Research in Economics 70 (2016) 262–280

Contents lists available at ScienceDirect

Research in Economics

journal homepage: www.elsevier.com/locate/rie

Counterfactual analysis in macroeconometrics: An empirical investigation into the effects of quantitative easing

M. Hashem Pesaran [a,b], Ron P. Smith [c,*]

[a] USC Dornsife INET, University of Southern California, United States
[b] Trinity College, Cambridge, United Kingdom
[c] Birkbeck, University of London, United Kingdom

ARTICLE INFO

Article history:
Received 8 October 2015
Accepted 19 January 2016
Available online 25 January 2016

Keywords:
Counterfactuals
Policy analysis
Policy ineffectiveness test
Macroeconomics
Quantitative easing (QE)

ABSTRACT

This paper develops tests of policy ineffectiveness when counterfactual outcomes, namely the predicted values of the target variable in the absence of a policy change, are obtained using reduced form or final form ‘policy response equations’. The policy response equation explains the target variable by lags of itself and by current and lagged values of the policy variable and policy-invariant exogenous variables. These tests complement those in Pesaran and Smith (2015), which are done in the context of complete systems of macroeconometric dynamic stochastic general equilibrium (DSGE) under rational expectations (RE). While there would be efficiency gains from using the complete system, were it fully known and correctly specified, this is rarely the case. Instead, since the counterfactual is a type of forecast and parsimonious models tend to forecast better than complex ones, we may obtain more reliable estimates of the counterfactual outcomes from a parsimonious policy response equation. We consider two types of policy intervention, a discretionary policy change where there is a deterministic change to the policy variable, and a rule-based policy change where one or more parameters of a stochastic policy rule are changed. We examine the asymptotic distributions and power of the tests under various assumptions. The proposed test for a discretionary policy change is illustrated with an application to the unconventional monetary policy known as quantitative easing (QE) adopted in the UK.

1. Introduction

This paper develops tests for the null hypothesis of the ineffectiveness of a policy intervention where the counterfactual, the predicted outcome for the target variable in the absence of a policy change, is provided by a parsimonious reduced form or final form “policy response equation”. It complements Pesaran and Smith (2015, PS), which considers tests of policy ineffectiveness in the context of complete macroeconometric dynamic stochastic general equilibrium (DSGE) systems under rational expectations (RE). The policy response equation, which can be derived from the complete system considered in PS, explains the target variable in terms of its lagged values; the policy variable, which may be endogenous or exogenous, and, if available, other exogenous variables and their lagged values that are invariant to the policy intervention.[1] We consider two

* Corresponding author.

[1] We assume that one can distinguish between those variables that are influenced by policy, and those that are invariant to policy, namely they remain unaffected by the policy change.

http://dx.doi.org/10.1016/j.rie.2016.01.004
1090-9443/

types of policy intervention: discretionary, where there is a deterministic change to the policy variable, and rule-based, where one of more parameters of a stochastic policy rule are changed.

In the case where the underlying model is static, the policy response equation is the reduced form model for the target variable, but in the dynamic case the reduced form also involves lagged values of the other endogenous variables. To obtain the final form equation, the lagged values of these other endogenous variables can be solved out, following the seminal work of Zellner and Palm (1974). This yields an infinite order distributed lag function relating the target variable to the policy and policy-invariant variables (if any). Assuming that the target variable adjusts to policy changes sufficiently fast, the lag orders can be truncated and estimated using standard time series techniques such as Akaike or Schwarz information criteria. Alternatively, following Berk (1974), the lag orders can be set to $T^{1/3}$, where T is the pre-intervention sample size.

To test for the effect of a policy change, we require (i) a model to construct counterfactuals for the target variable in the absence of the policy intervention and (ii) a way to determine whether the difference between the realized and counterfactual outcomes is larger than would have been expected by chance. When the structural model is fully known and correctly specified, using the complete model and imposing all the cross-equation restrictions implied by the DSGE model will yield more efficient estimates of the counterfactual outcomes than using reduced form or final form specifications. However, not only are we rarely certain of the correct specification of individual equations, we often have little knowledge of the subset of the large number of possible variables to include in the full model. In such circumstances, since the counterfactual outcome is a type of forecast, and all the evidence is that parsimonious models forecast better, using a parsimonious policy response equation may produce more reliable estimates of the counterfactual outcomes than a possibly misspecified large model. Accordingly in this paper, we propose tests for policy ineffectiveness based on single equation reduced or final forms and derive their asymptotic distributions both when the post-intervention sample is fixed, as the pre-intervention sample expands, and when both samples rise jointly but at different rates. We also investigate the power of the proposed tests.

A policy intervention can take a variety of forms. The intervention might involve changes to the parameters of either an endogenous policy rule, like the Taylor rule, or an exogenous rule, like a fixed money supply growth rule. Typically, these policy rules are stochastic, in the sense that the policy maker cannot necessarily control the realizations of the policy instrument exactly. Alternatively, the intervention might involve a discretionary deterministic change in the setting of the policy instrument or a shock to the error of a policy rule equation, as is often considered in generating impulse response functions. We will focus on policy interventions that can be characterized either as changes to the parameters of a policy rule or as discretionary policy changes. We note, however, that the testing procedure also applies when the policy intervention is defined in terms of shocks, as argued in PS.

The proposed policy ineffectiveness tests are based on the differences between the post-intervention realizations of the policy target(s) and associated counterfactual outcomes based on the parameters of a model estimated using data before the policy intervention. These differences are computed and averaged over a given policy evaluation horizon. The Lucas Critique is not an issue since the counterfactuals (obtained under the null of no policy change) are computed using pre-intervention parameter estimates, while the effects of policy change, either through changes in parameters or expectations, show up in the realized post-intervention outcomes. *The computation of the proposed tests does not require knowing the post-intervention parameters*.

As an illustration, we employ the policy ineffectiveness tests to investigate the effects of the quantitative easing (QE) introduced in the UK after March 2009. To construct the test, we employ an autoregressive distributed lag (ARDL) equation in the target variable, output growth, the policy variable, the spread between long and short rates, and US and euro area output variables, that we assume to be invariant to the policy change. We exclude other endogenous variables, that could be influenced by the policy. For instance, it would be wrong to include the exchange rate in the equation, because if QE was effective in reducing the spread then the exchange rate would almost certainly have been changed by it and we would have needed to allow for that effect by considering a separate equation that links the exchange rate to QE. By excluding the exchange rate from the policy response equation, we are in effect replacing the exchange rate (which is endogenously determined with the spread) by its determinants. The same argument also applies to any other endogenous variable which is affected by the policy change. This is the reverse of the usual misspecification argument, since we wish to attribute to policy the effects that are transmitted through the other endogenous variables.[2] Following the literature, we treat QE as a discretionary intervention which caused a deterministic 100 basis points reduction in the spread. This policy change has a positive impact effect on output growth of about one percentage point, but the policy impact is very quickly reversed.

The rest of the paper is organized as follows: Section 2 sets up a DSGE model with exogenous variables following PS and derives its solution which is the basis for the single equation policy response functions considered in the rest of the paper. Section 3 provides the framework for the policy ineffectiveness tests. Section 4 develops the test for the static model, where the reduced form equation is used to provide the counterfactual, under both discretionary and rule-based interventions and examines the power of the test. Section 5 extends the test to the dynamic case, where the final form equation is used to derive the counterfactual outcomes. Section 6 considers the empirical application, and Section 7 ends with some concluding remarks. The more technical derivations are given in the Appendix.

[2] A similar argument is developed in a continuous time regression context in Pesaran and Smith (2014).

2. Derivation of the policy response equation

Following Pesaran and Smith (2015) we begin from a canonical rational expectations (RE) model in the $k_y \times 1$ vector $\mathbf{y}_t$ of endogenous variables, determined in terms of their expected future values, past values, and a $k_s \times 1$ vector of exogenous variables, $\mathbf{s}_t$,

$$\mathbf{A}_0(\boldsymbol{\varphi})\mathbf{y}_t = \mathbf{A}_1(\boldsymbol{\varphi})E_t(\mathbf{y}_{t+1}) + \mathbf{A}_2(\boldsymbol{\varphi})\mathbf{y}_{t-1} + \mathbf{A}_3(\boldsymbol{\varphi})\mathbf{s}_t + \mathbf{u}_t, \tag{1}$$

where $E_t(\mathbf{y}_{t+1}) = E(\mathbf{y}_{t+1}|\Im_t)$, $\Im_t = (\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots; \mathbf{s}_t, \mathbf{s}_{t-1}, \ldots)$ is the information set, $\mathbf{A}_i(\boldsymbol{\varphi})$, for i=0,1,2, and $\mathbf{A}_3(\boldsymbol{\varphi})$ are, respectively, $k_y \times k_y$ and $k_y \times k_s$ coefficient matrices, $\mathbf{A}_0(\boldsymbol{\varphi})$ is non-singular, $\boldsymbol{\varphi}$ is a vector of structural or "deep" parameters, and $\mathbf{u}_t$ is a $k_y \times 1$ vector of structural shocks. The exogenous variables are assumed to follow the VAR(1) model

$$\mathbf{s}_t = \mathbf{a}(\boldsymbol{\rho}) + \mathbf{R}(\boldsymbol{\rho})\mathbf{s}_{t-1} + \boldsymbol{\eta}_t, \tag{2}$$

where $\mathbf{a}(\boldsymbol{\rho})$ is a $k_s \times 1$ vector of intercepts and $\mathbf{R}(\boldsymbol{\rho})$ is the $k_s \times k_s$ matrix of coefficients that depend on a vector of unknown coefficients, $\boldsymbol{\rho}$. The errors, $\mathbf{u}_t$ and $\boldsymbol{\eta}_t$, are assumed to be serially and cross sectionally uncorrelated, with zero means and constant variances, $\boldsymbol{\Sigma}_u$, and $\boldsymbol{\Sigma}_\eta$, respectively, and zero covariances, $\boldsymbol{\Sigma}_{\eta u} = \mathbf{0}$.

The above specification is sufficiently general and cover both DSGE rational expectations models, and simultaneous equation systems (when $\mathbf{A}_1(\boldsymbol{\varphi}) = \mathbf{0}$). It also allows for a distinction between endogenous and exogenous policy changes. An endogenous policy change, such as a change in the Taylor rule, changes one or more elements of $\boldsymbol{\varphi}$, while an exogenous policy change, such as changes to a fixed money supply rule or changes to the steady state inflation target, changes one or more elements of $\boldsymbol{\rho}$.

We distinguish between exogenous policy and non-policy variables and partition $\mathbf{s}_t$ as $\mathbf{s}_t = (\mathbf{x}_t', \tilde{\mathbf{w}}_t')'$, where $\mathbf{x}_t$ is a $k_x \times 1$ vector of exogenous policy variables, and $\tilde{\mathbf{w}}_t$ is a $k_{\tilde{w}} \times 1$ vector of non-policy exogenous variables that are invariant to policy interventions.

If the quadratic matrix equation,

$$\mathbf{A}_1(\boldsymbol{\varphi})\boldsymbol{\Phi}^2(\boldsymbol{\varphi}) - \mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi}) + \mathbf{A}_2(\boldsymbol{\varphi}) = \mathbf{0}, \tag{3}$$

has a solution, $\boldsymbol{\Phi}(\boldsymbol{\varphi})$, with all its eigenvalues inside the unit circle, then, as shown in PS, the RE model, (1) and (2), has the unique solution

$$\mathbf{y}_t = \boldsymbol{\Phi}(\boldsymbol{\varphi})\mathbf{y}_{t-1} + \boldsymbol{\Psi}_s(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{s}_t + \boldsymbol{\mu}_a(\boldsymbol{\varphi},\rho) + \boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{u}_t, \tag{4}$$

where

$$\boldsymbol{\Psi}_s(\boldsymbol{\varphi},\boldsymbol{\rho}) = \sum_{h=1}^{\infty} \mathbf{F}^h(\boldsymbol{\varphi})\boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{A}_3(\boldsymbol{\varphi})\mathbf{R}^h(\boldsymbol{\rho}), \tag{5}$$

$$\boldsymbol{\mu}_a(\boldsymbol{\varphi},\boldsymbol{\rho}) = \sum_{h=1}^{\infty} \mathbf{F}^h(\boldsymbol{\varphi})\boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{A}_3(\boldsymbol{\varphi})\left[\mathbf{I}_{k_s} - \mathbf{R}(\boldsymbol{\rho})\right]^{-1}\left[\mathbf{I}_{k_s} - \mathbf{R}(\boldsymbol{\rho})^h\right]\mathbf{a}(\boldsymbol{\rho}), \tag{6}$$

$$\boldsymbol{\Gamma}(\boldsymbol{\varphi}) = \left[\mathbf{A}_0(\boldsymbol{\varphi}) - \mathbf{A}_1(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi})\right]^{-1}, \tag{7}$$

and $\mathbf{F}(\boldsymbol{\varphi}) = \boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{A}_1(\boldsymbol{\varphi})$. Partitioning $\boldsymbol{\Psi}_s(\boldsymbol{\varphi},\boldsymbol{\rho}) = \left[\boldsymbol{\Psi}_x(\boldsymbol{\varphi},\boldsymbol{\rho}), \boldsymbol{\Psi}_{\tilde{w}}(\boldsymbol{\varphi},\boldsymbol{\rho})\right]$ conformably with the partitioning of $\mathbf{s}_t = (\mathbf{x}_t', \tilde{\mathbf{w}}_t')'$ we have

$$\mathbf{y}_t = \boldsymbol{\Phi}(\boldsymbol{\varphi})\mathbf{y}_{t-1} + \boldsymbol{\Psi}_x(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{x}_t + \boldsymbol{\Psi}_{\tilde{w}}(\boldsymbol{\varphi},\boldsymbol{\rho})\tilde{\mathbf{w}}_t + \boldsymbol{\mu}_a(\boldsymbol{\varphi},\boldsymbol{\rho}) + \boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{u}_t,$$

which we can write as the reduced form of a standard simultaneous equations model

$$\mathbf{y}_t = \boldsymbol{\Phi}(\boldsymbol{\varphi})\mathbf{y}_{t-1} + \boldsymbol{\Psi}_x(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{x}_t + \boldsymbol{\Psi}_w(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{w}_t + \boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{u}_t, \tag{8}$$

where $\mathbf{w}_t = (\tilde{\mathbf{w}}_t', 1)'$, and $\boldsymbol{\Psi}_w(\boldsymbol{\varphi},\boldsymbol{\rho}) = \left[\boldsymbol{\Psi}_{\tilde{w}}(\boldsymbol{\varphi},\boldsymbol{\rho}), \boldsymbol{\mu}_a(\boldsymbol{\varphi},\boldsymbol{\rho})\right]$. Whenever $\mathbf{a}(\boldsymbol{\rho} \neq \mathbf{0})$ then $\tilde{\mathbf{w}}_t$ has a non-zero mean. When there are no exogenous variables, the solution (8) corresponds to a vector autoregression and each endogenous variable depends on the lagged values of all the other endogenous variables.

