Granger Causality, Exogeneity, Cointegration, and Economic Policy Analysis

Halbert White
Department of Economics
University of California, San Diego

Davide Pettenuzzo
Department of Economics
Brandeis University

October 31, 2011

Abstract

Policy analysis had long been a main interest of Clive Granger's. Here, we present a framework for economic policy analysis that provides a novel integration of several fundamental concepts at the heart of Granger's contributions to time-series analysis. We work with a dynamic structural system analyzed by White and Lu (2010) with well defined causal meaning; under suitable conditional exogeneity restrictions, Granger causality coincides with this structural notion. The system contains target and control subsystems, with possibly integrated or cointegrated behavior. We ensure the invariance of the target subsystem to policy interventions using an explicitly causal partial equilibrium recursivity condition. Policy effectiveness is ensured by another explicit causality condition. These properties only involve the data generating process; models play a subsidiary role. Our framework thus complements that of of Ericsson, Hendry, and Mizon (1998) (EHM) by providing conditions for policy analysis alternative to weak, strong, and super-exogeneity. This makes possible policy analysis for systems that may fail EHM's conditions. It also facilitates analysis of the cointegrating properties of systems subject to policymaker control. We discuss a variety of practical procedures useful for analyzing such systems and illustrate with an application to a simple model of the U.S. macroeconomy.

1 Introduction

Although just three of Clive Granger's many papers explicitly focus on aspects of policy analysis (Granger, 1973; Granger, 1988; and Granger and Deutsch, 1992), a central and long-standing concern evident throughout his work is that econometric theory and practice should be informative and useful to policymakers. In this paper, we further this objective by providing a novel framework for economic policy analysis that blends together a number of concepts at the heart of Granger's contributions to time-series econometrics: causality, exogeneity, cointegration, and model specification.

We study a dynamic structural system with potentially cointegrated variables analyzed by White and Lu (2010) (WL) within which causal meanings are well defined. This system contains target and control subsystems, with possibly integrated or cointegrated behavior. We ensure the invariance of the target subsystem to policy interventions, obviating the Lucas critique, using an explicitly causal partial equilibrium recursivity condition. Another causality requirement ensures policy effectiveness. Causal effects are identified by a conditional form of exogeneity. These effects can be consistently estimated with a correctly specified model.

Following WL, we show that, given conditional exogeneity, Granger causality is equivalent to structural causality. On the other hand, given structural *non*-causality, Granger causality is equivalent to *failure* of conditional exogeneity. In this sense, Granger causality is not a fundamental system property requisite for reliable policy analysis, but an important consequence of necessary underlying structural properties.

By relying only on correct model specification and not weak exogeneity or its extensions (strong and superexogeneity), our framework complements the policy analytic framework of Ericsson, Hendry, and Mizon (1998) (EHM). Although giving up weak exogeneity may lead to loss of estimator efficiency, it also makes possible policy analysis for systems that may fail EHM's conditions (see Fisher, 1993). As we also show, our approach readily lends itself to analysis of the structural consequences of a variety of control rules that the policymaker may employ. Among other things, we find that proportional (P) control cannot modify the cointegrating properties of a target system, whereas proportional-integral (PI) control can. In fact, PI control can introduce, eliminate, or broadly modify the cointegrating properties of the uncontrolled target system. Whereas cointegration between target variables and policy instruments is possible but unusual with P control, PI control can easily induce causal cointegration between the target variables (Y_t) and the policy instruments (ΔZ_t).

The control mode also has interesting implications for estimation, inference, and specification testing in controlled systems. P control or a certain mode of PI control yields $\Delta Z_t \sim I(0)$, resulting in standard inference. Other modes of PI control yield $\Delta Z_t \sim I(1)$; the theory of Park and Phillips (1988, 1989) may be applied to these cases.

The plan of the paper is as follows. In Section 2, we introduce the data generating process (DGP) for the controlled system we study here, together with notions of structural causality and policy interventions natural in these systems. These enable us to formulate causal restrictions, essential for reliable policy analysis, obviating the Lucas critique and ensuring policy effectiveness. Section 3 discusses a conditional form of exogeneity that identifies causal effects of interest and forges links between structural and Granger causality. Section 4 reviews properties of relevant cointegrated systems, with particular attention to their structural content.

In Section 5, we give an explicit comparison of our framework with that of EHM, summarizing their similarities and differences and commenting on their relative merits. Section 6 analyzes

the structural consequences of various rules that may be employed by policymakers to control potentially cointegrated systems. We pay particular attention there to how the policy rules may introduce, modify, or eliminate cointegration within the target system and to the possible cointegrating relations that may hold between policy instruments and target variables, or among the policy instruments. Section 7 illustrates these methods with an application to a simple model of the U.S. macroeconomy, and Section 8 contains a summary and concluding remarks.

In what follows, we often refer to processes "integrated of order d," $I(d)$ processes for short. By this we mean a stochastic process that becomes $I(0)$ when differenced d times, where an $I(0)$ process is one that obeys the functional central limit theorem.

2 The DGP, Structural Causality, Policy Interventions, and Recursivity

In this section we present the key elements of our policy analysis framework. Our starting point is a dynamic structural system with potentially cointegrated variables, as analyzed in WL (2010). This setting allows for a clear definition of the structural links between the elements of the system, and allows us to introduce in section 2.1 the concepts of *intervention* and *structural causality*. These concepts are instrumental to showing how the structural system introduced in section 2.1 can serve for economic policy analysis. We next define in section 2.2 the concept of *policy intervention*, and introduce the causal property of *partial equilibrium recursivity*. This property defines a precise link between the structural reduced form parameters and the underlying "deep" parameters, and allows the structural system of section 2.1 to be an effective framework for policy analysis. As we will show, this approach permits study of the effectiveness of policy interventions for a broad class of cointegrated structural VAR systems.

2.1 The DGP and Structural Causality

We begin by specifying the data generating process (DGP). For concreteness, clarity, and to afford maximum comparability to EHM, we mainly work with a linear $N-$variate structural vector autoregression (VAR) with two lags:

$$X_t \equiv \left[\begin{array}{c} Y_t \\ Z_t \end{array} \right] = \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t, \qquad t = 1, 2, ..., \tag{1}$$

where Y_t represents observable "target" or "non-policy" variables[1] and Z_t represents observable "policy instruments" or "control variables" that may be useful for controlling Y_t. Both Y_t and Z_t are vectors, $N_1 \times 1$ and $N_2 \times 1$ respectively. Thus, $N = N_1 + N_2$. As an practical example, we might be interested in analyzing the effectiveness of the Fed's monetary policy, through open

---

[1] We follow EHM in referring to Y_t as "target" variables. This should not be confused with similar nomenclature appearing elsewhere in the literature, where "target *series*" means a sequence of desired values Y_t^* for Y_t or "*policy target*" means a desired value for $E(Y_t)$ or some aspect of Y_t or its distribution. When, for convenience, we refer simply to "targets" we always mean "target variables."

market operations, on some key macroeconomic variables. In that case, the policy instrument Z_t would be the effective Federal Funds rate, while the target variables Y_t could represent a variety of key macroeconomic indicators. We will study this particular example in section 7.

The vector $\delta_0 \equiv (\delta'_{10}, \delta'_{20})'$ includes intercepts and any deterministic trend components. (See EHM, eq.(4).) We partition the nonrandom coefficient matrices A_1 and A_2 as

$$A_1 = \left[\begin{array}{c} A_{11} \\ A_{21} \end{array} \right] = \left[\begin{array}{cc} A_{111} & A_{112} \\ A_{211} & A_{212} \end{array} \right] \quad \text{and} \quad A_2 = \left[\begin{array}{c} A_{12} \\ A_{22} \end{array} \right] = \left[\begin{array}{cc} A_{121} & A_{122} \\ A_{221} & A_{222} \end{array} \right].$$

As econometricians, we do not know the A's, nor do we observe the random "shocks" $\varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})'$. Although δ_0, A_1, and A_2 may depend deterministically on t, we leave this implicit to avoid further complicating the notation. We allow δ_0, A_1, and A_2 to generate unit root or other nonstationary processes, with or without cointegration. It is convenient to think of $\{X_t\}$ being (at most) $I(1)$ as EHM do, but this is not essential.

