry, D. F. (2011). On adding over-identifying instrumental variables to simultaneous equations. *Economics Letters*, **111**, 68–70.

Hendry, D. F., and Johansen, S. (2011). The properties of model selection when retaining theory variables. Unpublished paper, Economics Department, University of Oxford.

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. (2010). On the mathematical basis of inter-temporal optimization. Discussion paper 497, Economics Department, Oxford.

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.

Juselius, K. (2006). *The Cointegrated VAR Model: Methodology and Applications*. Oxford: Oxford University Press.

Kirman, A. (1989). The intrinsic limits of economic theory: The emperor has no clothes. *Economic Journal*, **99**, 126–139.

Mavroeidis, S. (2004). Weak identification of forward-looking models in monetary economics. *Oxford Bulletin of Economics and Statistics*, **66**, 609–635.

Mavroeidis, S. (2005). Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve. *Journal of Money, Credit, and Banking*, **37**, 421–448.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, **29**, 315–335.

Nickell, S. J. (2008). Aspects of monetary policy in the UK. Mimeo, Nuffield College, Oxford.

Pagan, A. R. (1984). Model evaluation by variable addition. In Hendry, D. F., and Wallis, K. F. (eds.), *Econometrics and Quantitative Economics*, pp. 103–135. Oxford: Basil Blackwell.

Pesaran, M. H. (1981). Identification of rational expectations models. *Journal of Econometrics*, **16**, 375–398.

Rudd, J., and Whelan, K. (2005). Does labor's share drive inflation?. *Journal of Credit and Banking*, **37**, 297–312.

Rudd, J., and Whelan, K. (2006). Can rational expectations sticky-price models explain inflation dynamics?. *American Economic Review*, **96**, 303–320.

Rudd, J., and Whelan, K. (2007). Modeling inflation dynamics: A critical review of recent work. *Journal of Credit and Banking*, **39**, 155–170.

Russell, B., Banerjee, A., Malki, I., and Ponomareva, N. (2010). A multiple break panel approach to estimating United States Phillips curves. Unpublished paper, Economic Studies, University of Dundee, Dundee, UK.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, **26**, 393–415.

Smets, F., and Wouters, R. (2002). An estimated stochastic dynamic general equilibrium model of the Euro Area. ECB Working Paper, 171, European Central Bank, Frankfurt.

Stock, J. H., and Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics*, **14**, 11–30.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.

11 Appendix calculations for the hybrid model

To obtain the reduced form parameterization, first set $\psi = 0$ in (9) and solve for the constant parameter reduced form:

$$y_t = \rho_0 + \rho_1 y_{t-1} + \varphi_0 z_t + \varphi_1 z_{t-1} + u_t \tag{37}$$

where the location shift, $d_{T_1,T_2,t}$ will be added to (37) when $\psi \neq 0$. Then:

$$y_{t+1} = \rho_0 + \rho_1 y_t + \varphi_0 z_{t+1} + \varphi_1 z_t + u_{t+1} \tag{38}$$

and hence:

$$\begin{aligned} y_{t+1} &= \rho_0 + \rho_1 \left(\rho_0 + \rho_1 y_{t-1} + \varphi_0 z_t + \varphi_1 z_{t-1} + u_t\right) \\ &\quad + \varphi_0 \left(\lambda_0 + \lambda_1 z_t + \lambda_2 z_{t-1} + \eta_{t+1}\right) + \varphi_1 z_t + u_{t+1} \\ &= \rho_1^2 y_{t-1} + \left(\rho_0 \left(1 + \rho_1\right) + \varphi_0 \lambda_0\right) + \left(\varphi_0 \left(\rho_1 + \lambda_1\right) + \varphi_1\right) z_t \\ &\quad + \left(\rho_1 \varphi_1 + \varphi_0 \lambda_2\right) z_{t-1} + \varphi_0 \eta_{t+1} + u_{t+1} + \rho_1 u_t \end{aligned} \tag{39}$$