By premultiplying both sides of (8) by $\mathbf{A}_0(\boldsymbol{\varphi})$, we can convert it to the structural form of a simultaneous equations model:

$$\mathbf{A}_0(\boldsymbol{\varphi})\mathbf{y}_t = \mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi})\mathbf{y}_{t-1} + \mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Psi}_x(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{x}_t + \mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Psi}_w(\boldsymbol{\varphi},\boldsymbol{\rho})\mathbf{w}_t + \mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{u}_t. \tag{9}$$

It is also worth noting that since

$$\mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Gamma}(\boldsymbol{\varphi}) = \mathbf{A}_0(\boldsymbol{\varphi})\left[\mathbf{A}_0(\boldsymbol{\varphi}) - \mathbf{A}_1(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi})\right]^{-1} = \left[\mathbf{I}_{k_y} - \mathbf{A}_1(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi})\mathbf{A}_0^{-1}(\boldsymbol{\varphi})\right]^{-1},$$

then (9) reduces to the identity matrix, $\mathbf{I}_{k_y}$ if $\boldsymbol{\Phi}(\boldsymbol{\varphi}) = \mathbf{0}$, namely if there are no dynamics.

Assume that the target variable of interest is a scalar, the first element of $\mathbf{y}_t$, and partition $\mathbf{y}_t$ as $\mathbf{y}_t = (y_{1t}, \mathbf{y}_{2t}')'$, with y_{1t} being the target variable, $\mathbf{y}_{2t}$ being the vector of remaining endogenous variables. Setting the first element of the first row of

$\mathbf{A}_0(\boldsymbol{\varphi})$ equal to 1 and collecting the remaining elements of the first row in $-\boldsymbol{\alpha}$, then the first equation of (9) can be written

$$y_{1t} = \boldsymbol{\alpha}'\mathbf{y}_{2t} + \phi y_{1,t-1} + \boldsymbol{\phi}_2'\mathbf{y}_{2,t-1} + \boldsymbol{\psi}_x'\mathbf{x}_t + \boldsymbol{\psi}_w'\mathbf{w}_t + v_{1t}, \tag{10}$$

where the first rows of $\mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Phi}(\boldsymbol{\varphi})$, $\mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Psi}_x(\boldsymbol{\varphi},\boldsymbol{\rho})$, and $\mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Psi}_w(\boldsymbol{\varphi},\boldsymbol{\rho})$, are denoted by $(\phi, \boldsymbol{\phi}_2')$, $\boldsymbol{\psi}_x'$, and $\boldsymbol{\psi}_w'$ respectively, and v_{1t} is the first element of $\mathbf{A}_0(\boldsymbol{\varphi})\boldsymbol{\Gamma}(\boldsymbol{\varphi})\mathbf{u}_t$.

The parameters of (10) are not themselves the object of interest. For testing the policy ineffectiveness hypothesis we only require to construct counterfactual outcomes for the target variable, y_{1t}, conditional on there not having been a change of policy. All the evidence is that simple parsimonious models forecast better than more complicated models like (10), which is likely to include a potentially very large number of other endogenous variables and their lags. In addition, were we to use (10), we would also have to specify the effect of any policy changes on $\mathbf{y}_{2t}$. When (1) and (3) are correctly specified, system estimation that imposes all the cross-equation restrictions implied by the structural RE model is more efficient. But in practice, we are rarely certain about the specification of the complete model. Given the possible spill-over of mis-specification across equations, one may obtain more robust estimates of the counterfactuals from a simpler single equation specification. This can be done by eliminating the effects of $\mathbf{y}_{2t}$ and its lagged value $\mathbf{y}_{2,t-1}$ from (10) following the procedure set out in Zellner and Palm (1974).[3] This is achieved at the expense of introducing lagged values of the remaining variables, y_{1t}, $\mathbf{x}_t$ and $\mathbf{w}_t$, into the analysis. The resultant equation, known as the 'final form', is given by the following infinite order distributed specification

$$y_{1t} = \sum_{i=1}^{\infty} \lambda_i(\boldsymbol{\theta}) y_{1,t-i} + \sum_{i=0}^{\infty} \boldsymbol{\pi}_{yx,i}'(\boldsymbol{\theta})\mathbf{x}_{t-i} + \sum_{i=0}^{\infty} \boldsymbol{\pi}_{yw,i}'(\boldsymbol{\theta})\mathbf{w}_{t-i} + v_{yt}, \tag{11}$$

where $\boldsymbol{\theta} = (\boldsymbol{\varphi}', \boldsymbol{\rho}')'$. However, in the case of stable dynamic models the coefficients of the lagged variables, $\{\lambda_i(\boldsymbol{\theta}), \boldsymbol{\pi}_{yx,i}(\boldsymbol{\theta}), \boldsymbol{\pi}_{yw,i}(\boldsymbol{\theta})\}$, decay exponentially and the infinite order distributed lagged model can be well approximated by the following finite order $ARDL(p_y, p_x, p_w)$ specification

$$y_{1t} = \sum_{i=1}^{p_y} \lambda_i(\boldsymbol{\theta}) y_{1,t-i} + \sum_{i=0}^{p_x} \boldsymbol{\pi}_{yx,i}'(\boldsymbol{\theta})\mathbf{x}_{t-i} + \sum_{i=0}^{p_w} \boldsymbol{\pi}_{yw,i}'(\boldsymbol{\theta})\mathbf{w}_{t-i} + v_{yt}, \tag{12}$$

where the lag orders (p_y, p_x, p_w) can be selected using Akaike and Schwarz information criteria. Alternatively, following Berk (1974), one can set the lag order $p = \max(p_y, p_x, p_w)$ such that $p^3/T \to \kappa$, where T is the sample size (defined below) and κ is a finite constant. This ensures that for sufficiently large T, the reduced form residuals, v_{yt}, are serially uncorrelated.

3. Tests of policy ineffectiveness using reduced or final forms

We will consider a policy change that occurs at time $t = T_0$, with its effects realized from period $T_0 + 1$ onwards. The pre- and post-intervention samples are defined by the periods $t = M, M+1, \ldots, T_0$, and $t = T_0+1, T_0+2, \ldots, T_0+H$, respectively. Therefore, the evaluation horizon is H and the sample size for estimation of the parameters before the policy change is $T = T_0 - M + 1$. This set up allows us to increase the sample size T (by letting $M \to -\infty$), while keeping the time of the policy change, end of T_0, fixed.

We first consider the static case, where we can use the reduced form, then the dynamic case where we use the final form. While dynamics are likely to be important in practice, the static case illuminates certain aspects of the procedure and enables us to derive an exact test that allows for parameter estimation uncertainty. For simplicity, we shall assume that there is a single scalar policy variable, x_t. We will denote the counterfactual outcomes of the policy variable by $\tilde{x}_{T_0+h}$, and the corresponding counterfactual outcomes for the target variable $\tilde{y}_{T_0+h}$, $h = 1, 2, \ldots, H$.

In the implementation of the test, we distinguish between two types of policy change. One type of intervention is a discretionary policy change that shifts the future values of the policy variable by a deterministic amount, so that $\tilde{x}_{T_0+h} = x_{T_0+h} + \delta_{T_0+h}$ for $h = 1, 2, \ldots, H$, where δ_{T_0+h} is set by the policy maker. The deterministic sequence $\{\delta_{T_0+h},\ h = 1, 2, \ldots, H\}$ is announced at time T_0 and may or may not change the parameters of the model. The other type of intervention is a rule-based policy change to one or more elements of $\boldsymbol{\rho}$, the parameters of a policy rule. Since the policy rule equations are stochastic with errors η_{xt}, this further complicates the treatment of the counterfactual. Testing the discretionary policy change is much simpler since the counterfactual values of x_t are known with certainty, while under the rule based policy change, the counterfactual values of x_t must be estimated and do not coincide with their realized values. The discretionary policy intervention could also involve changes to the parameters of a policy rule, so long as one can take the post-intervention counterfactual values of the policy variable, $\{\tilde{x}_{T_0+h},\ h = 1, 2, \ldots, H\}$, as known with certainty for all $h > 0$.

[3] It is well known that univariate representations of variables in a VAR are ARMA (autoregressive moving average) processes. For example, in the case where $\mathbf{y}_{2t}$ is a scalar variable and the RE model does not contain any exogenous variables, the univariate representation of y_{1t} will be an ARMA(2,1) process. However, in practice, such ARMA processes are approximated by high order AR processes.

4. The static case

In the static case, using (8) and setting $\mathbf{\Phi}(\boldsymbol{\varphi})=\mathbf{0}$, the reduced form is given by

$$\mathbf{y}_t=\boldsymbol{\psi}_x(\boldsymbol{\theta})x_t+\mathbf{\Psi}_w(\boldsymbol{\theta})\mathbf{w}_t+\mathbf{v}_t. \tag{13}$$

Partition $\mathbf{y}_t=(y_{1t},\mathbf{y}_{2t}')'$ where y_{1t} is the target variable and $\mathbf{y}_{2t}$ the remaining endogenous variables. We can write the first equation of (13) as the following policy response equation (using y_t for y_{1t} and v_t for v_{yt} the first element of $\mathbf{v}_t$ to simplify the notations)

$$y_t=\psi_{yx}(\boldsymbol{\theta})x_t+\boldsymbol{\psi}_{yw}'(\boldsymbol{\theta})\mathbf{w}_t+v_t, \tag{14}$$

which does not depend on $\mathbf{y}_{2t}$. But it is clear that the parameters of (14) depend on the structural and policy coefficients, $\boldsymbol{\theta}=(\boldsymbol{\varphi}',\boldsymbol{\rho}')'$.

A policy intervention which changes $\boldsymbol{\rho}^0$ to $\boldsymbol{\rho}^1$ will cause the policy response equation to exhibit a break at time T_0+1:

$$y_t=\psi_{yx}(\boldsymbol{\rho}^0)x_t+\boldsymbol{\psi}_{yw}'(\boldsymbol{\rho}^0)\mathbf{w}_t+v_t,\quad t=M,M+1,M+2,\ldots,T_0, \tag{15}$$

$$y_t=\psi_{yx}(\boldsymbol{\rho}^1)x_t+\boldsymbol{\psi}_{yw}'(\boldsymbol{\rho}^1)\mathbf{w}_t+v_t,\quad t=T_0+1,T_0+2,\ldots,T_0+H. \tag{16}$$

In the static case, the policy response equation is given by (15). Under the joint null hypothesis of (i) no policy intervention and (ii) no change in the other parameters, the counterfactual outcomes of y_t are the h-period forecasts, for $h=1,2,\ldots,H$, made at time T_0 conditional on $\tilde{x}_{T_0+h}$, and the realized values of the policy invariant variables, $\mathbf{w}_{T+h}$, namely

$$\tilde{y}_{T_0+h}=\psi_{yx}(\boldsymbol{\rho}^0)\tilde{x}_{T_0+h}+\boldsymbol{\psi}_{yw}'(\boldsymbol{\rho}^0)\mathbf{w}_{T_0+h},\quad h=1,2,\ldots,H. \tag{17}$$

The policy effects are

$$d_{T_0+h}=y_{T_0+h}-\tilde{y}_{T_0+h},\quad h=1,2,\ldots,H. \tag{18}$$

It is instructive to decompose the policy effects (18) into the direct part due to the change in the policy variable and the indirect part that arises due to the policy-induced parameter changes. Using (16) and (17) we have

$$d_{T_0+h}=\left[\psi_{yx}(\boldsymbol{\rho}^1)x_{T_0+h}-\psi_{yx}(\boldsymbol{\rho}^0)\tilde{x}_{T_0+h}\right]+\left[\boldsymbol{\psi}_{yw}(\boldsymbol{\rho}^1)-\boldsymbol{\psi}_{yw}(\boldsymbol{\rho}^0)\right]'\mathbf{w}_{T_0+h}+v_{T_0+h}, \tag{19}$$

which can be written equivalently as

$$d_{T_0+h}=\psi_{yx}(\boldsymbol{\rho}^0)(x_{T_0+h}-\tilde{x}_{T_0+h})+\left[\psi_{yx}(\boldsymbol{\rho}^1)-\psi_{yx}(\boldsymbol{\rho}^0)\right]x_{T_0+h}+\left[\boldsymbol{\psi}_{yw}(\boldsymbol{\rho}^1)-\boldsymbol{\psi}_{yw}(\boldsymbol{\rho}^0)\right]'\mathbf{w}_{T_0+h}+v_{T_0+h} \tag{20}$$

for $h=1,2,\ldots,H$. The first term captures the effects of the change in the policy variable, x_t, while the remaining terms capture the effects of policy-induced parameter changes. Only the first term would be present in the case of *ad hoc* policy changes that do not induce parameter changes in the policy response equation. In either case, the pure policy effect is augmented by the post-intervention random errors, v_{T_0+h}. But we can reduce the importance of such random influences both by using estimates of the mean policy effect,

$$\overline{d}_H=H^{-1}\sum_{h=1}^{H}d_{T_0+h}, \tag{21}$$

and by using the policy invariant variables, $\mathbf{w}_t$, when they are available, as in (19).

In the static case, the policy effects can be empirically evaluated by developing tests based on the individual estimates of the policy effects, d_{T_0+h}, for $h=1,2,\ldots,H$, defined by (19), or the mean policy effect defined by (21). The former is appropriate even if H is fixed and small relative to T (the pre-policy intervention sample size), but requires strong distributional assumptions on the random component of target variable, v_t, and the policy innovations, η_{xt} (when considering stochastic policy rules). When H is fixed and the test is based on the individual estimates of d_{T_0+h}, for $h=1,2,\ldots,H$, we shall assume that

$$\begin{pmatrix} v_t \\ \eta_{xt} \end{pmatrix}\sim N\left[\mathbf{0},\begin{pmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_{\eta_x}^2 \end{pmatrix}\right]. \tag{22}$$

However, the Gaussian assumption is not needed if the policy ineffectiveness test is based on the mean policy effects, and H and T are large.