By specifying that this is a structural system, we mean that it causally relates variables on the right to variables on the left. For example, consider an *intervention to* X_{t-1}, denoted $x_{t-1} \to x^*_{t-1}$ and defined as the pair (x_{t-1}, x^*_{t-1}). Then the *direct effect on* the key macroeconomic aggregates Y_t of the intervention $x_{t-1} \to x^*_{t-1}$ at (x_{t-1}, x_{t-2}, e_t) is defined as the difference

$$\begin{aligned} y^*_t - y_t &= (\delta_{10} + A_{11}x^*_{t-1} + A_{12}x_{t-2} + e_t) - (\delta_{10} + A_{11}x_{t-1} + A_{12}x_{t-2} + e_t) \\ &= A_{11}(x^*_{t-1} - x_{t-1}). \end{aligned}$$

We see that A_{11} fully determines the direct effects on Y_t of interventions to X_{t-1}. Indeed, its elements represent the direct effects of a one unit intervention to any given element of X_{t-1}, say $x_{jt-1} \to x_{jt-1} + 1$. Similarly, A_{12} fully determines the direct effects of interventions to X_{t-2}. We may therefore call A_{11} and A_{12} "matrices of effects." These concepts accord well with intuition, and they are especially straightforward because of the linear structure. Similar notions hold generally. See White and Chalak (2009) and WL for discussion of *settable systems*, which provide causal foundations, relied on here, for the general case.

Although A_1 and A_2 are reduced form coefficients, they nevertheless contain structural information and are directly relevant for policy analysis. Specifically, using the notion of causality just given, we can say that if $A_{112} = 0$, then Z_{t-1} *does not (directly) structurally cause* Y_t. Otherwise, Z_{t-1} *structurally causes* Y_t. If $A_{112} = 0$ and $A_{122} = 0$ then $Z^{t-1}_{t-2} \equiv (Z_{t-2}, Z_{t-1})$ does not structurally cause Y_t. Without structural causality from policy variables to target variables (i.e., without $A_{112} \neq 0$ or $A_{122} \neq 0$), policy cannot be effective. EHM (p.375) make a parallel observation, but stated in terms of Granger causality. We provide further discussion below, when we relate structural causality to Granger causality, using the framework of WL. The next section provides further motivation for examining reduced form parameters, "deep" parameters, and their relationship in studying policy analysis.

2.2 Policy Interventions and Recursivity

For economic policy analysis, we need the concept of a *policy intervention*. The rough idea, consistent with EHM, is that this is a change in the structure determining Z_t. In the Fed's monetary example we introduced earlier, this could be a change to the Fed's reaction function, impacting the way in which the Fed sets the target level of the Federal fund rate. To be sufficiently clear about how this works here, we posit an underlying "partial equilibrium" structure, compatible with the system (1). Although this leads us through some seemingly familiar territory, there are some perhaps subtle, but nevertheless important twists along the way.

We write the partial equilibrium structure as

$$\begin{aligned}
\tilde{Y}_t &= b_1 + B_{10}Z_t + B_{11}X_{t-1} + B_{12}X_{t-2} + \upsilon_{1t} \\
\tilde{Z}_t &= b_2 + B_{20}Y_t + B_{21}X_{t-1} + B_{22}X_{t-2} + \upsilon_{2t}, \qquad t = 1, 2, \dots .
\end{aligned} \tag{2}$$

This resembles a familiar system of simultaneous equations, but, in line with conventions of settable systems founded on the prescriptions of Strotz and Wold (1960), the right-hand side (RHS) and left-hand side (LHS) variables are *distinct*, as "responses" $\tilde{Y}_t$ and $\tilde{Z}_t$ appear on the left, whereas "settings" Y_t and Z_t (and their lags) appear on the right.

This seemingly minor notational difference reflects an important feature of such structures: they are *not* simultaneous, and thus avoid paradoxes associated with instantaneous causality and feedback. Nevertheless, just as in classical simultaneous equations, each equation represents the partial equilibrium and/or optimal joint response of the LHS variables to any admissible configuration of the RHS variables. Thus, $\tilde{Y}_t$ represents the (joint) outcome from whatever subsystem determines $\tilde{Y}_t$, when faced with variables outside that subsystem set to admissible values Z_t, X_{t-1}, X_{t-2}, and υ_{1t}. The meaning of $\tilde{Z}_t$ is similar. These responses are determined in isolation, without permitting full equilibrium[2], hence our designation "partial equilibrium."

Let $B_1 \equiv [b_1, B_{10}, B_{11}, B_{12}]$ and $B_2 \equiv [b_2, B_{20}, B_{21}, B_{22}]$. We now define a *structural change*, denoted $B_j \to B_j^*$ ($j = 1$ or 2) as a pair (B_j, B_j^*) of structural coefficients representing "old" (B_j) and "new" (B_j^*) regimes. We also call structural changes "structural shifts." A *policy intervention*, $B_2 \to B_2^*$, is a structural change in the policy equation, i.e., that determining $\tilde{Z}_t$. Our nomenclature is broadly consistent with that of Hendry and Massman (2006).

To specify the system's response when all LHS variables are determined jointly, rather than in isolation, we must specify how this joint determination is achieved. For this, we apply the equilibrium requirement of mutual consistency. In equilibrium, the structure (2) satisfies

$$\begin{aligned}
\tilde{Y}_t &= b_1 + B_{10}\tilde{Z}_t + B_{11}X_{t-1} + B_{12}X_{t-2} + \upsilon_{1t} \\
\tilde{Z}_t &= b_2 + B_{20}\tilde{Y}_t + B_{21}X_{t-1} + B_{22}X_{t-2} + \upsilon_{2t}, \qquad t = 1, 2, \dots .
\end{aligned} \tag{3}$$

---

[2]In settable systems language, partial equilibrium corresponds to the "agent partition," and full equilibrium corresponds to the "global partition." The partitions specify mutually exclusive subsystems, each of whose variables respond freely and jointly to variables outside that subsystem. See White and Chalak (2009) for details.

Although this resembles a classical system of structural equations, we explicitly do not view this as structural, because in the settable systems framework adopted here, structural relations necessarily embody causality. Interpreting eq.(3) causally requires instantaneous feedback (causality), and, consistent with Granger and Newbold's (1986, p.221) position on instantaneous feedback, settable systems do not allow this. Thus, eqs.(3) are not structural equations; instead they only represent the mutual consistency conditions necessary for equilibrium. Eq.(2) is the governing structural equation system.

If instantaneous feedback is ruled out, one must explain how mutual consistency can nevertheless be achieved. A standard approach is that taken in game theory, where each player has sufficient information to compute the equilibrium. Let "player" 1 (the public) determine $\tilde{Y}_t$ and "player" 2 (the policy authority) determine $\tilde{Z}_t$. Using (3), the full equilibrium, X_t, is given by the reduced form structural VAR, eq.(1):

$$X_t = \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t, \qquad t = 1, 2, ..., \qquad \text{where}$$

$$\delta_0 = \Delta \left[\begin{array}{c} b_1 \\ b_2 \end{array} \right] \qquad \Delta \equiv \left[\begin{array}{cc} I & -B_{10} \\ -B_{20} & I \end{array} \right]^{-1}$$

$$A_1 = \Delta \left[\begin{array}{c} B_{11} \\ B_{21} \end{array} \right] \qquad A_2 = \Delta \left[\begin{array}{c} B_{12} \\ B_{22} \end{array} \right], \qquad \text{and}$$

$$\varepsilon_t = \Delta \left[\begin{array}{c} v_{1t} \\ v_{2t} \end{array} \right]. \tag{4}$$

In this framework, it suffices for each player to know B_1, B_2, and $v_t \equiv (v'_{1t}, v'_{2t})'$.