Taking expectations:

$$\begin{aligned} \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} | z_t, \mathcal{I}_{t-1}\right] &= \mathsf{E}_{\mathsf{f}_t}\left[\rho_1^2 y_{t-1} + \left(\rho_0 \left(1 + \rho_1\right) + \varphi_0 \lambda_0\right) + \left(\varphi_0 \left(\rho_1 + \lambda_1\right) + \varphi_1\right) z_t \right. \\ &\qquad \left. + \left(\rho_1 \varphi_1 + \varphi_0 \lambda_2\right) z_{t-1} | z_t, \mathcal{I}_{t-1}\right] \\ &= \rho_1^2 y_{t-1} + \left(\rho_0 \left(1 + \rho_1\right) + \varphi_0 \lambda_0\right) \\ &\quad + \left(\varphi_0 \left(\rho_1 + \lambda_1\right) + \varphi_1\right) z_t + \left(\rho_1 \varphi_1 + \varphi_0 \lambda_2\right) z_{t-1} \end{aligned} \tag{40}$$

Using $y_{t+1}^e = \mathsf{E}_{\mathsf{f}_t}\left[y_{t+1} | z_t, \mathcal{I}_{t-1}\right]$ and substituting (40) in (9):

$$\begin{aligned} y_t &= \beta_1 \left(\rho_1^2 y_{t-1} + \left(\rho_0 \left(1 + \rho_1\right) + \varphi_0 \lambda_0\right) + \varphi_0 \left(\rho_1 + \lambda_1\right) + \varphi_1\right) z_t \\ &\quad + \left(\rho_1 \varphi_1 + \varphi_0 \lambda_2\right) z_{t-1} + \beta_1 y_{t-1} + \beta_3 z_t + \epsilon_t \\ &= \left(\beta_1 \rho_1^2 + \beta_2\right) y_{t-1} + \beta_1 \left(\rho_0 \left(1 + \rho_1\right) + \varphi_0 \lambda_0\right) \\ &\quad + \left(\beta_1 \varphi_0 \left(\rho_1 + \lambda_1\right) + \beta_1 \varphi_1 + \beta_3\right) z_t + \beta_1 \left(\rho_1 \varphi_1 + \varphi_0 \lambda_2\right) z_{t-1} + \epsilon_t \end{aligned} \tag{41}$$

Comparing coefficients in (37) and (41) using $1 - \beta_1 \rho_1 = \beta_1 \rho_2$, leads to the following set of restrictions:

$$\rho_0 = \varphi_0 \frac{\lambda_0}{(\rho_2 - 1)}$$

$$\rho_1 = \frac{(1 - \sqrt{1 - 4\beta_1\beta_2})}{2\beta_1} \qquad \rho_2 = \frac{(1 + \sqrt{1 - 4\beta_1\beta_2})}{2\beta_1}$$

$$\varphi_0 = \frac{\beta_3}{\beta_1}\left(\rho_2 - \lambda_1 - \lambda_2\rho_2^{-1}\right)^{-1} \qquad \varphi_1 = \varphi_0 \frac{\lambda_2}{\rho_2}.$$

The difference between y_{t+1} and $\mathsf{E}_{\mathsf{f}_t}\left[y_{t+1}|z_t, \mathcal{I}_{t-1}\right]$ is:

$$\varphi_0\eta_{t+1} + u_{t+1} + \rho_1 u_t \tag{42}$$

which has a variance:

$$\sigma_e^2 = \varphi_0^2\sigma_\eta^2 + \left(1 + \rho_1^2\right)\sigma_\epsilon^2 \tag{43}$$

as against σ_ϵ^2 when $\mathsf{E}_{\mathsf{f}_t}\left[y_{t+1}|z_t, \mathcal{I}_{t-1}\right]$ is known. The coefficient in (9) is β_1 so:

$$y_t = \beta_1 y_{t+1} + \beta_2 y_{t-1} + \beta_3 z_t + \epsilon_t - \beta_1\left(\varphi_0\eta_{t+1} + \epsilon_{t+1} + \rho_1\epsilon_t\right) \tag{44}$$

so the error variance is:

$$\sigma_\nu^2 = \sigma_\epsilon^2 + \beta_1^2\left(\varphi_0^2\sigma_\eta^2 + \left(1 + \rho_1^2\right)\sigma_\epsilon^2\right) - 2\beta_1\rho_1\sigma_\epsilon^2 \tag{45}$$