4.1. Discretionary policy change

Let $\mathbf{x}_{(0)}$ be the $T\times 1$ vector of observations on the policy variable before the intervention, and let $\mathbf{x}_{(1)}$ be the $H\times 1$ vector of realized values on x_t after the intervention. Similarly, let $\mathbf{W}_{(0)}$ to be the $T\times k_w$ matrix of observations on the policy invariant variables, $\mathbf{w}_t$, pre-intervention and let $\mathbf{W}_{(1)}$ be the $H\times k_w$ matrix of observations on $\mathbf{w}_t$ post-intervention. Initially, we assume that H is fixed. The vector of policy effects, (20), $\mathbf{d}_{(1)}=(d_{T_0+1},d_{T_0+2},\dots,d_{T_0+H})'$, can be written as

$$\mathbf{d}_{(1)}=\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}\right]\psi^0_{yx}+\mathbf{x}_{(1)}\left(\psi^1_{yx}-\psi^0_{yx}\right)+\mathbf{W}_{(1)}\left(\boldsymbol{\psi}^1_{yw}-\boldsymbol{\psi}^0_{yw}\right)+\mathbf{v}_{(1)} \tag{23}$$

where $\tilde{\mathbf{x}}_{(1)}=(\tilde{x}_{T_0+1},\tilde{x}_{T_0+2},\dots,\tilde{x}_{T_0+H})'$, $\psi^i_{yx}=\psi_{yx}(\boldsymbol{\rho}^i)$, and $\boldsymbol{\psi}^i_{yw}=\boldsymbol{\psi}_{yw}(\boldsymbol{\rho}^i)$, for i=0,1, and $\mathbf{v}_{(1)}=(v_{T_0+1},v_{T_0+2},\dots,v_{T_0+H})'$.

Under discretionary policy $\tilde{\mathbf{x}}_{(1)}=\mathbf{x}_{(1)}+\boldsymbol{\delta}_{(1)}$, where $\boldsymbol{\delta}_{(1)}=(\delta_{T_0+1},\delta_{T_0+2},\dots,\delta_{T_0+H})'$ is set by the policy maker and is non-stochastic. The policy effects can be estimated as

$$\hat{\mathbf{d}}_{(1)}=\mathbf{y}_{(1)}-\widehat{\tilde{\mathbf{y}}}_{(1)}, \tag{24}$$

where $\mathbf{y}_{(1)}=(y_{T_0+1},y_{T_0+2},\dots,y_{T_0+H})'$, and

$$\widehat{\tilde{\mathbf{y}}}_{(1)}=\tilde{\mathbf{x}}_{(1)}\hat{\psi}^0_{yx}+\mathbf{W}_{(1)}\hat{\boldsymbol{\psi}}^0_{yw}=\left[\mathbf{x}_{(1)}+\boldsymbol{\delta}_{(1)}\right]\hat{\psi}^0_{yx}+\mathbf{W}_{(1)}\hat{\boldsymbol{\psi}}^0_{yw}=\boldsymbol{\delta}_{(1)}\hat{\psi}^0_{yx}+\mathbf{S}_{(1)}\hat{\boldsymbol{\psi}}^0_{ys}, \tag{25}$$

$\mathbf{S}_{(1)}=\left[\mathbf{x}_{(1)},\mathbf{W}_{(1)}\right]$, and $\hat{\boldsymbol{\psi}}^0_{ys}=\left(\hat{\boldsymbol{\psi}}^{0\prime}_{yx},\hat{\boldsymbol{\psi}}^{0\prime}_{yw}\right)'$ is the least squares estimate of the coefficients in the regression of y_t on $\mathbf{s}_t=\left(\mathbf{x}'_t,\mathbf{w}'_t\right)'$, using the pre-intervention sample. More specifically,

$$\hat{\boldsymbol{\psi}}^0_{ys}=\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{y}_{(0)}, \tag{26}$$

where $\mathbf{y}_{(0)}=(y_M,y_{M+1},\dots,y_{T_0})'$, and $\mathbf{S}_{(0)}=\left(\mathbf{x}_{(0)},\mathbf{W}_{(0)}\right)$ is the $T\times(1+k_w)$ matrix of pre-intervention observations on $\mathbf{s}_t$. Post-intervention, we assume the same policy response equation holds, albeit with different parameter values, namely

$$\mathbf{y}_{(1)}=\mathbf{S}_{(1)}\boldsymbol{\psi}^1_{ys}+\mathbf{v}_{(1)}. \tag{27}$$

Hence, using (25) and (27) in (24), under discretionary policy we have

$$\hat{\mathbf{d}}_{(1)}=-\boldsymbol{\delta}_{(1)}\hat{\psi}^0_{yx}+\mathbf{S}_{(1)}\left(\boldsymbol{\psi}^1_{ys}-\hat{\boldsymbol{\psi}}^0_{ys}\right)+\mathbf{v}_{(1)}, \tag{28}$$

or

$$\hat{\mathbf{d}}_{(1)}=-\boldsymbol{\delta}_{(1)}\psi^0_{yx}+\mathbf{S}_{(1)}\left(\boldsymbol{\psi}^1_{ys}-\boldsymbol{\psi}^0_{ys}\right)-\boldsymbol{\delta}_{(1)}\left(\hat{\psi}^0_{yx}-\psi^0_{yx}\right)-\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\psi}}^0_{ys}-\boldsymbol{\psi}^0_{ys}\right)+\mathbf{v}_{(1)}.$$

Also

$$\begin{aligned}\boldsymbol{\delta}_{(1)}\left(\hat{\psi}^0_{yx}-\psi^0_{yx}\right)+\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\psi}}^0_{ys}-\boldsymbol{\psi}^0_{ys}\right)&=\boldsymbol{\delta}_{(1)}\left(\hat{\psi}^0_{yx}-\psi^0_{yx}\right)+\mathbf{x}_{(1)}\left(\hat{\psi}^0_{yx}-\psi^0_{yx}\right)+\mathbf{W}_{(1)}\left(\hat{\psi}^0_{yw}-\psi^0_{yw}\right)=\left[\mathbf{x}_{(1)}+\boldsymbol{\delta}_{(1)}\right]\left(\hat{\psi}^0_{yx}-\psi^0_{yx}\right)\\&+\mathbf{W}_{(1)}\left(\hat{\psi}^0_{yw}-\psi^0_{yw}\right)=\tilde{\mathbf{S}}_1\left(\hat{\boldsymbol{\psi}}^0_{ys}-\boldsymbol{\psi}^0_{ys}\right),\end{aligned}$$

where

$$\tilde{\mathbf{S}}_1=\left[\mathbf{x}_{(1)}+\boldsymbol{\delta}_{(1)},\mathbf{W}_{(1)}\right]=\left[\tilde{\mathbf{x}}_{(1)},\mathbf{W}_{(1)}\right]. \tag{29}$$

Further, from (26) we have

$$\hat{\boldsymbol{\psi}}^0_{ys}-\boldsymbol{\psi}^0_{ys}=\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{v}_{(0)}, \tag{30}$$

where $\mathbf{v}_{(0)}=(v_M,v_{M+1},\dots,v_{T_0})'$. Therefore, $\hat{\mathbf{d}}_{(1)}$ can be written as

$$\hat{\mathbf{d}}_{(1)}=\boldsymbol{\mu}_{(1)}+\boldsymbol{\xi}_{(1)}, \tag{31}$$

where

$$\boldsymbol{\mu}_{(1)}=-\boldsymbol{\delta}_{(1)}\psi^0_{yx}+\tilde{\mathbf{S}}_{(1)}\left(\boldsymbol{\psi}^1_{ys}-\boldsymbol{\psi}^0_{ys}\right), \tag{32}$$

and

$$\boldsymbol{\xi}_{(1)}=\mathbf{v}_{(1)}-\tilde{\mathbf{S}}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{v}_{(0)}, \tag{33}$$

with $\tilde{\mathbf{S}}_{(1)}$ defined by (29). Note that under a discretionary policy change, $\boldsymbol{\mu}_{(1)}$ is a deterministic sequence and $\boldsymbol{\xi}_{(1)}$ is the stochastic component of the policy effects.

To derive the distribution of $\hat{\mathbf{d}}_{(1)}$ we assume that all the classical assumptions apply to the policy response equation during the pre-intervention sample ($t=M,M+1,\dots,T_0$), namely $\mathbf{s}_t$ and $v_{t'}$ are uncorrelated for all t and t', and v_t are serially uncorrelated with a constant variance, σ^2_v. We also assume that $\mathbf{W}_{(1)}$ and $\mathbf{v}=\left[\mathbf{v}'_{(0)},\mathbf{v}'_{(1)}\right]'$ are uncorrelated, and $E\left[\mathbf{v}_{(1)}\mathbf{v}'_{(0)}\right]=\mathbf{0}$, $E\left[\mathbf{v}_{(0)}\mathbf{v}'_{(0)}\right]=\sigma^2_{0v}\mathbf{I}_T$, and $E\left[\mathbf{v}_{(1)}\mathbf{v}'_{(1)}\right]=\sigma^2_{1v}\mathbf{I}_H$. Note that in this case, we do not need to make any assumptions concerning the realized values of $\mathbf{x}_t$, over the post-intervention sample. In what follows, we also assume that the policy change does not alter the error variances, namely $\sigma^2_{0v}=\sigma^2_{1v}$. Accordingly, we define the implicit null of the policy ineffectiveness test as

$$\mathcal{H}_0{:}\,\boldsymbol{\mu}_{(1)}=\mathbf{0},\quad \sigma^2_{0v}=\sigma^2_{1v}. \tag{34}$$

Under $\mathcal{H}_0$,

$$\hat{\mathbf{d}}_{(1)} = \mathbf{v}_{(1)} - \tilde{\mathbf{S}}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{v}_{(0)}, \tag{35}$$

and assuming that the above classical assumptions hold we have $\hat{\mathbf{d}}_{(1)} \sim N(\mathbf{0}, \Omega_d)$, where

$$\mathbf{\Omega}_d = \sigma^2_{0v}\left\{\mathbf{I}_H + \tilde{\mathbf{S}}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\tilde{\mathbf{S}}'_{(1)}\right\}.$$

A test can now be based on all the individual H elements of $\hat{\mathbf{d}}_{(1)}$ which yields the joint test statistic

$$X^2_{d,H} = \frac{\hat{\mathbf{d}}'_{(1)}\left\{\mathbf{I}_H + \tilde{\mathbf{S}}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\tilde{\mathbf{S}}'_{(1)}\right\}^{-1}\hat{\mathbf{d}}_{(1)}}{\sigma^2_{0v}}. \tag{36}$$

Under the policy ineffectiveness hypothesis, H_0, and assuming that $\mathbf{v} = \left[\mathbf{v}'_{(0)}, \mathbf{v}'_{(1)}\right]'$ is normally distributed, then $X^2_{d,H}$ is distributed as a Chi-square variate with H degrees of freedom. If σ^2_{0v} is replaced by its unbiased estimator based on the pre-intervention sample,

$$\hat{\sigma}^2_{0v} = \frac{\left[\mathbf{y}_{(0)} - \mathbf{S}_{(0)}\hat{\boldsymbol{\psi}}^0_{ys}\right]'\left[\mathbf{y}_{(0)} - \mathbf{S}_{(0)}\hat{\boldsymbol{\psi}}^0_{ys}\right]}{T-1-k_w}, \tag{37}$$

we obtain the feasible test statistic

$$F_{d,H} = \frac{\hat{\mathbf{d}}'_{(1)}\left\{\mathbf{I}_H + \tilde{\mathbf{S}}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\tilde{\mathbf{S}}'_{(1)}\right\}^{-1}\hat{\mathbf{d}}_{(1)}}{H\hat{\sigma}^2_{0v}}, \tag{38}$$

which under the null hypothesis, $\mathcal{H}_0$, is distributed as $F(H, T-k_w-1)$. A proof is provided in Appendix A1. For future reference, it is also worth noting that for sufficiently large T and a fixed H the above test statistic is asymptotically equivalent to[4]

$$X^2_{d,H} = \frac{\hat{\mathbf{d}}'_{(1)}\hat{\mathbf{d}}_{(1)}}{\hat{\sigma}^2_{0v}} \overset{a}{\sim} \chi^2_H. \tag{39}$$

Alternatively, one can base a test of $\mathcal{H}_0$ on linear combinations of the elements of $\hat{\mathbf{d}}_{(1)}$, such as the mean given by (21), which can be estimated by

$$\bar{\hat{d}}_H = H^{-1}\boldsymbol{\tau}'_H\hat{\mathbf{d}}_{(1)}, \tag{40}$$

where $\boldsymbol{\tau}_H$ is a vector of ones of length H. The policy ineffectiveness test statistic for the mean effect is given by

$$t_{d,H} = \frac{\sqrt{H}\bar{\hat{d}}_H}{\hat{\sigma}_{0v}\sqrt{1 + T^{-1}\left\{H^{-1}\boldsymbol{\tau}'_H\mathbf{S}^0_{(1)}\left[T^{-1}\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}^{0\prime}_{(1)}\boldsymbol{\tau}_H\right\}}}. \tag{41}$$

For this test, the assumption that $\mathbf{v}_{(1)}$ is normally distributed can be relaxed, so long as H is sufficiently large. This result holds irrespective of the relative rate at which H and $T \to \infty$. It is now easily seen that under the null of policy ineffectiveness, $t_{d,H} \overset{a}{\sim} N(0,1)$ when both H and $T \to \infty$. In the case where T is large relative to H, estimation uncertainty can be ignored and the second term in the denominator of (41) becomes relatively negligible and we have

$$t^a_{d,H} = \sqrt{H}\frac{\bar{\hat{d}}_H}{\hat{\sigma}_{0v}} \overset{a}{\sim} N(0,1), \tag{42}$$

where $\hat{\sigma}_{0v}$ is defined by (37).

A test based on $\bar{\hat{d}}_H$ is not a classical test since although the implicit null of no policy change, (34), is equality of the parameters before and after the policy intervention, this is not tested directly, but only indirectly through the effect of the change on the estimated policy effects. Nor is the test based on the H individual elements of $\hat{\mathbf{d}}_{(1)}$ a predictive failure test, since the counterfactual outcomes are not *ex ante* forecasts, being computed conditional on the realizations of $\mathbf{w}_{T_0+h}$ and not their predictions. Nor are the counterfactuals *ex post* forecasts either, since they are based on estimated projected values for the policy variable and the initial values of the target variable.

Finally, in applications where the sign of the expected effect of the policy change on the target variable is known, a one sided version of the above tests can be entertained.

[4] Note that, for a fixed H, the second term inside $\{\cdot\}$ in (38) is $O_p(T^{-1})$.

4.2. Power of tests for a discretionary policy change

In considering the power of policy ineffectiveness tests, we need to distinguish between the F test defined by (38) where the test is based on individual elements of the policy effects, $\hat{\mathbf{d}}_{(1)}$, and the standard normal test defined by (41) which is based on the mean policy effect. In both cases the power of the test depends on the size of the policy change, the state of the economy at the time of the policy change, the distribution of post policy intervention shocks, $\mathbf{v}_{(1)}$, and the evaluation horizon, H. The extent of parameter uncertainty can also affect the power of the tests, but its effect will be of second-order in importance. However, for the tests to be consistent it is required that $H/T \to 0$, as $H \to \infty$ and $T \to \infty$.