We now have sufficient foundation to embark on policy analysis, that is, the study of the consequences of policy interventions. We focus here on policy interventions that introduce changes to the policymaker's subsystem of the DGP; these in turn imply changes to the full equilibrium structural VAR for Z_t. A crucial requirement for traditional policy analysis is that the full equilibrium structural VAR for Y_t is invariant to the policy intervention. That is, meaningful policy analysis requires the public subsystem of the DGP to be invariant to the policy intervention. If this were not the case, after the intervention the public would likely reoptimize their behavior to accommodate the new regime introduced by the intervention, and it would be no longer possible to disentangle the true policy effect of the intervention from the change in the public's behavior. As Lucas (1976) pointed out, without this required invariance, policy implications from such interventions would likely be misleading. In the EHM framework, superexogeneity ensures this invariance, and EHM (section 3) give compelling evidence that this invariance holds empirically. In the monetary policy example given above, this property

requires that the structural VAR describing the dynamics of the key macroeconomic indicators Y_t is invariant to changes in the Fed's reaction function.

Our use of settable systems permits ensuring the required invariance using an approach alternative to superexogeneity. First, observe that because the structural reduced form A's depend on all the underlying "deep parameters" B, a policy intervention $B_2 \to B_2^*$ generally leads to a structural shift $(\delta_{10}, A_{11}, A_{12}) \to (\delta_{10}^*, A_{11}^*, A_{12}^*)$ in the full equilibrium structural VAR for Y_t, violating invariance. The desired invariance is impossible without some further restriction. Here, the restriction analogous to superexogeneity is that $B_{10} = 0$. We call this *partial equilibrium recursivity*, or, more simply, just *recursivity*. With recursivity,

$$\begin{aligned} \tilde{Y}_t &= b_1 + B_{11}X_{t-1} + B_{12}X_{t-2} + \upsilon_{1t} \\ \tilde{Z}_t &= b_2 + B_{20}Y_t + B_{21}X_{t-1} + B_{22}X_{t-2} + \upsilon_{2t} \qquad t = 1, 2, ... \end{aligned} \tag{5}$$

This condition is sufficient for invariance to policy interventions of the reduced form structural VAR for Y_t, as $B_{10} = 0$ implies $\delta_{10} = b_1$, $A_{11} = B_{11}$ and $A_{12} = B_{12}$. Recursivity is also necessary for invariance to policy interventions of the reduced form VAR for Y_t, in the following sense:

Proposition 2.1 *Suppose B_{10} and B_{20} are such that Δ exists and that equations (4) hold. Then $B_{10} = 0$ if and only if for all $[b_1, B_{11}, B_{12}]$, we have $[\delta_{10}, A_{11}, A_{12}] = [b_1, B_{11}, B_{12}]$.*

Recursivity is informationally plausible, as it allows the public not to know the policymaker's response function and shock. Instead, the public need only know its own optimal response coefficients, B_1. Though υ_{1t} may include components known only to the public, it may also contain an "implementation error" or "tremble" that the public does not control or know.

Recursivity is also behaviorally plausible. Indeed, experimental evidence in economics supports the hypothesis that interacting agents do not arrive at fully rational Nash equilibria, but there is instead a "level$-k$" hierarchy of agents, who adopt strategies of varying sophistication (Stahl and Wilson, 1994; Crawford and Iriberri, 2007). Recursivity is consistent with viewing the public as a level k player and the policy authority as a level $k + 1$ player. This ordering is plausible, as the public is not a single monolithic rational agent, but an aggregate of agents of varying objectives and sophistication. On the other hand, the policymaker is typically a well-defined government entity with more or less coherent objectives and with resources sufficient to know or learn player 1's coefficients B_1, which it may use to determine its coefficients B_2.

In what follows, then, we take partial equilibrium recursivity to be a maintained assumption, analogous to superexogeneity in the EHM framework.

3 Conditional Exogeneity and Granger Causality

Our discussion so far references only the DGP. Models, as distinct from the DGP, have played no role. We view this as an advantage, as delaying the introduction of models until absolutely

necessary not only accords with Occam's principle, but also yields a theory with broader potential applicability. We now discuss two further properties of the DGP, conditional exogeneity and Granger causality, that bear directly on policy analysis.

3.1 Conditional Exogeneity

In the settable systems approach, exogeneity plays a crucial role in identifying causal effects. Here, identification means the notion of "correspondence to the desired entity" as discussed by Hendry (1995) and Hendry, Lu, and Mizon (2009), based on notions of Wright (1915). The particular correspondence relevant here is that between aspects (e.g., functions of moments) of the joint distribution of observable variables, e.g., $\{(Y_t, Z_t)\}$, and the structural information embodied in δ_0, A_1, and A_2. WL give results implying that structural coefficients (δ_0, A_1, A_2) can be identified when data are generated as in (1), provided that (X_{t-1}, X_{t-2}) is independent of ε_t given covariates[3] W_t, or, in Dawid's (1979) notation,

$$(X_{t-1}, X_{t-2}) \perp \varepsilon_t \mid W_t. \tag{6}$$

This is a time-series analog of the selection on observables condition (Barnow, et al. 1980).

When (6) holds, we say that (X_{t-1}, X_{t-2}) is *conditionally exogenous with respect to ε_t given W_t*, or just *conditionally exogenous*. This is a conditional form of the strict exogeneity relation,

$$(X_{t-1}, X_{t-2}) \perp \varepsilon_t, \tag{7}$$

where W_t is absent. For example, (7) holds for (1) when $\{\varepsilon_t\}$ is independent and identically distributed (IID), as in EHM, and $\{\varepsilon_t\}$ is independent of (X_0, X_{-1}), a standard assumption.

When strict exogeneity fails, conditional exogeneity can nevertheless hold, as WL discuss in detail; see also White (2006a). Suitable choices for W_t are proxies for ε_t, including not only current and lagged values of variables that may also be driven by ε_t but also their *leads* (see White and Kennedy, 2009). W_t should not be driven by lagged X_t's.

Observe that conditional exogeneity is distinct from weak, strong, or superexogeneity (Engle, Hendry, and Richard, 1983), as these concepts are defined strictly with respect to a model. In contrast, conditional exogeneity is a property solely of the DGP.

To see how conditional exogeneity ensures identification of structural coefficients, we write

$$
\begin{aligned}
E(X_t \mid X_{t-1}, X_{t-2}, W_t) &= E(\delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t \mid X_{t-1}, X_{t-2}, W_t) \\
&= \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + E(\varepsilon_t \mid X_{t-1}, X_{t-2}, W_t) \\
&= \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + E(\varepsilon_t \mid W_t) \\
&= \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + C_0 W_t.
\end{aligned} \tag{8}
$$

---

[3] Covariates are sometimes called "control variables," as they "control for" the influence of otherwise omitted variables. Here, we avoid confusion by reserving the designation "control variables" for those variables Z_t that control the target variables Y_t.

The third equality uses (6), as this implies $E(\varepsilon_t \mid X_{t-1}, X_{t-2}, W_t) = E(\varepsilon_t \mid W_t)$. The final equality invokes a simplifying linearity assumption, $E(\varepsilon_t \mid W_t) = c_0 + C_0 W_t$, with $c_0 = 0$. Linearity is by no means essential, but it keeps our notation and discussion simple. When c_0 differs from zero, then the structural intercept (i.e., the non-trend component of δ_0) becomes unidentified; this need not be a serious difficulty, however.

Thus, regressing X_t on X_{t-1}, X_{t-2}, and W_t will yield consistent estimates of δ_0, A_1, A_2, and C_0, under suitable conditions. These conditions can even permit structural shifts. As the details are somewhat involved, we leave this aside for now. The regression model implicitly referenced here must be correctly specified for the *sequence* of conditional expectations $\{E(X_t \mid X_{t-1}, X_{t-2}, W_t)\}$, in keeping with the discussion of White (1994, pp.141-147, especially p.144).

Note that models have just appeared for the first time and that weak exogeneity plays no role. The only model condition we explicitly require is correct specification for the conditional mean sequence $\{E(X_t \mid X_{t-1}, X_{t-2}, W_t)\}$. This condition does not apply directly to the structural system (1). Nevertheless, knowledge of important features of the DGP (1) plays a key role in achieving correct specification. This knowledge includes (i) which variables are economically meaningful choices for Y_t and Z_t; and (ii) which variables W_t, driven by unobservable drivers of Y_t and Z_t, may plausibly suffice for (6), conditional exogeneity. Specification issues of functional form, numbers of lags, cointegration (discussed later), and even structural shift locations, among others, may be resolved from the data.