The key component of the power analysis is the size of the policy change, $\boldsymbol{\mu}_{(1)}$, defined by (32). To see this note that under the alternative of policy intervention (denoted by $\mathcal{H}_1$) we have (assuming H is fixed and T sufficiently large)

$$F_{d,H} \overset{a}{\sim} \frac{\left[\boldsymbol{\mu}_{(1)}+\mathbf{v}_{(1)}\right]'\left[\boldsymbol{\mu}_{(1)}+\mathbf{v}_{(1)}\right]}{H\sigma_{0v}^2}+O_p\left(T^{-1/2}\right),$$

which yields a non-central χ^2_H with the non-centrality parameter, $H^{-1}\left[\boldsymbol{\mu}_{(1)}'\boldsymbol{\mu}_{(1)}/\sigma_{0v}^2\right]$. For a given value of H, the power of the test is governed by $\boldsymbol{\mu}_{(1)}/\sigma_{0v}$ which measures the effects of the policy change normalized by the standard deviation of the shocks, σ_{0v}. Also, note that the F test assumes that the post-intervention innovations, $\mathbf{v}_{(1)}$, are Gaussian.

Consider now the test based on mean policy effects and assume that H is sufficiently large. Using (31) in (42) and after some algebra we have

$$\begin{aligned} t_{d,H}^a &= \frac{\sqrt{H}\bar{\hat{d}}_H}{\hat{\sigma}_{0v}} = \frac{H^{-1/2}\boldsymbol{\tau}_H'\hat{\mathbf{d}}_{(1)}}{\sigma_{0v}}+O_p\left(T^{-1/2}\right) = \frac{H^{-1/2}\boldsymbol{\tau}_H'\left[\boldsymbol{\mu}_{(1)}+\mathbf{v}_{(1)}-\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\psi}}_{ys}^0-\psi_{ys}^0\right)\right]}{\sigma_{0v}}+O_p\left(T^{-1/2}\right) \\ &= \frac{H^{-1/2}\boldsymbol{\tau}_H'\boldsymbol{\mu}_{(1)}}{\sigma_{0v}} - \frac{H^{-1/2}\boldsymbol{\tau}_H'\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\psi}}_{ys}^0-\psi_{ys}^0\right)}{\sigma_{0v}} + \frac{H^{-1/2}\boldsymbol{\tau}_H'\mathbf{v}_{(1)}}{\sigma_{0v}}+O_p\left(T^{-1/2}\right). \end{aligned}$$

The third term tends to a normal distribution as $H \to \infty$ with a zero mean. Here we do not require the innovations to be Gaussian. It is also easily seen that since $\hat{\boldsymbol{\psi}}_{ys}^0-\psi_{ys}^0 = O_p(T^{-1/2})$, then the second term is of order $O_p(H^{1/2}T^{-1/2})$ and therefore vanishes as T and $H \to \infty$, so long as $H/T \to 0$. Hence, the power of the test depends on the first term. Using (32) we have,

$$\frac{H^{-1/2}\boldsymbol{\tau}_H'\boldsymbol{\mu}_{(1)}}{\sigma_{0v}} = \frac{-H^{-1/2}\boldsymbol{\tau}_H'\boldsymbol{\delta}_{(1)}\psi_{yx}^0}{\sigma_{0v}} + \frac{H^{-1/2}\boldsymbol{\tau}_H'\mathbf{S}_{(1)}\left(\boldsymbol{\psi}_{ys}^1-\boldsymbol{\psi}_{ys}^0\right)}{\sigma_{0v}}, \tag{43}$$

which can be written as

$$\frac{H^{-1/2}\boldsymbol{\tau}_H'\boldsymbol{\mu}_{(1)}}{\sigma_{0v}} = -\sqrt{H}\left[\frac{\psi_{yx}^0\bar{\delta}_{(1),H}}{\sigma_{0v}}\right] + \sqrt{H}\left[\frac{\left(\boldsymbol{\psi}_{ys}^1-\boldsymbol{\psi}_{ys}^0\right)'\bar{\mathbf{s}}_{(1),H}}{\sigma_{0v}}\right]$$

where

$$\bar{\delta}_{(1),H} = H^{-1}\sum_{h=1}^{H}\delta_{T_0+h}, \quad \text{and} \quad \bar{\mathbf{s}}_{(1),H} = H^{-1}\sum_{h=1}^{H}\mathbf{s}_{T_0+h}.$$

Not surprisingly, the power of the test depends on the nature of the policy change. The test is powerful if $\bar{\delta}_{(1),H} \neq 0$ and/or $\boldsymbol{\psi}_{ys}^1 \neq \boldsymbol{\psi}_{ys}^0$, and rises with $\sqrt{H}$ if

$$p\lim_{H\to\infty}\left[\left(\boldsymbol{\psi}_{ys}^1-\boldsymbol{\psi}_{ys}^0\right)'\bar{\mathbf{s}}_{(1),H}-\psi_{yx}^0\bar{\delta}_{(1),H}\right] \neq \mathbf{0}. \tag{44}$$

In the case where the policy is purely discretionary and only alters the policy variable and not the coefficients of the policy equation, then the test has power if $\psi_{yx}^0 \lim_{H\to\infty}\left[\bar{\delta}_{(1),H}\right] \neq \mathbf{0}$. But in general, the power of the test could be enhanced or diminished depending on the compensating impact of the policy change on the coefficients of policy and non-policy exogenous variables. When condition (44) is met, a test based on $t_{d,H}^a$ will be consistent, in the sense that under the alternative, that the policy change has been effective, the probability of rejecting the null of no policy effect will tend to unity with H and $T \to \infty$, so long as $H/T \to 0$.

4.3. Testing for changes to a policy rule

In this case, we assume the policy rule is formulated in terms of changes in $\boldsymbol{\rho}$, the parameters of the policy equation determining x_t, from $\boldsymbol{\rho}^0$ to $\boldsymbol{\rho}^1$. For instance, suppose the policy variable, x_t, follows the AR(1) process

$$x_t = \mu_x(1-\rho_x)+\rho_x x_{t-1}+\eta_{xt}, \quad |\rho_x|<1, \tag{45}$$

where $\boldsymbol{\rho} = (\mu_x, \rho_x)'$, and $\eta_{xt} \sim IID(0, \sigma_{\eta_x}^2)$. Then the counterfactual values of the policy variable under the null hypothesis of no change in policy are given by

$$\tilde{x}_{T_0+h}(\boldsymbol{\rho}^0)=\mu_x^0\left[1-(\rho_x^0)^h\right]+(\rho_x^0)^h x_{T_0}, \quad \text{for} \quad h=1,2,\ldots,H, \tag{46}$$

where $\boldsymbol{\rho}^0=(\mu_x^0,\rho_x^0)'$. The vector of the counterfactual values of y_{T_0+h}, for $h=1,2,\ldots,H$, is now given by

$$\widehat{\tilde{\mathbf{y}}}_{(1)}=\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\hat{\psi}_{yx}^0+\mathbf{W}_{(1)}\hat{\boldsymbol{\psi}}_{yw}^0=\left[\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)-\mathbf{x}_{(1)}\right]\hat{\psi}_{yx}^0+\mathbf{S}_{(1)}\hat{\boldsymbol{\psi}}_{ys}^0,$$

where $\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)=\left[\tilde{x}_{T_0+1}(\hat{\boldsymbol{\rho}}^0),\tilde{x}_{T_0+2}(\hat{\boldsymbol{\rho}}^0),\ldots,\tilde{x}_{T_0+H}(\hat{\boldsymbol{\rho}}^0)\right]'$. As compared to (25), the above expression for $\widehat{\tilde{\mathbf{y}}}_{(1)}$ has the stochastic component $\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)-\mathbf{x}_{(1)}$, instead of the discretionary deterministic component, $\boldsymbol{\delta}_{(1)}$. This introduces further randomness into the analysis which depends on the policy rule assumed for x_t.

Hence,

$$\hat{\mathbf{d}}_{(1)}=\mathbf{y}_{(1)}-\widehat{\tilde{\mathbf{y}}}_{(1)}=\mathbf{y}_{(1)}-\left[\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)-\mathbf{x}_{(1)}\right]\hat{\psi}_{yx}^0-\mathbf{S}_{(1)}\hat{\boldsymbol{\psi}}_{ys}^0, \tag{47}$$

which upon using (27) yields

$$\hat{\mathbf{d}}_{(1)}=\mathbf{S}_{(1)}\left(\boldsymbol{\psi}_{ys}^1-\boldsymbol{\psi}_{ys}^0\right)+\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\hat{\psi}_{yx}^0+\boldsymbol{\xi}_{(1)}, \tag{48}$$

where $\boldsymbol{\xi}_{(1)}$ is now given by

$$\boldsymbol{\xi}_{(1)}=\mathbf{v}_{(1)}-\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\psi}}_{ys}^0-\boldsymbol{\psi}_{ys}^0\right). \tag{49}$$

Consider now the second term of $\hat{\mathbf{d}}_{(1)}$ and note that

$$\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}\left(\hat{\boldsymbol{\rho}}^0\right)\right]\hat{\psi}_{yx}^0=\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)+\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\psi_{yx}^0. \tag{50}$$

Also since

$$x_{T_0+h}=\mu_x^1\left[1-(\rho_x^1)^h\right]+(\rho_x^1)^h x_{T_0}+\sum_{j=0}^{h-1}(\rho_x^1)^j\eta_{x,t+h-j},$$

then

$$\mathbf{x}_{(1)}=\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^1)+\boldsymbol{\Gamma}(\boldsymbol{\rho}^1)\boldsymbol{\eta}_{(1)},$$

where

$$\boldsymbol{\Gamma}(\rho_x^1)=\begin{pmatrix}1 & 0 & \cdots & 0\\ \rho_x^1 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ (\rho_x^1)^{h-1} & (\rho_x^1)^{h-2} & \cdots & 1\end{pmatrix}, \quad \boldsymbol{\eta}_{(1)}=\begin{pmatrix}\eta_{x,t+1}\\ \eta_{x,t+2}\\ \vdots\\ \eta_{x,t+H}\end{pmatrix}.$$

Hence,

$$\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}\left(\hat{\boldsymbol{\rho}}^0\right)\right]=\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^1)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]+\boldsymbol{\Gamma}(\rho_x^1)\boldsymbol{\eta}_{(1)}.$$

Using the above result in (50), we obtain

$$\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}\left(\hat{\boldsymbol{\rho}}^0\right)\right]\hat{\psi}_{yx}^0=\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^1)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)+\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)\boldsymbol{\Gamma}(\rho_x^1)\boldsymbol{\eta}_{(1)}+\left[\mathbf{x}_{(1)}-\tilde{\mathbf{x}}_{(1)}\left(\hat{\boldsymbol{\rho}}^0\right)\right]\psi_{yx}^0,$$

which, if used in (48), yields

$$\hat{\mathbf{d}}_{(1)}=\mathbf{S}_{(1)}\left(\boldsymbol{\psi}_{ys}^1-\boldsymbol{\psi}_{ys}^0\right)+\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^1)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)+\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)\boldsymbol{\Gamma}(\rho_x^1)\boldsymbol{\eta}_{(1)}+\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^1)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\psi_{yx}^0+\psi_{yx}^0\boldsymbol{\Gamma}(\rho_x^1)\boldsymbol{\eta}_{(1)}+\boldsymbol{\xi}_{(1)}.$$

But under the null hypothesis of no policy change $\mathcal{H}_0:\boldsymbol{\rho}^1=\boldsymbol{\rho}^0$ (and hence $\boldsymbol{\psi}_{ys}^1=\boldsymbol{\psi}_{ys}^0$) and $\hat{\mathbf{d}}_{(1)}$ simplifies to

$$\hat{\mathbf{d}}_{(1)}=\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^0)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)+\left(\hat{\psi}_{yx}^0-\psi_{yx}^0\right)\boldsymbol{\Gamma}(\rho_x^0)\boldsymbol{\eta}_{(1)}+\psi_{yx}^0\left[\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^0)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)\right]+\psi_{yx}^0\boldsymbol{\Gamma}(\rho_x^0)\boldsymbol{\eta}_{(1)}+\boldsymbol{\xi}_{(1)}.$$

Furthermore, using (46) the hth element of $\tilde{\mathbf{x}}_{(1)}(\boldsymbol{\rho}^0)-\tilde{\mathbf{x}}_{(1)}(\hat{\boldsymbol{\rho}}^0)$ is given by

$$\mu_x^0\left[1-(\rho_x^0)^h\right]-\hat{\mu}_x^0\left[1-\left(\hat{\rho}_x^0\right)^h\right]+\left[(\rho_x^0)^h-\left(\hat{\rho}_x^0\right)^h\right]x_{T_0}.$$

In the present case, we can only derive the distribution of $\hat{\mathbf{d}}_{(1)}$ when T is sufficiently large. This is due to the dynamic nature of the policy rule. More specifically, assuming $|\rho_x^i| < 1$, for $i=0,1$, and x_{T_0} is given we have (for a fixed H)

$$\tilde{x}_{(1)}(\boldsymbol{\rho}^0) - \tilde{x}_{(1)}(\hat{\boldsymbol{\rho}}^0) = O_p(T^{-1/2}), \quad \text{and} \quad \hat{\psi}_{yx}^0 - \psi_{yx}^0 = O_p(T^{-1/2}).$$

Therefore, under the null of no policy change we have

$$\hat{\mathbf{d}}_{(1)} = \psi_{yx}^0 \boldsymbol{\Gamma}(\boldsymbol{\rho}^0)\boldsymbol{\eta}_{(1)} + \boldsymbol{\xi}_{(1)} + O_p(T^{-1/2}).$$

But using (30) in (49),

$$\boldsymbol{\varepsilon}_{(1)} = \mathbf{v}_{(1)} - \mathbf{S}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{v}_{(0)} = \mathbf{v}_{(1)} - \frac{1}{\sqrt{T}}\mathbf{S}_{(1)}\left[T^{-1}\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}T^{-1/2}\mathbf{S}'_{(0)}\mathbf{v}_{(0)} = \mathbf{v}_{(1)} + O_p\left(T^{-1/2}\right), \quad \text{for a fixed } H.$$

Therefore, under the null of no policy change and for a fixed H

$$\hat{\mathbf{d}}_{(1)} = \mathbf{v}_{(1)} + \psi_{yx}^0 \boldsymbol{\Gamma}(\rho_x^0)\boldsymbol{\eta}_{(1)} + O_p(T^{-1/2}). \tag{51}$$

To obtain the distribution of $\hat{\mathbf{d}}_{(1)}$ we now need to make some assumptions regarding the distribution of $\mathbf{v}_{(1)}$ and $\boldsymbol{\eta}_{(1)}$. Assuming that these are normally distributed (see (22)), for sufficiently large T and a fixed H), we have

$$\hat{\mathbf{d}}'_{(1)}\left[\sigma_v^2 \mathbf{I}_H + \left(\psi_{yx}^0\right)^2 \sigma_{\eta_x}^2 \boldsymbol{\Gamma}(\rho_x^0)\boldsymbol{\Gamma}'(\rho_x^0)\right]^{-1}\hat{\mathbf{d}}_{(1)} \overset{a}{\sim} \chi_H^2.$$

A feasible version of the above test statistic can be obtained under the assumption that σ_v^2 and $\sigma_{\eta_x}^2$ have remained stable over the full sample period (pre- and post-policy change). It is given by

$$\hat{\mathbf{d}}'_{(1)}\left[\hat{\sigma}_{0v}^2 \mathbf{I}_H + \left(\hat{\psi}_{yx}^0\right)^2 \hat{\sigma}_{0\eta_x}^2 \boldsymbol{\Gamma}\left(\hat{\rho}_x^0\right)\boldsymbol{\Gamma}'\left(\hat{\rho}_x^0\right)\right]^{-1}\hat{\mathbf{d}}_{(1)} \overset{a}{\sim} \chi_H^2, \tag{52}$$

where $\hat{\sigma}_{0v}^2$ is given by (37), and

$$\hat{\sigma}_{0\eta_x}^2 = \frac{\sum_{t=1}^T \left[x_t - \hat{\mu}_x^0(1-\hat{\rho}_x^0) - \hat{\rho}_x^0 x_{t-1}\right]^2}{T-2}.$$

It is easily seen that (52) reduces to (39) if we abstract from the future uncertainty of the counterfactual values of the policy variable.