It is especially noteworthy that some, but not all, of the regression coefficients in (8) have structural meaning. Specifically, δ_0, A_1, and A_2 are structural coefficients directly relevant for policy analysis, whereas C_0 has no structural meaning. Instead, C_0 yields optimal predictions.

Policy analysis may be conducted without full knowledge of δ_0, A_1, and A_2. For example, interest may attach just to A_{112} and A_{122} in

$$Y_t = \delta_{10} + A_{111} Y_{t-1} + A_{112} Z_{t-1} + A_{121} Y_{t-2} + A_{122} Z_{t-2} + \varepsilon_{1t}, \qquad t = 1, 2, ...,$$

as A_{112} and A_{122} determine whether policy is effective or not. The milder exogeneity condition

$$(Z_{t-1}, Z_{t-2}) \perp \varepsilon_{1t} \mid (Y_{t-1}, Y_{t-2}, W_t). \tag{9}$$

identifies A_{112} and A_{122}. Now we only require a correctly specified model for the sequence $\{E(Y_t \mid X_{t-1}, X_{t-2}, W_t)\}$; again, weak exogeneity is not required.

3.2   Granger Causality

WL give results implying that given (1) and conditional exogeneity, Granger causality is equivalent to structural causality. WL also give results implying that given (1) and in the *absence* of structural causality, Granger causality is equivalent to the failure of conditional exogeneity. We now make these claims precise and discuss their implications for policy analysis.

The relevant equivalence of structural and Granger causality is as follows:

Proposition 3.1 *Suppose that $\{(W_t, X_t, \varepsilon_t)\}$ is a stochastic process satisfying (1) and (9), and that (Z_{t-1}, Z_{t-2}) is not solely a function of (Y_{t-1}, Y_{t-2}, W_t). Then (Z_{t-1}, Z_{t-2}) does not structurally cause Y_t (i.e., $A_{112} = 0$ and $A_{122} = 0$) if and only if*

$$Y_t \perp (Z_{t-1}, Z_{t-2}) \mid (Y_{t-1}, Y_{t-2}, W_t),$$

that is, (Z_{t-1}, Z_{t-2}) **does not finite-order** $G-$**cause** Y_t **with respect to** (Y_{t-1}, Y_{t-2}, W_t).

The finite-order Granger non-causality condition $Y_t \perp (Z_{t-1}, Z_{t-2}) \mid (Y_{t-1}, Y_{t-2}, W_t)$ is not classical G non-causality. In the notation here, the classical condition is

$$Y_t \perp Z^{t-1} \mid Y^{t-1}, W^{t-1},$$

where $Z^{t-1} \equiv (Z_{t-1}, Z_{t-2}, ...)$ is the "$t-1$ history" of $\{Z_t\}$ and W_t contains no leads. As WL explain in detail, finite-order G non-causality is the extension of the classical condition most directly relevant for "Markov" structures such as (1), in the sense that this is the condition equivalent to structural non-causality, given conditional exogeneity. The classical condition corresponds to more general structures under different but related exogeneity conditions. As WL also explain, the covariates W_t can contain both lags and leads relative to time t, without violating the causal direction of time. Thus, the presence of W_t in the finite-order definition does not conflict with the spirit (or causal content) of the classical definition.

Without further conditions, neither $G-$causality property is necessary nor sufficient for the other. As WL note, the finite-order condition is that usually tested in the literature.

Thus, in the presence of the conditional exogeneity required to identify specific causal effects, statements about $G-$causality (specifically, the applicable finite-order $G-$causality) are essentially statements about structural causality. Given conditional exogeneity, it may therefore be possible to determine whether the policy and target variables have genuine causal links (as required by Granger and Deutsch, 1992; see also EHM, p.375), by testing whether (Z_{t-1}, Z_{t-2}) (or some other suitable finite history) finite-order $G-$causes Y_t (with respect to (Y_{t-1}, Y_{t-2}, W_t)). We qualify this statement by saying that this determination "may" be possible to signal that certain control scenarios can interfere with use of $G-$causality for this purpose (see Sargent, 1976; Buiter, 1984; Granger, 1988; and Ermini, 1992), namely that (Y_{t-1}, Y_{t-2}, W_t) completely determines (Z_{t-1}, Z_{t-2}). We provide further discussion in Section 7, where we discuss estimating and testing controlled systems. For the time being, we treat such cases as special.

Similarly, one can test whether the policy variables (Y_{t-1}, Y_{t-2}) (or some other suitable finite history) structurally cause Z_t (i.e., $A_{211} \neq 0$ or $A_{221} \neq 0$) by testing whether (Y_{t-1}, Y_{t-2}) finite-order $G-$causes Z_t (with respect to (Z_{t-1}, Z_{t-2}, W_t)). As EHM (p.375) comment, "actual policy simulations may or may not assume such feedback," although past values of target variables

typically do influence policy-making behavior. Nevertheless, as we see in Section 6, feedback is not necessary for policy effectiveness.

By itself, however, Granger causality is *not* enough to ensure the presence of the genuine causal links required for policy effectiveness. The reason is that when structural causality is absent, Granger causality can still appear, as a consequence of the failure of conditional exogeneity. In fact, the two properties are equivalent in this case:

Proposition 3.2 *Suppose that $\{(W_t, X_t, \varepsilon_t)\}$ is a stochastic process satisfying (1) and that (Z_{t-1}, Z_{t-2}) is not solely a function of (Y_{t-1}, Y_{t-2}, W_t).*

(i) Suppose (Z_{t-1}, Z_{t-2}) does not structurally cause Y_t (i.e., $A_{112} = 0$ and $A_{122} = 0$). Then (Z_{t-1}, Z_{t-2}) does not finite-order $G-$cause Y_t with respect to (Y_{t-1}, Y_{t-2}, W_t) if and only if

$$(Z_{t-1}, Z_{t-2}) \perp \varepsilon_{1t} \mid (Y_{t-1}, Y_{t-2}, W_t).$$

(ii) Suppose (X_{t-1}, X_{t-2}) does not structurally cause Y_t (i.e., $A_{11} = 0$ and $A_{12} = 0$). Then (X_{t-1}, X_{t-2}) does not finite-order $G-$cause Y_t with respect to W_t if and only if

$$(X_{t-1}, X_{t-2}) \perp \varepsilon_{1t} \mid W_t.$$

Thus, when EHM (p.375) state, "Without Granger causality from instruments to targets, policy is unlikely to be effective," one must , recognize that, in the present context, the accuracy of this statement rests on the strict exogeneity $(X_{t-1}, X_{t-2}) \perp \varepsilon_{1t}$ ensured by their specification of the DGP (that $\{\varepsilon_t\}$ is IID in (1); see EHM, p.373). Otherwise, the presence of Granger causality has nothing necessarily to say about policy effectiveness, because it has nothing necessarily to say about the structural causality required for policy effectiveness. Instead, $G-$causality may simply be signalling exogeneity failure.

4   System Estimation With and Without Cointegration

Consider a generic structural VAR (i.e., we permit but do not require $X_t = (Y_t', Z_t')'$):

$$X_t = \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t, \quad t = 1, 2, ...,$$

and suppose that ΔX_t is $I(0)$ and that there exist $r < N$ cointegrating relations such that $\beta' X_t$ is also $I(0)$, where β is an $N \times r$ matrix with full column rank. We emphasize that here the cointegrating relations are dynamic properties of the data generating process. They are explicitly not causal, as also emphasized by EHM, p.378. When $\beta' X_t$ is $I(0)$, there also exists an $N \times r$ matrix α with full column rank (Johansen, 1988) such that

$$\alpha \beta' = A_1 + A_2 - I.$$

11

Letting $\Gamma \equiv -A_2$ then gives the standard error-correction cointegrating representation

$$\Delta X_t = \delta_0 + \alpha\,\Psi_{t-1} + \Gamma\Delta X_{t-1} + \varepsilon_t, \tag{10}$$

where $\Delta X_t \equiv X_t - X_{t-1}$ and $\Psi_{t-1} \equiv \beta' X_{t-1}$. Because eq. (10) is a rewriting of a structural VAR that preserves the directionality of the structural relations, it retains its structural meaning. In this form, it represents the causal relations holding between the response ΔX_t and any admissible settings of RHS variables Ψ_{t-1}, ΔX_{t-1}, and ε_t. Thus, the matrices α and Γ embody the effects on ΔX_t of interventions to Ψ_{t-1} (long-run equilibrium departures) and ΔX_{t-1}, respectively. On the other hand, β does not embody causal effects among elements of X_t, as the cointegrating relations are not causal.