Finally, as before we can also base the test of policy ineffectiveness on the mean policy effect, $\bar{\hat{d}}_H = H^{-1}\boldsymbol{\tau}'_H\hat{\mathbf{d}}_{(1)}$, where $\hat{\mathbf{d}}_{(1)}$ is defined by (47). But due to the dynamic nature of the policy rule, we need to consider the case where H and T are sufficiently large. Using (51) we obtain the test statistic

$$t_{d,H}^a = \frac{\sqrt{H}\bar{\hat{d}}_H}{\left[\hat{\sigma}_{0v}^2 + \left(\hat{\psi}_{yx}^0\right)^2 \hat{\sigma}_{0\eta_x}^2 H^{-1}\boldsymbol{\tau}'_H \boldsymbol{\Gamma}\left(\hat{\rho}_x^0\right)\boldsymbol{\Gamma}'\left(\hat{\rho}_x^0\right)\boldsymbol{\tau}'_H\right]^{1/2}}. \tag{53}$$

Again it is clear that (53) reduces to (42) in the absence of uncertainty about the counterfactual values of the policy variable. As before, it is possible to show that $t_{d,H}^a \to_d N(0,1)$ for $H = \kappa T^{\epsilon}$, with $\epsilon \le 1/2$, as T and $H \to \infty$, where κ is a finite positive constant. See also the derivations in Section 5.2.

5. The dynamic case

Returning to the case with dynamics, we will use the ARDL specification given by (12), again writing y_t for y_{1t} and using the policy effects, (18), $d_{T_0+h} = y_{T_0+h} - \tilde{y}_{T_0+h}$, $h = 1, 2, \dots, H$, which will be subject to the post-intervention random errors, v_{y,T_0+h}. To simplify the exposition we set $p_y = p_x = p_w = p = 1$, and consider an ARDL(1,1,1) specification to derive counterfactual values of y_t. The analysis readily extends to higher-order ARDL specification, as set out in Appendix A2.

For the pre- and post-intervention samples we have[5]

$$y_t = \lambda^0 y_{t-1} + \pi_{yx0}^0 x_t + \pi_{yx1}^0 x_{t-1} + \boldsymbol{\pi}_{yw0}^{0\prime}\mathbf{w}_t + \boldsymbol{\pi}_{yw1}^{0\prime}\mathbf{w}_{t-1} + v_{yt}, \quad \text{for } t = M, M+1, \dots, T_0, \tag{54}$$

$$y_t = \lambda^1 y_{t-1} + \pi_{yx0}^1 x_t + \pi_{yx1}^1 x_{t-1} + \boldsymbol{\pi}_{yw0}^{1\prime}\mathbf{w}_t + \boldsymbol{\pi}_{yw1}^{1\prime}\mathbf{w}_{t-1} + v_{yt}, \quad \text{for } t = T_0+1, \dots, T_0+H, \tag{55}$$

where $\left|\lambda^j\right| < 1$ for $j=0,1$. We will also assume that the estimate of λ^0, denoted by $\hat{\lambda}^0$, satisfies the stationary condition, $\left|\hat{\lambda}^0\right| < 1$. The pre-intervention sample size is given by $T = T_0 - M + 1$, and the post-intervention sample size is given by H. In what follows, we focus on the case of a discretionary policy change under which the future values of the policy variable are assumed as given.

[5] As is clear from (8), a change in policy parameters can change all the parameters.

The ARDL specification for the pre-intervention sample can be written as

$$\mathbf{y}_{(0)} = \lambda^0 \mathbf{y}_{-1,(0)} + \mathbf{S}_{(0)} \pi^0_{ys} + \mathbf{v}_{(0)},$$

where $\mathbf{y}_{(0)}$ and $\mathbf{v}_{(0)}$ are defined as before, $\mathbf{y}_{-1,(0)} = (y_{M-1}, y_M, \ldots, y_{T_0-1})'$, $\mathbf{S}_{(0)} = \left(\mathbf{X}_{(0)}, \mathbf{W}_{(0)}\right), \mathbf{X}_{(0)} = \left(\mathbf{x}_{(0)}, \mathbf{x}_{-1,(0)}\right)$, $\mathbf{x}_{(0)} = (x_M, x_{M+1}, \ldots, x_{T_0})'$, $\mathbf{x}_{-1,(0)} = (x_{M-1}, x_M, \ldots, x_{T_0-1})'$, $\mathbf{W}_{(0)} = \left(\mathbf{w}_{(0)}, \mathbf{w}_{-1,(0)}\right)$, $\mathbf{w}_{(0)} = (\mathbf{w}_M, \mathbf{w}_{M+1}, \ldots, \mathbf{w}_{T_0})'$, $\mathbf{w}_{-1,(0)} = (\mathbf{w}_{M-1}, \mathbf{w}_M, \ldots, \mathbf{w}_{T_0-1})'$, and $\boldsymbol{\pi}^0_{ys} = \left(\pi^0_{yx0}, \pi^0_{yx1}, \boldsymbol{\pi}^{0\prime}_{yw0}, \boldsymbol{\pi}^{0\prime}_{yw1}\right)'$. Further lagged values of x_t as well as deterministic components such as an intercept and linear trends can also be included in $\mathbf{S}_{(0)}$.

Based on this specification and given counterfactual values of the policy variables and their lagged values over the post-intervention sample, which we denote by $\tilde{\mathbf{X}}_{(1)}$, by forward iterations from $t = T_0$ we obtain the following counterfactual outcomes:

$$\hat{\tilde{\mathbf{y}}}_{(1)} = \hat{\boldsymbol{\Lambda}}^0_H \left[\mathbf{e}_1 \hat{\lambda}^0 y_{T_0} + \tilde{\mathbf{X}}_{(1)} \hat{\boldsymbol{\pi}}^0_{yx} + \mathbf{W}_{(1)} \hat{\boldsymbol{\pi}}^0_{yw},\right] = \hat{\boldsymbol{\Lambda}}^0_H \left[\mathbf{e}_1 \hat{\lambda}^0 y_{T_0} + \tilde{\mathbf{S}}_{(1)} \hat{\boldsymbol{\pi}}^0_{ys}\right], \tag{56}$$

where $\tilde{\mathbf{S}}_{(1)} = \left[\tilde{\mathbf{X}}_{(1)}, \mathbf{W}_{(1)}\right]$, the predicted values of the policy variables and the realized values of the policy-invariant variables for the post-intervention sample, and $\hat{\Lambda}^0_H$ is the $H \times H$ lower triangular matrix

$$\hat{\boldsymbol{\Lambda}}^0_H = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \hat{\lambda}^0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \left(\hat{\lambda}^0\right)^{H-1} & \left(\hat{\lambda}^0\right)^{H-2} & \cdots & 1 \end{pmatrix}, \tag{57}$$

$\mathbf{e}_1 = (1, 0, \ldots, 0)'$, and $\hat{\lambda}^0$, $\hat{\boldsymbol{\pi}}^0_{ys} = (\hat{\boldsymbol{\pi}}^{0\prime}_{yx}, \hat{\boldsymbol{\pi}}^{0\prime}_{yw})'$ are least square estimates of λ^0, $\boldsymbol{\pi}^0_{ys} = (\boldsymbol{\pi}^{0\prime}_{yx}, \boldsymbol{\pi}^{0\prime}_{yw})'$ in the dynamic policy equation, (54), based on the pre-intervention sample. More specifically, setting $\varphi^0 = (\lambda^0, \boldsymbol{\pi}^{0\prime}_{ys})'$, and $\mathbf{Q}_{(0)} = \left[\mathbf{y}_{-1,(0)}, \mathbf{S}_{(0)}\right]$, we have

$$\hat{\boldsymbol{\varphi}}^0 = \left[\mathbf{Q}_{(0)} \mathbf{Q}'_{(0)}\right]^{-1} \mathbf{Q}'_{(0)} \mathbf{y}_{(0)}. \tag{58}$$

For future reference, we also note that under fairly general conditions on the error terms, $\mathbf{v}_{(0)}$, and assuming that $\left|\lambda^0\right| < 1$, and $\left|\hat{\lambda}^0\right| < 1$, then as $T \to \infty$ we have

$$\sqrt{T}\left(\hat{\varphi}^0 - \varphi^0\right) \to_d N(\mathbf{0}, \sigma^2_{0v} \Sigma_0^{-1}), \tag{59}$$

where, as before, $\mathbf{v} = \left[\mathbf{v}'_{(0)}, \mathbf{v}'_{(1)}\right]'$, $E(\mathbf{v}\mathbf{v}') = \sigma^2_{0v} \mathbf{I}_{T+H}$, H is finite and $\Sigma_0 = \text{plim}_{T \to \infty} \left[T^{-1} \mathbf{Q}_{(0)} \mathbf{Q}'_{(0)}\right]$ is a positive definite matrix.

The estimates of the policy effects are now given by

$$\hat{\mathbf{d}}_{(1)} = \mathbf{y}_{(1)} - \hat{\boldsymbol{\Lambda}}^0_H \left[y_{T_0} \hat{\lambda}^0 \mathbf{e}_1 + \tilde{\mathbf{S}}_{(1)} \hat{\boldsymbol{\pi}}^0_{ys}\right].$$

As before, this can be decomposed into a systematic effect of the policy, the random components due to $\mathbf{v}_{(1)}$ and the sampling uncertainty in estimation of $\hat{\lambda}^0$, $\hat{\boldsymbol{\pi}}^0_{yx}$, and $\hat{\boldsymbol{\pi}}^0_{yw}$.

Using the forward recursive approach, we first note that

$$\mathbf{y}_{(1)} = \boldsymbol{\Lambda}^1_H \left[y_{T_0} \lambda^1 \mathbf{e}_1 + \mathbf{X}_{(1)} \boldsymbol{\pi}^1_{yx} + \mathbf{W}_{(1)} \boldsymbol{\pi}^1_{yw} + \mathbf{v}_{(1)}\right] = \boldsymbol{\Lambda}^1_H \left[y_{T_0} \lambda^1 \mathbf{e}_1 + \mathbf{S}_{(1)} \boldsymbol{\pi}^1_{ys} + \mathbf{v}_{(1)}\right],$$

where $\mathbf{S}_{(1)} = [\mathbf{X}_{(1)}, \mathbf{W}_{(1)}]$. Using the above results we have

$$\hat{\mathbf{d}}_{(1)} = \boldsymbol{\Lambda}^1_H \left[y_{T_0} \lambda^1 \mathbf{e}_1 + \mathbf{S}_{(1)} \boldsymbol{\pi}^1_{ys}\right] - \hat{\boldsymbol{\Lambda}}^0_H \left[y_{T_0} \hat{\lambda}^0 \mathbf{e}_1 + \tilde{\mathbf{S}}_{(1)} \hat{\boldsymbol{\pi}}^0_{ys}\right] + \boldsymbol{\Lambda}^1_H \mathbf{v}_{(1)},$$

which can be written as

$$\hat{\mathbf{d}}_{(1)} = \boldsymbol{\mu}_{(1)} - \boldsymbol{\xi}_{(1)} + \boldsymbol{\Lambda}^1_H \mathbf{v}_{(1)}, \tag{60}$$

where

$$\boldsymbol{\mu}_{(1)} = y_{T_0} \left(\boldsymbol{\Lambda}^1_H \lambda^1 - \boldsymbol{\Lambda}^0_H \lambda^0\right) \mathbf{e}_1 + \left[\boldsymbol{\Lambda}^1_H \mathbf{S}_{(1)} \boldsymbol{\pi}^1_{ys} - \boldsymbol{\Lambda}^0_H \tilde{\mathbf{S}}_{(1)} \boldsymbol{\pi}^0_{ys}\right], \tag{61}$$

and

$$\boldsymbol{\xi}_{(1)} = \hat{\boldsymbol{\Lambda}}^0_H \left[y_{T_0} \hat{\lambda}^0 \mathbf{e}_1 + \mathbf{S}_{(1)} \hat{\boldsymbol{\pi}}^0_{ys}\right] - \boldsymbol{\Lambda}^0_H \left[y_{T_0} \lambda^0 \mathbf{e}_1 + \mathbf{S}_{(1)} \boldsymbol{\pi}^0_{ys}\right]. \tag{62}$$

In the dynamic case, the implicit null of the policy ineffectiveness hypothesis is given by

$$\boldsymbol{\mu}_{(1)} = y_{T_0} \left(\boldsymbol{\Lambda}^1_H \lambda^1 - \boldsymbol{\Lambda}^0_H \lambda^0\right) \mathbf{e}_1 + \left[\boldsymbol{\Lambda}^1_H \mathbf{S}_{(1)} \boldsymbol{\pi}^1_{ys} - \boldsymbol{\Lambda}^0_H \mathbf{S}_{(1)} \boldsymbol{\pi}^0_{ys}\right] = \mathbf{0}. \tag{63}$$

The third term of (60), $\boldsymbol{\Lambda}^1_H \mathbf{v}_{(1)}$, is the vector of the random shocks during post-intervention period, and the implementation of the test of policy ineffectiveness hypothesis in the dynamic case requires making the additional assumption that under H_0, we also have $\lambda^1 = \lambda^0$, as well as $E(\mathbf{v}_{(1)} \mathbf{v}'_{(1)}) = \sigma^2_{0v} \mathbf{I}_H$, the assumption already made in the static case. Finally, $\boldsymbol{\xi}_{(1)}$ captures the effects of sampling uncertainty associated with the estimation λ^0 and $\boldsymbol{\pi}_{ys}$. In the dynamic case, the null hypothesis of policy ineffectiveness is given by

$$H_0{:}\boldsymbol{\mu}_{(1)}=\mathbf{0},\quad \sigma_{0v}^2=\sigma_{1v}^2,\quad \lambda^0=\lambda^1. \tag{64}$$