When cointegration is present, the relevant exogeneity conditions permit estimation along standard lines. Specifically, suppose (6) holds. It follows from Dawid (1979, lemma 4.2(i)) that

$$(X_{t-1}, \Delta X_{t-1}) \perp \varepsilon_t \mid W_t \quad \text{and} \quad (\Psi_{t-1}, \Delta X_{t-1}) \perp \varepsilon_t \mid W_t.$$

Thus, for example, we have

$$
\begin{aligned}
E(\Delta X_t \mid \Psi_{t-1}, \Delta X_{t-1}, W_t) &= \delta_0 + \alpha\Psi_{t-1} + \Gamma\Delta X_{t-1} + E(\varepsilon_t \mid \Psi_{t-1}, \Delta X_{t-1}, W_t) \\
&= \delta_0 + \alpha\Psi_{t-1} + \Gamma\Delta X_{t-1} + E(\varepsilon_t \mid W_t) \\
&= \delta_0 + \alpha\Psi_{t-1} + \Gamma\Delta X_{t-1} + C_0 W_t.
\end{aligned}
$$

As for the Engle-Granger estimator (Engle and Granger, 1987), one can apply a two-stage procedure, estimating β in a first stage by least squares (Stock, 1987), forming an estimate $\hat{\Psi}_{t-1}$ of Ψ_{t-1}, and then regressing ΔX_t on an intercept and $\hat{\Psi}_{t-1}$, ΔX_{t-1}, W_t to obtain standard estimators of δ_0, α, Γ, and C_0. An interesting feature of this regression is that conditional exogeneity justifies the inclusion of covariates W_t, as above, which may include both lags and leads with respect to time t. To the best of our knowledge, this possibility has not previously been noted. As above, C_0 has no structural meaning, whereas the remaining coefficients have the desired structural interpretation.

Similarly, one can apply methods of Johansen (1988, 1995), but also including covariates W_t as regressors along with X_{t-1} and ΔX_{t-1}.

When cointegration does not hold, quasi-maximum likelihood methods nevertheless apply to deliver useful estimators of coefficients of interest. We saw above that

$$E(X_t \mid X_{t-1}, X_{t-2}, W_t) = \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + C_0 W_t.$$

Further, observe that with $\eta_t \equiv X_t - E(X_t \mid X_{t-1}, X_{t-2}, W_t) = \varepsilon_t - E(\varepsilon_t \mid W_t)$, the exogeneity condition $(X_{t-1}, X_{t-2}) \perp \varepsilon_t \mid W_t$ implies $(X_{t-1}, X_{t-2}) \perp \eta_t \mid W_t$, so that

$$E(\eta_t \eta_t' \mid X_{t-1}, X_{t-2}, W_t) = E(\eta_t \eta_t' \mid W_t).$$

It is plausible that this conditional heteroskedasticity can be exploited to yield a relatively efficient GLS-like estimator, based on a suitable specification for $E(\eta_t \eta_t' \mid W_t)$. Observe that conditional exogeneity simplifies the modeling, as X_{t-1} and X_{t-2} do not contribute to the conditional variance. On the other hand, since W_t is explicitly chosen to predict ε_t, we should generally expect it to predict $\eta_t \eta_t'$ as well, affording the opportunity for possible efficiency gains.

The normal quasi-maximum likelihood estimator (QMLE) is the solution to the problem

$$
\begin{aligned}
\max_{\theta \in \Theta} \ L_T(\theta) \quad \equiv \quad & T^{-1} \sum_{t=1}^{T} -.5 \ln \det(\Sigma(W_t; \theta_2)) \\
& -.5(X_t - \mu(X_{t-1}, X_{t-2}, W_t; \theta_1))'[\Sigma(W_t; \theta_2)]^{-1}(X_t - \mu(X_{t-1}, X_{t-2}, W_t; \theta_1)),
\end{aligned}
$$

where $\theta \equiv (\theta_1', \theta_2')$, $\Sigma(W_t; \theta_2)$ is a parametrization for $E(\eta_t \eta_t' \mid W_t)$, and $\mu(X_{t-1}, X_{t-2}, W_t; \theta_1)$ is a parametrization (e.g., linear) of $E(X_t \mid X_{t-1}, X_{t-2}, W_t)$.

Although this is the usual normal QMLE, its asymptotic properties will vary, depending on those of $\{X_t\}$, which may contain trends, unit roots, and possible unsuspected cointegration. Generally, the QMLE will be consistent, but its asymptotic distribution need not be normal. Asymptotic theory sufficiently general to handle this QMLE for the strictly exogenous case (W_t absent) can be found in Park and Phillips (1988, 1989), Ahn and Reinsel (1990), Li, Ling, and Wong (2001), and Sin (2004). Developing theory for the fully general conditionally exogenous case in the absence of cointegration is an interesting topic for future research.

5   A Comparison with EHM

The settable systems-based policy analysis framework laid out above contains a variety of elements in common with the framework set forth by EHM. Nevertheless, the relation and roles of these elements differ between the two approaches. There are also elements in each that are not shared by the other. In this section, we briefly summarize the similarities and differences of these systems and comment on their relative merits.

The goal of both our approach and EHM's is to specify conditions under which one can analyze the effects of policy interventions through the use of an econometric model. Both approaches start by specifying the DGP. For clarity and concreteness, both we and EHM work with a linear $N-$variate structural VAR with two lags, eq.(1). For expositional convenience, EHM restrict $\{X_t\}$ to be (at most) $I(1)$. We emphasize that this is just for convenience; in the next section we see how an $I(2)$ process for Z_t can arise naturally.

As EHM note, a necessary condition for policy analysis is that the policy instruments and targets have genuine causal links (EHM, p.375, condition 1). In our framework, this requirement is literally enforced by a structural causality condition: the causal effects B_{112} and B_{122} of policy instruments Z_{t-1} and Z_{t-2} on the partial equilibrium response $\tilde{Y}_t$ must not both be zero.

Otherwise, the policy instruments have no causal effect on the target variable. In contrast, EHM (p.375) link this requirement to Granger causality: "Without Granger causality from instruments to targets, policy is unlikely to be effective." The qualification "unlikely" properly reflects the lack of perfect correspondence between structural causality and Granger causality; the two are not the same. The present framework draws the needed distinction, based on work of WL, who show that structural and $G-$causality are equivalent, given a suitable conditional form of exogeneity. In this sense, $G-$causality is a derivative requirement that may be useful for testing the structural causality of policy instruments, the fundamental requirement here.

Another necessary condition is that the policy intervention "does not alter the econometric model in a self-contradictory way," ensuring that the Lucas (1976) critique does not hold (EHM, p.375, condition 3). EHM enforce this requirement by imposing superexogeneity to ensure the necessary invariance (Engle, Hendry, and Richard, 1983, definition 2.9). Superexogeneity combines the properties of weak exogeneity and invariance to a specified set of parameter interventions. Thus, superexogeneity is undefined without weak exogeneity. Weak exogeneity, however, is a property of a correctly specified model relative to a DGP that acts primarily to ensure estimator efficiency (see White, 1994, pp. 141-147). A significant concern is that imposing weak exogeneity can rule out important structures directly relevant for policy analysis. For example, Fisher (1993) shows that weak exogeneity is violated when dynamic stability is imposed in cointegrated structural VAR models.

As a simplified version of Fisher's (1993) example, let both Y_t and Z_t be scalars, and consider the integrated structural VAR

$$\begin{bmatrix} \Delta Y_t \\ \Delta Z_t \end{bmatrix} = \begin{bmatrix} \delta_{10} \\ \delta_{20} \end{bmatrix} + \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} \Delta Y_{t-1} \\ \Delta Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{11}$$

where $[\varepsilon_{1t}, \varepsilon_{2t}]' \sim N(0, \Sigma)$. Central to the EHM approach is the reparametrization of (11) in terms of conditional $(\Delta Y_t \mid \Delta Z_t)$ and marginal (ΔZ_t) distributions. Here this yields

$$\Delta Y_t = \theta_0 + \theta_1 \Delta Z_t + \theta_2 \Delta Y_{t-1} + \theta_3 \Delta Z_{t-1} + \zeta_t \tag{12}$$

$$\Delta Z_t = \delta_{20} + \Gamma_{21} \Delta Y_{t-1} + \Gamma_{22} \Delta Z_{t-1} + \varepsilon_{2t}, \tag{13}$$

with $\zeta_t \equiv \varepsilon_{1t} - \theta_1 \varepsilon_{2t} \sim N(0, \Omega)$, and where $\Omega = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.[4] Also,

$$\theta_1 = \Sigma_{12}\Sigma_{22}^{-1} \qquad \theta_0 = \delta_{10} - \theta_1 \delta_{20}$$

$$\theta_2 = \Gamma_{11} - \theta_1 \Gamma_{21} \qquad \theta_3 = \Gamma_{12} - \theta_1 \Gamma_{22}. \tag{14}$$

Without dynamic stability restrictions, the parameters of the conditional and marginal distributions, $\lambda_1 = (\theta_0, \theta_1, \theta_2, \theta_3, \Omega)$ and $\lambda_2 = (\delta_{20}, \Gamma_{21}, \Gamma_{22}, \Sigma_{22})$, define a sequential cut between the

---

[4] Here Σ_{ij} identifies the generic $i, j-$th term of the matrix Σ.

conditional model (12) and the marginal model (13). The cross restrictions stemming from (14) impose no specific restrictions on the elements of either λ_1 or λ_2. Hence, weak exogeneity holds.

However, if short-run dynamic stability is imposed in (12), then we require $|\theta_2| < 1$. This, together with (14), implies $|\Gamma_{11} - \theta_1 \Gamma_{21}| < 1$. Now λ_1 and λ_2 are no longer variation free, and weak exogeneity of Y_{t-1} no longer holds. EHM policy analysis is not possible in this system.

In contrast, our approach enforces the needed invariance by imposing the partial equilibrium recursivity restriction $B_{10} = 0$, i.e., Z_t does not structurally cause $\tilde{Y}_t$ in partial equilibrium. This ensures that the coefficients of the full equilibrium reduced form *data generating process* for Y_t are invariant to policy interventions. This is not a property of the model, so we are not imposing invariance on the econometric *model*, as in condition 3 of EHM. But our requirement does imply that an invariant model for Y_t can be correctly specified, since recursivity ensures that invariance holds for the DGP. We view the transparency and plausibility of the partial equilibrium recursivity condition $B_{10} = 0$ as a further advantage.

As we do not require weak exogeneity, our approach applies to the structurally stable VAR above or to Fisher's (1993) more elaborate example with cointegration. Although efficiency may not be achieved without weak exogeneity, in our view this sacrifice is worth the gain of permitting the analysis of important policy-relevant DGPs. Even without full efficiency, relative efficiency gains are often possible. And if weak exogeneity does hold, then nothing is lost.

Although we do not require estimator efficiency, adept policy analysis minimally requires consistent coefficient estimation. This is required for policymakers seeking to implement effective policy, as these policies typically depend on the coefficients of the structural VAR for the target. This is also required for econometricians seeking to understand how the components of the controlled system behave, both individually and jointly. To ensure consistent estimation, the coefficients of interest must be identified in the sense previously described. Otherwise, no model can inform us about these. Given identification, we then require a correctly specified model for certain aspects of the distribution of observables (e.g., specific conditional expectations). To ensure identification of the various effects of interest in our framework, we rely on conditional exogeneity requirements. Thus, whereas EHM rely on weak exogeneity in a correctly specified model to arrive at efficient estimates of weakly exogenous parameters, we rely on conditional exogeneity to identify structural effects (coefficients) of interest and a correctly specified model to consistently estimate these. In this way, our approach satisfies EHM's condition 2 for a policy analytic framework to be of value, namely that "the model represents the economy closely enough that its policy predictions reasonably match outcomes."

EHM additionally require that "the policy experiment is feasible" (condition 4) and that "the policy instruments are manipulable" (condition 5). These conditions are also in force here, but with the difference that because policy interventions (experiments) here are structural shifts to the policy subsystem of the DGP, the model is not directly involved, as it is for EHM.

15

(Note that with settable systems, a sufficiently flexible DGP can readily accommodate policy interventions; one need not posit a separate DGP for each policy intervention.) Feasibility here means that the contemplated intervention to the policy subsystem is itself compatible with the DGP. Manipulability means that the policy instruments can in fact be set by the policy authority to the value specified by the policy rule. In the discussion of Section 6, where we study implications of various policy rules, we take feasibility and manipulability for granted.

Thus, EHM work with the DGP-based properties of Granger causality and cointegration, together with the model properties of weak exogeneity and superexogeneity to pursue policy analysis. Here, we pursue policy analysis using the DGP-based properties of structural causality, conditional exogeneity, and cointegration, together with the model property of correct specification. There are two causal requirements: (i) causality of lagged policy instruments for the partial equilibrium target response and (ii) recursivity, i.e., non-causality of current policy instruments for the partial equilibrium target response. Invariance is a consequence of recursivity; Granger causality is an implication of structural causality and conditional exogeneity.

6   Some Structural Implications of Policy Control Rules

So far, we have taken as given the dynamics of the control subsystem determining Z_t, that is, the policymaker's behavior. But policymakers may follow specific rules to attain their policy objectives; these rules generally have implications for the integration and cointegration properties of the system and its components. We now demonstrate the utility of the present framework by examining the consequences of various policymaker behaviors, all directed toward achieving the goal of a desired long-run expected value for the target variable. Despite its simplicity, this case usefully illustrates a variety of interesting features of the controlled system. The analysis is facilitated by not having to account for a priori considerations of weak exogeneity.

We begin by recalling the recursive partial equilibrium structural system specified earlier,

$$\tilde{Y}_t = b_1 + B_{11}X_{t-1} + B_{12}X_{t-2} + \upsilon_{1t},$$
$$\tilde{Z}_t = b_2 + B_{20}Y_t + B_{21}X_{t-1} + B_{22}X_{t-2} + \upsilon_{2t} \qquad t = 1, 2, ...$$

Next, we translate this system to a form called the *canonical* recursive representation. For this, we equate settings with responses, i.e., $Y_t = \tilde{Y}_t$ and $Z_t = \tilde{Z}_t$, so that

$$Y_t = b_1 + B_{111}Y_{t-1} + B_{112}Z_{t-1} + B_{121}Y_{t-2} + B_{122}Z_{t-2} + \upsilon_{1t},$$
$$Z_t = b_2 + B_{20}Y_t + B_{211}Y_{t-1} + B_{212}Z_{t-1} + B_{221}Y_{t-2} + B_{222}Z_{t-2} + \upsilon_{2t} \qquad t = 1, 2, ... \ (15)$$

where we have also exploited the partitioning of the four random matrices B_{11}, B_{12}, B_{21}, and B_{22}. In particular,$B_{11} = \begin{bmatrix} B_{111} & B_{112} \end{bmatrix}$, $B_{12} = \begin{bmatrix} B_{121} & B_{122} \end{bmatrix}$, $B_{21} = \begin{bmatrix} B_{211} & B_{212} \end{bmatrix}$, and $B_{22} = \begin{bmatrix} B_{221} & B_{222} \end{bmatrix}$. We explicitly rule out instantaneous causation by requiring that υ_{1t} is

16

realized prior to Y_t, and that Y_t and v_{2t} are realized prior to Z_t. These realizations can be viewed as occurring within the period, that is, after $t-1$ and before t. We emphasize this requirement by referring to this as *contemporaneous* rather than instantaneous causation. These equations now represent the natural system evolution in a form making it particularly suitable for describing policymaker behavior and for studying the implications of this behavior.