5.1. The distribution of the test statistic in the case where H is fixed

We now derive the asymptotic distribution of $\hat{\mathbf{d}}_{(1)}$ under $\mathcal{H}_0$, initially assuming that H is fixed. Under H_0

$$\hat{\mathbf{d}}_{(1)}=\boldsymbol{\Lambda}_H^1\mathbf{v}_{(1)}-\boldsymbol{\xi}_{(1)}, \tag{65}$$

and using (62) we have

$$\boldsymbol{\xi}_{(1)}=y_{T_0}\left(\hat{\lambda}^0-\lambda^0\right)\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\mathbf{e}_1+\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\tilde{\mathbf{S}}_{(1)}\left(\hat{\boldsymbol{\pi}}_{ys}^0-\boldsymbol{\pi}_{ys}^0\right)+y_{T_0}\lambda^0\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\mathbf{e}_1+\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\tilde{\mathbf{S}}_{(1)}\boldsymbol{\pi}_{ys}^0+y_{T_0}\left(\hat{\lambda}^0-\lambda^0\right)\boldsymbol{\Lambda}_H^0\mathbf{e}_1$$
$$+\boldsymbol{\Lambda}_H^0\tilde{\mathbf{S}}_{(1)}\left(\hat{\boldsymbol{\pi}}_{ys}^0-\boldsymbol{\pi}_{ys}^0\right).$$

Also estimating the dynamic regression model, (54), by least squares under standard assumption we have

$$\hat{\lambda}^0=\lambda^0+a_TT^{-1/2},\quad\text{and}\quad \hat{\boldsymbol{\pi}}_{ys}^0=\boldsymbol{\pi}_{ys}^0+\mathbf{b}_TT^{-1/2}, \tag{66}$$

where a_T and $\mathbf{b}_T$ are random variables bounded in T. Hence, under H_0

$$\hat{\mathbf{d}}_{(1)}=\boldsymbol{\Lambda}_H^1\mathbf{v}_{(1)}-y_{T_0}\lambda^0\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\mathbf{e}_1-y_{T_0}\left(\hat{\lambda}^0-\lambda^0\right)\boldsymbol{\Lambda}_H^0\mathbf{e}_1-\left(\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0\right)\mathbf{S}_{(1)}\boldsymbol{\pi}_{ys}^0-\boldsymbol{\Lambda}_H^0\mathbf{S}_{(1)}\left(\hat{\boldsymbol{\pi}}_{ys}^0-\boldsymbol{\pi}_{ys}^0\right)+O_p(T^{-1}). \tag{67}$$

Using a Taylor series expansion we have (for H fixed)

$$\hat{\boldsymbol{\Lambda}}_H^0-\boldsymbol{\Lambda}_H^0=\frac{\partial\boldsymbol{\Lambda}_H^0}{\partial\lambda^0}\left(\hat{\lambda}^0-\lambda^0\right)+O_p\left(\frac{1}{T}\right),$$

where

$$\frac{\partial\boldsymbol{\Lambda}_H^0}{\partial\lambda^0}=\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0\\ 1 & 0 & \cdots & 0 & 0 & 0\\ 2\lambda_0 & 1 & \cdots & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ (H-2)\left(\lambda^0\right)^{H-3} & (H-3)\left(\lambda^0\right)^{H-4} & \cdots & 1 & 0 & 0\\ (H-1)\left(\lambda^0\right)^{H-2} & (H-2)\left(\lambda^0\right)^{H-3} & \cdots & 2\lambda_0 & 1 & 0 \end{pmatrix}.$$

Using this result in (67) now yields

$$\hat{\mathbf{d}}_{(1)}=\boldsymbol{\Lambda}_H^1\mathbf{v}_{(1)}-\tilde{\mathbf{D}}_{(1)}\left(\hat{\lambda}^0-\lambda^0\right)-\boldsymbol{\Lambda}_H^0\tilde{\mathbf{S}}_{(1)}\left(\hat{\boldsymbol{\pi}}_{ys}^0-\boldsymbol{\pi}_{ys}^0\right)+O_p\left(\frac{1}{T}\right),$$

where

$$\tilde{\mathbf{D}}_{(1)}=y_{T_0}\left(\lambda^0\frac{\partial\boldsymbol{\Lambda}_H^0}{\partial\lambda^0}+\boldsymbol{\Lambda}_H^0\right)\mathbf{e}_1+\frac{\partial\boldsymbol{\Lambda}_H^0}{\partial\lambda^0}\tilde{\mathbf{S}}_{(1)}\boldsymbol{\pi}_{ys}^0.$$

Writing the above result more compactly, we have

$$\hat{\mathbf{d}}_{(1)}=\boldsymbol{\Lambda}_H^1\mathbf{v}_{(1)}-\boldsymbol{\Psi}_{(1)}^0\left(\hat{\varphi}^0-\boldsymbol{\varphi}^0\right)+O_p\left(\frac{1}{T}\right),$$

where $\boldsymbol{\Psi}_{(1)}^0=\left(\tilde{\mathbf{D}}_{(1)},\boldsymbol{\Lambda}_H^0\tilde{\mathbf{S}}_{(1)}\right)$.

When H is fixed and T, sufficiently large, a test can be based on all the elements of $\hat{\mathbf{d}}_{(1)}$ if $\mathbf{v}_{(1)}$ has a known distribution. This corresponds to the test statistic (36) in the static case. But in general, we do not know the distribution of $\mathbf{v}_{(1)}$ and, as in the static case, we need to base the test of policy ineffectiveness on some average of $\hat{\mathbf{d}}_{(1)}$. Again using the policy mean effect statistic, $\overline{\hat{d}}_H$, defined by (40), under H_0 defined by (64) we have

$$\sqrt{H}\overline{\hat{d}}_H=\frac{1}{\sqrt{H}}\boldsymbol{\tau}_H'\boldsymbol{\Lambda}_H^0\mathbf{v}_{(1)}-\frac{1}{\sqrt{H}}\boldsymbol{\tau}_H'\boldsymbol{\Psi}_{(1)}^0\left(\hat{\boldsymbol{\varphi}}^0-\boldsymbol{\varphi}^0\right)+O_p\left(\frac{\sqrt{H}}{T}\right). \tag{68}$$

Note that under the null hypothesis, $\left(\hat{\boldsymbol{\varphi}}^0-\boldsymbol{\varphi}^0\right)=O_p(T^{-1/2})$, and for a finite H and a sufficiently large T the distribution of $\sqrt{H}\overline{\hat{d}}_H$ is determined by the distribution of $H^{-1/2}\boldsymbol{\tau}_H'\boldsymbol{\Lambda}_H^0\mathbf{v}_{(1)}$. In the case where $\mathbf{v}_{(1)}$ is normally distributed, we can use the following test statistic:

$$\mathcal{T}^a_{d,H} = \frac{\sqrt{H}\hat{\bar{d}}_H}{\hat{\sigma}_{0v}\left\{\left(\frac{\boldsymbol{\tau}'_H\hat{\boldsymbol{\Lambda}}^0_H\hat{\boldsymbol{\Lambda}}^{0\prime}_H\boldsymbol{\tau}_H}{H}\right)+\frac{1}{T}\left[\frac{\boldsymbol{\tau}'_H\hat{\boldsymbol{\Psi}}^0_{(1)}\left(T^{-1}\mathbf{Q}_{(0)}\mathbf{Q}'_{(0)}\right)^{-1}\hat{\boldsymbol{\Psi}}^{0\prime}_{(1)}\boldsymbol{\tau}'_H}{H}\right]\right\}^{1/2}} \to_d N(0,1), \tag{69}$$

where $\hat{\lambda}^0$ and $\hat{\sigma}_{0v}$ are the least squares estimates of λ and σ_v based on the pre-intervention sample,

$$\hat{\boldsymbol{\Psi}}^0_{(1)} = \left[y_{T_0}\left(\hat{\lambda}^0\frac{\partial\hat{\boldsymbol{\Lambda}}^0_H}{\partial\lambda^0}+\hat{\boldsymbol{\Lambda}}^0_H\right)\mathbf{e}_1+\frac{\partial\hat{\boldsymbol{\Lambda}}^0_H}{\partial\lambda^0}\tilde{\mathbf{S}}_{(1)}\hat{\boldsymbol{\pi}}^0_{ys},\ \hat{\boldsymbol{\Lambda}}^0_H\tilde{\mathbf{S}}_{(1)}\right], \tag{70}$$

and $\mathbf{Q}_{(0)} = \left(\mathbf{y}_{-1,(0)}, \mathbf{S}_{(0)}\right)$. In the case where T is reasonably large relative to H, the second term in the denominator of (69) will be negligible and the test statistic simplifies to

$$\mathcal{T}^a_{d,H} = \frac{\sqrt{H}\hat{\bar{d}}_H}{\hat{\sigma}_{0v}\left(\frac{\boldsymbol{\tau}'_H\hat{\boldsymbol{\Lambda}}^0_H\hat{\boldsymbol{\Lambda}}^{0\prime}_H\boldsymbol{\tau}_H}{H}\right)^{1/2}} \to_d N(0,1), \tag{71}$$

where $\hat{\sigma}_{0v}$ is the estimate of σ_{0v} computed using the pre-intervention sample, and

$$\frac{\boldsymbol{\tau}'_H\hat{\boldsymbol{\Lambda}}^0_H\hat{\boldsymbol{\Lambda}}^{0\prime}_H\boldsymbol{\tau}_H}{H} = \frac{1}{\left(1-\hat{\lambda}^0\right)^2}\left\{1-\frac{2}{H}\left[\frac{\left(\hat{\lambda}^0\right)^{H+1}-\hat{\lambda}^0}{1-\hat{\lambda}^0}\right]+\frac{1}{H}\left[\frac{\left(\hat{\lambda}^0\right)^{2H+2}-\left(\hat{\lambda}^0\right)^2}{\left[1-\left(\hat{\lambda}^0\right)^2\right]}\right]\right\}. \tag{72}$$

5.2. The distribution of the test statistic in the case where H is sufficiently large

In the case where H rises with T, the derivations are best carried out in terms of the individual elements of $\hat{\mathbf{d}}_{(1)}$ defined by (60), which we write as

$$\hat{d}_{T_0+h} = -\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^1\right)^h\right]y_{T_0}-\sum_{j=0}^{h-1}\left[\left(\hat{\lambda}^0\right)^h\hat{\boldsymbol{\pi}}^0_{ys}-\left(\lambda^1\right)^h\boldsymbol{\pi}^1_{ys}\right]'\mathbf{s}_{T_0+h-j}+\sum_{j=0}^{h-1}\left(\lambda^1\right)^h v_{T_0+h-j},$$

where $\mathbf{s}_t = (x_t, x_{t-1}, \mathbf{w}_t, \mathbf{w}_{t-1})'$. The mean policy effect test statistic can now be written as

$$\sqrt{H}\hat{\bar{d}}_H = -H^{-1/2}\sum_{h=1}^{H}\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^1\right)^h\right]y_{T_0}-H^{-1/2}\sum_{h=1}^{H}\sum_{j=0}^{h-1}\left[\left(\hat{\lambda}^0\right)^h\hat{\boldsymbol{\pi}}^0_{ys}-\left(\lambda^1\right)^h\boldsymbol{\pi}^1_{ys}\right]'\mathbf{s}_{T_0+h-j}+H^{-1/2}\sum_{h=1}^{H}\sum_{j=0}^{h-1}\left(\lambda^1\right)^h v_{T_0+h-j}. \tag{73}$$

Under the null hypothesis $\lambda^1 = \lambda^0$, $\boldsymbol{\pi}^1_{ys} = \boldsymbol{\pi}^0_{ys}$, and $\sigma^2_{1v} = \sigma^2_{0v}$, we first note that

$$\sum_{h=1}^{H}\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^0\right)^h\right] = \frac{\hat{\lambda}^0-\lambda^0+\hat{\lambda}^0\lambda^0\left[\left(\hat{\lambda}^0\right)^H-\left(\lambda^0\right)^H\right]-\left[\left(\hat{\lambda}^0\right)^{H+1}-\left(\lambda^0\right)^{H+1}\right]}{(1-\lambda^0)(1-\hat{\lambda}^0).}$$

Also using results in Lemma 3 in PS, and since $\hat{\lambda}^0-\lambda^0 = a^0_T/\sqrt{T}$, $(\lambda^0 \neq 0)$

$$H^{-1/2}\left|\left(\hat{\lambda}^0\right)^H-\left(\lambda^0\right)^H\right| \le H^{1/2}\left|\hat{\lambda}^0\right|^{H-1}\left|\hat{\lambda}^0-\lambda^0\right| \le \left(\frac{H}{T}\right)^{1/2}|a^0_T|\left|\lambda^0\right|^H\left|1+\frac{a^0_T}{\lambda^0\sqrt{T}}\right|^H,$$

where $|a^0_T|$ is bounded in T. But $\left|1+a^0_T/\lambda^0\sqrt{T}\right|^H$ tends to a bounded random variable if $H/\sqrt{T}$ tends to a fixed constant, or equivalently if $H = \kappa T^\epsilon$, with $\epsilon \le 1/2$ as T and $H \to \infty$, jointly. Under this condition, we have $H^{-1/2}\left|\left(\hat{\lambda}^0\right)^H-\left(\lambda^0\right)^H\right| \to_p 0$ since $\left|\lambda^0\right| < 1$. Similarly, under the null hypothesis

$$H^{-1/2}\sum_{h=1}^{H}\sum_{j=0}^{h-1}\left[\left(\hat{\lambda}^0\right)^h\hat{\boldsymbol{\pi}}^0_{ys}-\left(\lambda^1\right)^h\boldsymbol{\pi}^1_{ys}\right]'\mathbf{s}_{T_0+h-j} = H^{-1/2}\sum_{h=1}^{H}\sum_{j=0}^{h-1}\left\{\begin{array}{c}\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^0\right)^h\right]\left(\hat{\boldsymbol{\pi}}^0_{ys}-\boldsymbol{\pi}^0_{ys}\right)+\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^0\right)^h\right]\boldsymbol{\pi}^0_{ys}\\+\left(\lambda^0\right)^h\left(\hat{\boldsymbol{\pi}}^0_{ys}-\boldsymbol{\pi}^0_{ys}\right)\end{array}\right\}'\mathbf{s}_{T_0+h-j},$$

we note that (using Lemma 1 in PS)

$$H^{-1/2}\sum_{h=1}^{H}\sum_{j=0}^{h-1}\left[\left(\hat{\lambda}^0\right)^h-\left(\lambda^0\right)^h\right]\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+h-j}=\left[\frac{H^{1/2}\left(\hat{\lambda}^0-\lambda^0\right)}{\left(1-\lambda^0\right)\left(1-\hat{\lambda}^0\right)}\right]\left(H^{-1}\sum_{j=1}^{H}\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right)$$
$$-\left\{H^{-1/2}\sum_{j=1}^{H}\left[\left(1-\hat{\lambda}^0\right]^{-1}\left(\hat{\lambda}^0\right)^{H-j+1}-\left(1-\lambda^0\right)^{-1}\left(\lambda^0\right)^{H-j+1}\right)\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right\}.$$