We start by studying a system of the particular form

$$
\begin{aligned}
\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + \Gamma_{11}\Delta Y_{t-1} + \Gamma_{12}\Delta Z_{t-1} + v_{1t}, \\
\Delta Z_t &= b_2 + \Gamma_{20}\Delta Y_t + \Gamma_{12}^{-1}[\alpha_1\lambda_1' + \kappa_1(\beta_1 + \lambda_1)']Y_t \quad t = 1, 2, ...
\end{aligned}
\tag{16}
$$

where both λ_1 and κ_1 are $N_1 \times r$ matrices. The system (16) is a restricted version of (15), with $\alpha_1\beta_1' = (B_{111} + B_{121} - I_{N_1})$, $\Gamma_{11} = -B_{121}$, $\Gamma_{12} = B_{112}$, $\Gamma_{20} = B_{20} - B_{112}^{-1}[\alpha_1\lambda_1' + \kappa_1(\beta_1 + \lambda_1)']$, and with $B_{122} = -B_{112}$, $B_{211} = -\Gamma_{20}$, $B_{212} = I_{N_2}$, $B_{221} = 0$ $(N_2 \times N_1)$, $B_{222} = 0$ $(N_2 \times N_2)$. For simplicity, we take $v_{2t} = 0$, so the policymaker is able to precisely implement their policy.

We will see shortly that (16) represents a quite general control rule; studying its properties will also help shed light on other types of control rules. For simplicity, we assume here that the shocks $\{v_{1t}\}$ are IID with mean zero. Thus, v_{1t} is independent of (X_{t-1}, X_{t-2}) and estimation is standard (e.g., Ahn and Reinsel, 1990).

With no control ($\Delta Z_t = 0$), we observe the "open-loop" target dynamics,

$$
\Delta Y_t = b_1 + \alpha_1\beta_1' Y_{t-1} + \Gamma_{11}\Delta Y_{t-1} + v_{1t}.
$$

When $\beta_1 = 0$, we have an integrated open-loop system for Y_t without cointegration. Otherwise, with $0 < r \equiv \mathrm{rk}(\beta_1) < N_1$, the open-loop target system exhibits cointegration.

We suppose the policymaker seeks to attain $E(\Delta Y_t) = \gamma_o$ in the long run, as, for example, when the policymaker targets an inflation rate or a GDP growth rate. The question of how to adjust ΔZ_t to achieve a desired long-run policy goal or even a series of desired target values $\{\Delta Y_t^*\}$ in dynamic systems is the subject of the theory of optimal control. There is a vast literature in this area; the classical theory developed in engineering and related fields, was adopted early into economics (Simon, 1952), and has transformed in ways relevant to specific challenges in economics. See Ermini (1992) and Pagan (1997) for a discussion of this evolution.

Despite its potential importance for policy analysis, the study of control of cointegrated systems has only received modest attention so far. Besides Ermini (1992) and EHM, works considering various aspects of this topic are those of Granger (1988), Karunaratne (1996), Johansen and Juselius (2001) (JJ), and Monti (2003). JJ and Monti (2003) in particular give sophisticated treatments of control in cointegrated systems. For conciseness, we do not reiterate the foundations of this theory. Instead we only note the proliferation of a variety of control rules, with different degrees of complexity and sophistication. In particular, the classes of proportional-integral (PI) and proportional-integral-derivative (PID) methods are commonly adopted in the

engineering literature, both of which can also be expressed as constrained versions of eq.(15). Here, we focus on the PI rule and take as our starting point its most general version, which is indeed (16).[5] Hereafter, we refer to (16) as PI_1. Throughout, our analysis is fairly elementary. We refer the interested reader to JJ or Monti (2003) for deeper analysis.

The PI_1-controlled "closed-loop" target system, obtained by substitution in (16), is given by

$$
\begin{aligned}
\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + \Gamma_{11} \Delta Y_{t-1} + \Gamma_{12}(b_2 + \Gamma_{20}\Delta Y_{t-1} + \Gamma_{12}^{-1}[\alpha_1 \lambda_1' + \kappa_1(\beta_1 + \lambda_1)']Y_{t-1}) + \upsilon_{1t} \\
&= b_1 + \Gamma_{12}b_2 + \left((\alpha_1 + \kappa_1)(\beta_1 + \lambda_1)'\right) Y_{t-1} + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\Delta Y_{t-1} + \upsilon_{1t} \\
&= (b_1 + \Gamma_{12}b_2) + \widetilde{\alpha}_1 \widetilde{\beta}_1' Y_{t-1} + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\Delta Y_{t-1} + \upsilon_{1t}.
\end{aligned}
$$

where the cointegration vector is $\widetilde{\beta}_1 \equiv \beta_1 + \lambda_1$ and the effect of equilibrium departures is $\widetilde{\alpha}_1 \equiv \alpha_1 + \kappa_1$. As long as ΔZ_{t-1} structurally causes ΔY_t ($\Gamma_{12} \neq 0$), the dynamics of the open-loop and closed-loop systems can be very different. Because of partial equilibrium recursivity, the open-loop deep parameters $(b_1, \Gamma_{11}, \Gamma_{12})$ and dynamics are invariant to policy interventions. On the other hand, the closed-loop coefficients and dynamics are not invariant to policy interventions, as the policymaker's coefficients (b_2, Γ_{20}) determine the closed-loop dynamics.

In what follows, we pay particular attention to comparing target system properties with and without control, that is, to comparing the closed- and open-loop dynamics.

6.1 PI Control without Open-Loop Cointegration

With PI_1 control and in the absence of cointegration, the closed-loop target system is

$$
\Delta Y_t = (b_1 + \Gamma_{12}b_2) + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\Delta Y_{t-1} + \kappa_1 \lambda_1' Y_{t-1} + \upsilon_{1t}. \tag{17}
$$

To attain the policymaker's long run goal $E(\Delta Y_t) = \gamma_0$, take expectations on both sides of (17) at the steady state, which gives

$$
\gamma_o = (b_1 + \Gamma_{12}b_2) + \kappa_1 \mu_1 + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\gamma_o, \tag{18}
$$

where $\mu_1 \equiv \lambda_1' E(Y_{t-1})$. To determine μ_1, observe that

$$
\lambda_1' E(Y_t) = \lambda_1' E(Y_{t-1}) + \lambda_1' E(\Delta Y_t),
$$

so that in the steady state

$$
\lambda_1' \gamma_o = 0.
$$

This restriction defines the feasible long-run policy targets. We call these "cointegration feasible." Any policy goal not satisfying this condition is unattainable in this system. Essentially, the cointegrating relations remove $r < N_1$ degrees of freedom from the policymaker's discretion.

This may enable the policymaker to focus on controlling a linear combination of ΔY_t using a smaller set of policy control variables. We discuss this case in the appendix.