Since by assumption $\left|H^{-1}\sum_{j=1}^{H}\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right|<K$, and $\hat{\lambda}^0-\lambda^0=O_p(T^{-1/2})$, then the first term of the above expression tends to zero in probability of $H/T\to 0$, which is satisfied if $H=\kappa T^{\epsilon}$, with $\epsilon\le 1/2$. Consider the second term of the above expression and note that

$$\left|H^{-1/2}\sum_{j=1}^{H}\left[\left(1-\hat{\lambda}^0\right)^{-1}\left(\hat{\lambda}^0\right]^{H-j+1}-\left(1-\lambda^0\right)^{-1}\left(\lambda^0\right)^{H-j+1}\right)\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right|\le\sup_j\left|\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right|H^{-1/2}\sum_{j=1}^{H}\left|\left(1-\hat{\lambda}^0\right)^{-1}\left(\hat{\lambda}^0\right)^{H-j+1}\right.$$
$$\left.-\left(1-\lambda^0\right)^{-1}\left(\lambda^0\right)^{H-j+1}\right|\le\frac{\sup_j\left|\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right|}{\left|(1-\lambda^0)(1-\hat{\lambda}^0)\right|}\left[H^{1/2}\sum_{j=1}^{H}\left|\left(\hat{\lambda}^0\right)^j-\left(\lambda^0\right)^j\right|+\left|\lambda^0\right|\left|\hat{\lambda}^0\right|H^{1/2}\sum_{j=1}^{H}\left|\left(\hat{\lambda}^0\right)^{j-1}-\left(\lambda^0\right)^{j-1}\right|\right].$$

But using results in Lemma 3 and 4 in PS we have

$$H^{1/2}\sum_{j=1}^{H}\left|(\hat{\lambda}^0)^j-\left(\lambda^0\right)^j\right|\le H^{1/2}\left|\hat{\lambda}^0-\lambda^0\right|\sum_{j=1}^{H}j\left|\hat{\lambda}^0\right|^{j-1}=H^{1/2}\left|\hat{\lambda}^0-\lambda^0\right|\left[\frac{1-\left|\hat{\lambda}^0\right|^H}{\left(1-\left|\hat{\lambda}^0\right|\right)^2}-\frac{H\left|\hat{\lambda}^0\right|^{H+1}}{\left(1-\left|\hat{\lambda}^0\right|\right)}\right],$$

and using similar arguments as before it follows that

$$H^{1/2}\sum_{j=1}^{H}\left|(\hat{\lambda}^0)^j-\left(\lambda^0\right)^j\right|\to_p 0,$$

if $\sup_j\left|\boldsymbol{\pi}_{ys}^{0\prime}\mathbf{s}_{T_0+j}\right|$ is bounded in T, and $H=\kappa T^{\epsilon}$, with $\epsilon\le 1/2$, as T and $H\to\infty$, jointly. Notice that since, by assumption, both $\hat{\lambda}^0$ and λ^0 are less than one in absolute value, $\left|\left(\hat{\lambda}^0\right)^j-\left(\lambda^0\right)^j\right|$ declines exponentially in j.

Under these conditions and employing the above results in (73), and using (72) with $H\to\infty$, we finally obtain the large T and H test statistic

$$\mathcal{T}_{d,H}^{b}=\left|1-\hat{\lambda}^0\right|\sqrt{H}\frac{\bar{\hat{d}}_H}{\hat{\sigma}_{0v}}\to_d N(0,1). \tag{74}$$

It is easily verified that the test based on $\mathcal{T}_{d,H}^{b}$ is equivalent to the test based on $\mathcal{T}_{d,H}^{a}$ given by (71), when H is sufficiently large.

6. An empirical application: testing the effects of quantitative easing

We will illustrate the policy-ineffectiveness test with an investigation into the effect of an unconventional monetary policy (UMP), known as quantitative easing (QE), in the UK introduced in March 2009.[6] We will use a final form approach and an ARDL (1,1) model, as in Section 5. Following the literature, such as Kapetanios et al. (2012), we treat QE as an intervention that reduced the spread between the long and short term interests by 100 basis points. Thus it is a discretionary policy intervention of the type discussed in Section 4.1. Note that our approach allows for the policy change to alter the parameters of the underlying model and the Lucas Critique does not apply to our test since it is computed under the null of no policy change. Policy induced parameter changes only affect the power of our test, and do not have any consequence for the properties of the policy ineffectiveness test

[6] See, also, the November 2012 Special Issue of the *Economic Journal* on “Unconventional Monetary Policy after the Financial Crisis”.

under the null hypothesis. This is clear from (44), which shows that the power depends on the average size of the change in the policy variable and the policy induced changes in the coefficients of the exogenous variables.

UMPs have tended to be adopted when central banks have hit the zero lower bound for the policy interest rate, though in principle they could be adopted even if interest rates were not at the lower bound. The term quantitative easing was used by the Bank of Japan to describe its policies from 2001. See, for example, Bowman et al. (2011). During the financial crisis, starting in 2007, and particularly after the failure of Lehman Brothers in 2008, many central banks adopted UMP. Examples include the large scale asset purchase programme by Federal Reserve in the US and the long term repo operations and emergency liquidity assistance by the European Central Bank. The central banks differed in the specific measures used and had different theoretical perceptions of what the policy interventions were designed to achieve and the transmission mechanisms involved.[7] Borio and Disayatat (2010) classify such policies as balance sheet policies, as distinct from interest rate policies, and describe the variety of different types of measures adopted by seven central banks during the financial crisis. There has been considerable controversy over two questions: (i) what were the effects of UMP on various interest rates? usually answered using "event studies" and (ii) what was the effect of those interest rate changes on real output and inflation? We shall consider question (ii) to illustrate our proposed test taking the answer to (i) as given.

In the UK, QE involved exchanging one liability of the state – government bonds (gilts) – for another – claims on the central bank. That change in the quantities of the two assets could not only cause a rise in the price of gilts and a decline in their yields, but also cause a rise in the prices of substitute assets such as corporate bonds and equities. The Bank of England believed that QE boosted demand by increasing wealth and by reducing the cost of finance to companies.[8]

Event studies documented in Joyce et al. (2011) suggest that QE reduced the spread of long over short term government interest rates (the "spread") by 100 basis points from its introduction in March 2009.[9] Thus, the counterfactual we consider is the effect on log real output, Y_t, of there not having been a 100 basis points reduction in the spread. The estimate that QE reduced the spread by 100 basis points is not uncontroversial, Meaning and Zhu (2011) estimate a smaller impact of about 25 basis points, but our estimates could be easily scaled downwards to match this alternative estimate. Like other studies, we treat this as a discretionary policy change of the type discussed in Section 4.1, where there is a deterministic change in a policy variable, the spread in this case. Thus our counterfactual value for the policy variable is of the form $\tilde{x}_{T_0+h} = x_{T_0+h} - \delta$, where δ is a constant (set to 100 basis points), so that the variance of the policy implementation errors, which appears in (53), is zero.

In examining QE, we model the growth rate of output, $y_t = Y_t - Y_{t-1}$, because log output appears to have a unit root (and there is no long-run relationship between log real output and the spread). The test is then be based on (40), a mean policy effect computed over the post-intervention horizon $T_0 + h$, for $h = 1, 2, \ldots, H$, namely

$$\overline{\hat{d}}_H = \frac{1}{H} \sum_{h=1}^{H} \hat{d}_{T_0+h},$$

where

$$\hat{d}_{T_0+h} = y_{T_0+h} - \hat{\tilde{y}}_{T_0+h}, \quad h = 1, 2, \ldots, H.$$

However, our analysis still applies to log output. To see this, note that

$$\begin{aligned}\overline{\hat{d}}_H &= H^{-1}\left[\sum_{h=1}^{H}\left(y_{T_0+h} - \hat{\tilde{y}}_{T_0+h}\right)\right] = H^{-1}\left[(Y_{T_0+H} - Y_{T_0+H-1}) + (Y_{T_0+H-1} - Y_{T_0+H-2}) + \cdots + (Y_{T_0+1} - Y_{T_0})\right] \\ &\quad - H^{-1}\left[\left(\hat{\tilde{Y}}_{T_0+H} - \hat{\tilde{Y}}_{T_0+H-1}\right) - \left(\hat{\tilde{Y}}_{T_0+H-1} - \hat{\tilde{Y}}_{T_0+H-2}\right) + \cdots + \left(\hat{\tilde{Y}}_{T_0+1} - Y_{T_0}\right)\right] = H^{-1}\left(Y_{T_0+H} - \hat{\tilde{Y}}_{T_0+H}\right).\end{aligned}$$

Thus $\overline{\hat{d}}_H$ measures the average effect of the policy on log real output over the given post policy horizon, H. The policy ineffectiveness test statistics are given by (74) or (71), depending whether H is sufficiently large.

Kapetanios et al. (2012) , who examine the effects of QE on UK output growth and inflation, also use a reduction in spread of 100 basis points. They use three time-varying vector autoregressions, VARs, that include other endogenous $\mathbf{y}_{2t}$ type variables and allow for parameter change in different ways. Baumeister and Benati (2013) also use time varying VARs to assess the macroeconomic effects of QE in the US and UK, assuming the effect of QE in the UK was to reduce the spread by 50 basis points. But as our theoretical analysis highlights, the effects of structural breaks due to factors other than the policy change must be distinguished from the structural breaks that could result from the policy intervention. Goodhart and Ashworth (2012) challenge the view that the official long rate is the proper measure of the effect of QE on the economy, and argue that the transmission was through other variables such as credit risk spreads. We do not rule out that QE might have had an impact on other such variables, the $\mathbf{y}_{2t}$ in our notation, with subsequent effects on output growth, but, as explained above, such effects are allowed for in our approach, albeit indirectly. Goodhart and Ashworth (2012) also argue that external effects are important and we allow for these effects through the inclusion of foreign variables as candidates for the policy invariant variables, $\mathbf{w}_t$.

[7] For instance Giannone et al. (2011), who discuss the euro area, distinguish the Eurosystem's actions from the QE adopted by other Central Banks.
[8] For instance, see the Financial Times article 4 May 2012 by Charlie Bean, then the Bank's Deputy Governor.
[9] Subsequent asset purchases seem to have been anticipated and do not seem to have such a clear effect on the spread.

Table 1
ARDL in UK growth (y) and spread (x) augmented with US and Euro area growth rates, t ratios in parentheses.

Sample	1980Q3–2008Q4	1980Q3–2011Q2
y_{t-1}	0.34688	0.3822
	(3.94)	(5.12)
Δx_t	−0.94559	−0.8197
	(−3.05)	(−2.63)
x_{t-1}	–	0.18960
	–	(1.61)
y_t^{US}	0.15509	0.1465
	(2.07)	(2.00)
y_t^{Euro}	0.11040	0.1636
	(1.89)	(3.047)
$\overline{R}^2$	0.333	0.446
LM test res. serial corr.	0.414	0.332
$\hat{\sigma}_v$	0.0051	0.0050

The data are taken from the Global VAR data set, starting in 1979Q2 and ending in 2011Q2.[10] Growth, y_t, is measured by the quarterly change in the logarithm of real GDP. In calculating the spread between the short and long government interest rates, the rates are expressed as $0.25\log(1+R_t/100)$, where R_t is the annual percent rate. For the conditioning variables, $\mathbf{w}_t=(y_t^{US},y_t^{Euro})'$, we use US and euro area output growth as they are unlikely to have been significantly affected by UK QE, but their inclusion allows for the possible indirect effects of UMPs implemented in US and euro area on UK output growth. Over the full sample, the correlation between UK growth and US growth is 0.47, in the post 1999 sample it is 0.76. For euro growth, the correlations are 0.36 and 0.73. Like Kapetanios et al. (2012) we assume that the reduction in the spread is permanent. But other time profiles for the policy effects of QE on spreads could also be considered.

We use an ARDL in output growth (y_t) and the spread between long and short government interest rates (x_t) augmented by current euro and US growth rates. Pesaran and Shin (1999) show that ARDL estimates are robust to endogeneity and robust to the fact that y_t (stationary) and x_t (near unit root) have different degrees of persistence. Under the null of no policy change the analysis can be based on the full sample or the pre-policy change sample. Using the full sample 1980Q3–2011Q2, both AIC and SBC selected the lag order of 1. The ARDL equation for the pre-policy change period is given by (54),

$$y_t=\lambda^0 y_{t-1}+\pi_{yx0}^0 x_t+\pi_{yx1}^0 x_{t-1}+\boldsymbol{\pi}_{yw}^{0\prime}\mathbf{w}_t+v_{yt},\quad \text{for } t=1,2,\ldots,T_0.$$

The estimates pass the diagnostic tests for serial correlation and heteroskedasticity, but fail (at 5% level) tests of error normality and functional form. The restriction that it is the spread that matters, rather than short and long government interest rates separately, is not rejected (pval=0.23). We also find a positive long run effect of the spread on output growth which is implausible, but is statistically insignificant; the t-statistic for testing the long run restriction $\pi_{yx0}+\pi_{yx1}=0$ is $t=1.61$, which is not statistically significant. This restriction is imposed on the model used in obtaining the pre-policy estimates. The estimates for the full sample and the pre-policy sample, using the change in spread and the lagged spread, are shown in Table 1.

The estimates suggest that a permanent 100 basis points reduction in the spread increases predicted growth by almost 1% on impact, although this effect is quickly reversed and disappears altogether within two years. Although they do not emphasise this feature, the estimates of Kapetanios et al. (2012), tell very much the same story: the beneficial effects of QE on growth are of a similar size and rather short-lived. However, this predicted positive effect on growth of reducing the spread is small relative to the large negative equation errors, the estimated equation over-predicts growth during the recession. So the actual is below the counterfactual outcome without QE in all but one post-intervention quarter. For $H=10$, (2009Q1−2011Q2), $\overline{\hat{d}}_H=-0.00315$; and from Table 1, $\hat{\lambda}^0=0.347$ and $\hat{\sigma}_{0v}=0.0051$. So the test statistic (71)

$$\mathcal{T}_{d,H}^a=\frac{\sqrt{H}\overline{\hat{d}}_H}{\hat{\sigma}_{0v}\left(\frac{\boldsymbol{\tau}_H'\hat{\boldsymbol{\Lambda}}_H^0\hat{\boldsymbol{\Lambda}}_H^{0\prime}\boldsymbol{\tau}_H}{H}\right)^{1/2}},$$

is −1.17.

The small H adjustment used in (71) does not make very much difference and the test statistic without it (74)

$$\mathcal{T}_{d,H}^b=\frac{\left|1-\hat{\lambda}^0\right|\sqrt{H}\overline{\hat{d}}_H}{\hat{\sigma}_{0v}},$$

is −1.28. Also taking account of the sampling uncertainty associated with parameter estimation, captured by the second

[10] Described in Dees et al. (2007), with updates available at https://sites.google.com/site/gvarmodelling/.

term in the denominator of (69), does not alter the test outcome. Firstly, with $T=114$ this second term is likely to be small, and secondly given that it is positive its inclusion can only reduce the statistical significance of the test.

As a result, we conclude that given our model specification, the null that the QE policy intervention was ineffective cannot be rejected. Of course, we need to bear in mind that, as with all statistical tests, the null hypothesis being tested is a joint null, assuming that under the null hypothesis either no other major policy changes were put into effect, or such additional policy changes were also ineffective. Separating the effects of QE from other policy developments, such as the austerity measures that were put into effect by the Coalition Government in the UK would be difficult.

7. Conclusion

In this paper, we have derived tests for the null hypothesis of the ineffectiveness of a policy intervention using single equation reduced form and final form policy response equations. These are simpler to implement than complete structural models and tests using them could be more robust than the tests based on possibly misspecified complete systems of equations. We propose estimating an unrestricted final form policy equation which makes the target variable a function of lagged values of the target variable, as well as current and lagged values of the policy and policy-invariant exogenous variables. We consider both discretionary policy interventions and ones that involve changes in the parameters of a policy rule. This proposed final form analysis of policy can also be viewed as a prelude to a more complicated policy analysis using the DSGE framework advanced in PS.

The tests are based on the differences, over a given policy evaluation horizon, between the post-intervention realizations of the target variable and the associated counterfactual outcomes based on the parameters estimated using data before the policy intervention. The Lucas Critique is not an issue since the counterfactual, given by conditional predictions from the model estimated on pre-intervention data, will embody pre-intervention parameters, while the actual post-intervention outcomes will embody any effect of the change in policy, the change in parameters, and the consequent change in expectations. The proposed tests do not require knowing the post-intervention parameters.

We derive the asymptotic distribution of the policy ineffectiveness tests under alternative assumptions concerning the type of model, the future error processes and the pre and post-intervention sample sizes. We also develop a policy ineffectiveness test based on the mean policy effect which is robust to the distribution of future errors, but requires the post-intervention, policy evaluation horizon to be reasonably large. In the case of a static model, we also derive an exact test allowing for the estimation uncertainty and analyze the factors that determine the power of the test.

We use the test to assess the effectiveness of the quantitative easing which was introduced in the UK in March 2009. We follow the Bank of England in assuming that QE caused a permanent 100 basis points reduction in the spread of long interest rates over short interest rates after March 2009. Thus, we treat it as a discretionary deterministic intervention. We estimate models explaining UK output growth over two sample periods, one ending in 2008Q4 (before QE), and the other ending in 2011Q2. We use an ARDL(1,1) specification between output growth and the change in the spread of long government interest rates over short rates, augmented by current values of US and euro area output growth. The model indicates that QE had an immediate positive effect on growth, but this effect tended to disappear quite quickly, certainly within a year. The estimates of Kapetanios et al. (2012) for the time profiles of the effects of the QE tell very much the same story, namely the beneficial effects of QE are rather short-lived. This conclusion is supported by the outcomes of the policy ineffectiveness tests. We cannot reject the null of policy ineffectiveness, although we recognize the possible low power of such tests in the present application.

The counterfactuals and tests considered in this paper apply to a policy intervention for a single unit (country). It would be interesting to consider panel version of such tests for the analysis of policy across many different units that allow for dynamics and cross-sectional dependence.

Acknowledgments

We are grateful for comments made on earlier versions of this paper by Alex Chudik, Karrar Hussain, Oscar Jorda, and Ivan Petrella.

Appendix A

A.1. *Derivation of the distribution of $\mathcal{F}_{d,H}$ defined by (38)*

Under the null hypothesis $\mathcal{H}_0$ defined by (34)

$$\hat{\mathbf{d}}_{(1)} = \mathbf{v}_{(1)} - \mathbf{S}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)}\mathbf{v}_{(0)} = \mathbf{G}'\mathbf{v},$$

$$\mathbf{y}_{(0)}-\mathbf{S}_{(0)}\hat{\boldsymbol{\pi}}^{0}_{ys}=\left\{\mathbf{I}_T-\mathbf{S}_{(0)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}_{(0)}\right\}\mathbf{v}_{(0)}=\mathbf{Q}\mathbf{v},$$

where $\mathbf{v}=(\mathbf{v}'_{(0)},\mathbf{v}'_{(1)})'$, and

$$\mathbf{G}'=\left\{-\mathbf{S}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(0)},\ \mathbf{I}_H\right\},$$

$$\mathbf{Q}=\begin{pmatrix}\mathbf{I}_T-\mathbf{S}_{(0)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}_{(0)} & \mathbf{0}_{T\times H}\\ \mathbf{0}_{H\times T} & \mathbf{0}_{H\times H}\end{pmatrix}.$$

Using these results in $F_{d,H}$ defined by (38), we have

$$\mathcal{F}_{d,H}=\frac{T-1-k_w}{H}\frac{\hat{\mathbf{d}}'_{(1)}\left\{\mathbf{I}_H+\mathbf{S}_{(1)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}'_{(1)}\right\}^{-1}\hat{\mathbf{d}}_{(1)}}{\mathbf{v}'_{(0)}\left\{\mathbf{I}_H-\mathbf{S}_{(0)}\left[\mathbf{S}'_{(0)}\mathbf{S}_{(0)}\right]^{-1}\mathbf{S}_{(0)}\right\}\mathbf{v}_{(0)}}=\frac{T-1-k_w}{H}\frac{\boldsymbol{\xi}'\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\boldsymbol{\xi}}{\boldsymbol{\xi}'\mathbf{Q}\boldsymbol{\xi}},$$

where $\boldsymbol{\xi}=\mathbf{v}/\sigma_{0v}\sim N(\mathbf{0},\mathbf{I}_{T+H})$. It is also easily seen that $\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'$ and $\mathbf{Q}$ are idempotent matrices that are orthogonal (namely, $\mathbf{G}'\mathbf{Q}=\mathbf{0}$) with ranks H and $T-1-k_w$, which establish that $F_{d,H}$ has a F distribution with H and $T-1-k_w$ degrees of freedom.

A.2. Derivation of the distribution of $\mathcal{T}_{d,H}$ in the general dynamic case

Consider the general ARDL specification given by (12), with a common lag length p, to simplify the notation. The post-intervention model can be written as

$$\boldsymbol{\Gamma}^1_H\mathbf{y}_{(1)}=\boldsymbol{\Upsilon}^1_H\mathbf{y}^*_p+\mathbf{S}_{(1)}\boldsymbol{\pi}^1_{ys}+\mathbf{v}_{(1)},$$

where $\mathbf{y}^*_p$ is the $H\times 1$ vector containing the p initial observations, $\mathbf{y}^*_p=(y_{T_0},y_{T_0-1},\dots,y_{T_0-p+1},0,\dots,0)'$, $\boldsymbol{\Gamma}^1_H=\boldsymbol{\Gamma}(\boldsymbol{\lambda}^1)$, $\boldsymbol{\Upsilon}^1_H=\boldsymbol{\Upsilon}_H(\boldsymbol{\lambda}^1)$, $\boldsymbol{\lambda}=(\lambda_1.\lambda_2,\dots,\lambda_p)'$,

$$\boldsymbol{\Gamma}_H(\boldsymbol{\lambda})=\begin{pmatrix}1&0&0&0&0&0&0&0\\-\lambda_1&1&0&0&0&0&0&0\\-\lambda_2&-\lambda_1&1&0&0&0&0&0\\\vdots&\vdots&\vdots&\vdots&0&0&0&0\\-\lambda_p&-\lambda_{p-1}&\cdots&1&0&0&0&0\\0&0&\cdots&0&0&0&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&-\lambda_p&-\lambda_{p-1}&\cdots&1&0\\0&0&\cdots&0&-\lambda_p&\cdots&-\lambda_1&1\end{pmatrix},$$

$$\boldsymbol{\Upsilon}_H(\boldsymbol{\lambda})=\begin{pmatrix}\lambda_1&\lambda_2&\cdots&0&\lambda_p\\\lambda_2&\lambda_3&\cdots&\lambda_p&0\\\vdots&\vdots&\cdots&\vdots&\vdots\\\lambda_p&0&\cdots&0&0\\0&0&\cdots&0&0\\\vdots&\vdots&\cdots&\vdots&\vdots\\0&0&\cdots&0&0\end{pmatrix},$$

and as before, $\mathbf{S}_{(1)}$ represents the $H\times k$ matrix of post-intervention observations on the current and lagged exogenous variables, $\mathbf{x}_t$ and $\mathbf{w}_t$, with $k=(1+p)(k_x+k_w)$. Also

$$\boldsymbol{\Gamma}^{-1}_H(\boldsymbol{\lambda})=\boldsymbol{\Lambda}_H(\boldsymbol{\lambda})=\begin{pmatrix}a_0&0&0&0&0\\a_1&a_0&0&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\\a_{H-2}&a_{H-3}&\cdots&a_0&0\\a_{H-1}&a_{H-2}&\cdots&a_1&a_0\end{pmatrix},\tag{75}$$

where α_h is obtained recursively using the difference equation

$$a_h=\sum_{i=1}^{p}\lambda_i a_{h-i},\quad h=1,2,\dots,H,$$

with $a_0=1$ and $a_i=0$, for all $i<0$. It is easily verified that for $p=1$, $\boldsymbol{\Lambda}_H(\boldsymbol{\lambda})$ defined by (75) reduces to the matrix defined by (57). Using the above set up, the estimated counterfactual outcomes are given by

$$\hat{\mathbf{y}}^0_{(1)}=\boldsymbol{\Lambda}_H(\hat{\boldsymbol{\lambda}}^0)\left[\boldsymbol{\Upsilon}_H(\hat{\boldsymbol{\lambda}}^0)\mathbf{y}^*_p+\tilde{\mathbf{X}}_{(1)}\hat{\boldsymbol{\pi}}^0_{yx}+\mathbf{W}_{(1)}\hat{\boldsymbol{\pi}}^0_{yw},\right]=\hat{\boldsymbol{\Lambda}}^0_H\left[\hat{\boldsymbol{\Upsilon}}^0_H\mathbf{y}^*_p+\tilde{\mathbf{S}}_{(1)}\hat{\boldsymbol{\pi}}^0_{ys}\right],$$

where $\hat{\boldsymbol{\varphi}}^0=\left(\hat{\boldsymbol{\lambda}}^{0'},\hat{\boldsymbol{\pi}}^{0\prime}_{ys}\right)'$ is obtained estimating the ARDL regression using the pre-intervention sample. The mean policy ineffectiveness test can be defined as before, $\mathcal{T}_{d,H}=\sqrt{H}\hat{\bar{d}}_H/\hat{\omega}_0$, where $\hat{\bar{d}}_H=H^{-1}\boldsymbol{\tau}'_H\hat{\mathbf{d}}_{(1)}$, $\hat{\mathbf{d}}_{(1)}=\mathbf{y}_{(1)}-\hat{\mathbf{y}}^0_{(1)}$,

$$\hat{\omega}^2_0=\hat{\sigma}^2_{0v}\left[H^{-1}\boldsymbol{\tau}'_H\boldsymbol{\Lambda}_H(\hat{\boldsymbol{\lambda}}^0)\boldsymbol{\Lambda}'_H(\hat{\boldsymbol{\lambda}}^0)\boldsymbol{\tau}_H\right],$$

and $\hat{\sigma}^2_{0v}$ is the estimate of σ^2_{0v} based on the pre-intervention sample. Following the same line of reasoning as in Section 5, it follows that $\mathcal{T}_{d,H}\to_d N(0,1)$, under the null hypothesis of policy ineffectiveness (including the restrictions $\boldsymbol{\lambda}^0=\boldsymbol{\lambda}^1$ and $\sigma^2_{0v}=\sigma^2_{1v}$), and as T and $H\to\infty$, jointly such that $H=\kappa T^{\epsilon}$, with $\epsilon\geq 1/2$.

References

Baumeister, C., Benati, L., 2013. Unconventional monetary policy and the great recession: estimating the macroeconomic effects of spread compression at the zero lower bound. Int. J. Central Bank. 9, 165–212.

Berk, K.N., 1974. Consistent autoregressive spectral estimates. Ann. Stat. 2, 489–502.

Borio, C., Disayatat, P., 2010. Unconventional monetary policies: an appraisal. Manchester School, 53–89. Supplement.

Bowman, D., Cai, F., Davies, S., Kamin, S., 2011. Quantitative Easing and Bank Lending: evidence from Japan, Board of Governors of the Federal Reserve System, International Finance Discussion Papers No. 1018.

Dees, S., di Mauro, F., Pesaran, M.H., Smith, L.V., 2007. Exploring the international linkages of the euro area: a global VAR analysis. J. Appl. Economet. 22, 1–38.

Giannone, D., Lenza, M., Pill, H., Reichlin, L., 2011. Non-Standard Monetary Policy Measures and Monetary Developments. ECB Working Paper No. 1290.

Goodhart, C.A.E., Ashworth, J.P., 2012. QE: a successful start may be running into diminishing returns. Oxford Rev. Econ. Policy 28, 640–670.

Joyce, M.A.S., Lasaosa, A., Stevens, I., Tong, M., 2011. The financial market impact of Quantitative Easing in the UK. Int. J. Central Bank. 7, 113–161.

Kapetanios, G., Mumtaz, H., Stevens, I., Theodoridis, K., 2012. Assessing the economy wide effects of quantitative easing. Econ. J. 122, F316–F347.

Meaning, J., Zhu, F., 2011. The impact of recent central bank asset purchase programs. BIS Q. Rev., 73–83. December.

Pesaran, M.H., Shin, Y., 1999. An autoregressive distributed-lag modelling approach to cointegration analysis. In: Strom, S. (Ed.), Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, Cambridge University Press, Cambridge, pp. 371–413. Chapter 11.

Pesaran, M.H., Smith, R.P., 2014. Signs of impact effects in time series regression models. Econ. Lett. 122, 150–153.

Pesaran, M.H., Smith, R.P., 2015. Tests of Policy Ineffectiveness in Macroeconometrics, Unpublished Manuscript, September.

Zellner, A., Palm, F., 1974. Time series analysis and simultaneous equation econometric models. In: Zellner, A., Palm, F.C. (Eds.), The Structural Econometric Time Series Analysis Approach, Cambridge University Press, Cambridge.