To solve for μ_1, multiply both sides of (18) by λ_1'. This gives

$$\lambda_1'(b_1 + \Gamma_{12}b_2) + \lambda_1'\kappa_1\mu_1 + \lambda_1'(\Gamma_{11} + \Gamma_{12}\Gamma_{20})\gamma_o = 0 \qquad \text{so that}$$

$$\mu_1 = -(\lambda_1'\kappa_1)^{-1}\lambda_1'[(\Gamma_{11} + \Gamma_{12}\Gamma_{20})\gamma_o + (b_1 + \Gamma_{12}b_2)],$$

where we use the fact that κ_1 and λ_1 have full column rank, ensuring that $\lambda_1'\kappa_1$ is nonsingular. Substituting this into (18) gives

$$\gamma_o = (b_1 + \Gamma_{12}b_2) - \kappa_1(\lambda_1'\kappa_1)^{-1}\lambda_1'[(\Gamma_{11} + \Gamma_{12}\Gamma_{20})\gamma_o + (b_1 + \Gamma_{12}b_2)] + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\gamma_o.$$

Collecting terms, we get

$$\{I - J_1(\Gamma_{11} + \Gamma_{12}\Gamma_{20})\}\gamma_o = J_1(b_1 + \Gamma_{12}b_2), \tag{19}$$

where $J_1 \equiv I - \kappa_1(\lambda_1'\kappa_1)^{-1}\lambda_1'$ is the "long-run impact matrix" (see JJ, eq.(4)). Observe that $J_1(\Gamma_{11} + \Gamma_{12}\Gamma_{20})$ plays the role of a "first order autocorrelation". This determines the speed of convergence to the policy steady state, provided the roots of the associated characteristic equation are inside the unit circle. Generally, J_1 is singular, so expressing b_2 in terms of Γ_{20} gives

$$J_1\Gamma_{12}b_2 = [I - J_1(\Gamma_{11} + \Gamma_{12}\Gamma_{20})]\gamma_o - J_1 b_1.$$

This is an under-determined system of equations, so there are generally many ways to choose b_2 satisfying these equations for given B_{20} and γ_o. One way to proceed in such cases is to minimize a convex function of b_2 (for example, $b_2'b_2$) subject to (19). We emphasize that only choices for γ_o satisfying $\lambda_1'\gamma_o = 0$ give valid choices for Γ_{20} and b_2. On the other hand, we have

Proposition 6.1 *When the open-loop target system is integrated but not cointegrated, PI_1 control can initiate cointegration in the closed-loop target system.*

Observe that with PI_1 control, since Y_t is $I(1)$, $\Delta Z_t = b_2 + \Gamma_{20}\Delta Y_t + \Gamma_{12}^{-1}\kappa_1\lambda_1'Y_t$ is generally $I(1)$, so Z_t is $I(2)$. There could be linear combinations of Z_t that are $I(1)$, but as the details are involved and the circumstances special, we do not pursue this. Though the target and control are generally of different orders, we see immediately from the control equations that $\Delta Z_t - \Gamma_{12}^{-1}\kappa_1\lambda_1'Y_t$ is $I(0)$, so ΔZ_t and Y_t are cointegrated. Interestingly, this cointegration is generated by the causal control relation between ΔZ_t and Y_t.

6.2 PI Control with Open-Loop Cointegration

Next, consider a PI_1 system with a cointegrated open-loop target. This is the general case of equation (16), which we restate here for convenience:

$$\begin{aligned}
\Delta Y_t &= b_1 + \alpha_1\beta_1'Y_{t-1} + \Gamma_{11}\Delta Y_{t-1} + \Gamma_{12}\Delta Z_{t-1} + \upsilon_{1t}, \\
\Delta Z_t &= b_2 + \Gamma_{20}\Delta Y_t + \Gamma_{12}^{-1}[\alpha_1\lambda_1' + \kappa_1(\beta_1 + \lambda_1)']Y_t \quad t = 1, 2, \ldots
\end{aligned}$$

We have already seen that the closed-loop system takes the form

$$\Delta Y_t = (b_1 + \Gamma_{12}b_2) + \widetilde{\alpha}_1\widetilde{\beta}_1' Y_{t-1} + (\Gamma_{11} + \Gamma_{12}\Gamma_{20})\Delta Y_{t-1} + v_{1t}.$$

Here, the cointegrating vector is $\tilde{\beta}_1 \equiv \beta_1 + \lambda_1$, and the effect of equilibrium departure is $\tilde{\alpha}_1 \equiv \alpha_1 + \kappa_1$. Not only does PI_1 control permit the policymaker to modify the cointegration-feasible policies, it permits the policymaker to adjust the response to equilibrium departures. We have

Proposition 6.2 *When the open-loop target system is integrated or cointegrated, PI_1 control can modify the closed-loop target system cointegrating coefficients $(\alpha_1 + \kappa_1)$, $(\beta_1 + \lambda_1)$.*

As for the case of PI_1 control without open loop cointegration, Z_t is generally $I(2)$. We do not pursue an analysis of the possible cointegrating properties of Z_t as these can arise only under very special circumstances. Also, we see immediately from the control equations that $\Delta Z_t - B_{12}^{-1}[\alpha_1\lambda_1' + \kappa_1(\beta_1 + \lambda_1)']Y_t$ is $I(0)$, so ΔZ_t and Y_t are again causally cointegrated.

Parallel to the PI_1 control case without open loop cointegration, here the policymaker can attain γ_o such that $\tilde{\beta}_1'\gamma_o = 0$, where $\tilde{\beta}_1 = \beta_1 + \lambda_1$. Here, we chose b_2 and B_{20} such that

$$\{I - \tilde{J}_1(B_{11} + B_{12}B_{20})\}\gamma_o = \tilde{J}_1(b_1 + B_{12}b_2),$$

where now $\tilde{J}_1 \equiv I - \tilde{\alpha}_1(\tilde{\beta}_1'\tilde{\alpha}_1)^{-1}\tilde{\beta}_1'$.

Before concluding our discussion of PI control, we note certain other special cases of PI control that one may encounter in applications. A special case of PI_1 control fixes κ_1 at zero, with $\lambda_1 \neq 0$. We call this PI_2 control. This is limited, in that PI_2 cannot initiate cointegration in the closed-loop target system if cointegration is absent in the open-loop target system. Also, PI_2 does not allow the policymaker to alter the effect α_1 of equilibrium departures on the target system. Another special case of PI_1 control fixes λ_1 at zero, with $\kappa_1 \neq 0$. We call this PI_3 control. This is the special case of PI_1 where the policymaker only modifies the speed of error correction. Both PI_2 and PI_3's properties can be inferred from our discussion of PI_1.

Finally, we mention the class of P control policy rules, which take the form

$$\Delta Z_t = b_2 + \Gamma_{20}\Delta Y_t.$$

This is a restricted version of PI control, where the Y_t term is dropped from the policymaker's control system. The famous Taylor rule (Taylor, 1993) is of this form. The properties of P control follow from our discussion of PI_1 control, so for conciseness we do not discuss these further here. We refer the interested reader to the online appendix available at [*online link here*], where we cover the cases of P control as well PI_2 and PI_3 in more detail.

7 Illustrative application

The effectiveness of U.S. Federal Reserve policy has been the focus of many previous theoretical and empirical studies. See, e.g., Bernanke and Blinder (1992), Christiano, Eichenbaum, and

Evans (1996), Leeper, Sims, and Zha (1996), and Hamilton (2008), and the references given there. Here, we apply our framework to illustrate how one can examine this issue. We emphasize that because our goal here is only to illustrate useful methods, we will not push this investigation as far as would be required to obtain a model sufficiently well specified to deliver definitive insights about Fed policy. Thus, we pay attention to indicators of model shortcomings without necessarily resolving the issues identified. As will become apparent, resolving the issues uncovered in fact require an extensive modeling effort well beyond what we can feasibly undertake here.

For our illustration, we examine the impact of Fed policy on macroeconomic variables Y_t (inflation, unemployment, output, and oil prices) through the Federal Funds rate, Z_t. While the Fed does not directly control this rate, it sets its target value; daily open market operations then align the Fed Funds rate closely to the target value. This corresponds exactly to the case of imperfect control examined above.

Specifically, we let Z_t be the effective Federal Funds rate (taken from the Board of Governors of the Federal Reserve System), and we let Y_t include (i) the natural logarithm of the total US industrial production index, seasonally adjusted (taken from the Board of Governors of the Federal Reserve System); (ii) the natural logarithm of the seasonally adjusted US total unemployment rate for all individuals aged 16 and over (taken from the Bureau of Labor Statistics); (iii) the US inflation rate, computed as the 12-month difference of the natural logarithm of the consumer price index for all urban consumers (taken from the Bureau of Labor Statistics); and (iv) the natural logarithm of the Cushing, OK WTI Spot Price FOB (taken from the Energy Information Administration). The data are monthly, covering January, 1986 through December, 2007, a total of $T = 262$ observations, adjusting for lags and differencing.

We start by focusing on a DGP accommodating PI control of the target first differences,

$$\Delta Y_t = b_1 + D_1 Y_{t-1} + \Gamma_{11} \Delta Y_{t-1} + \Gamma_{12} \Delta Z_{t-1} + \upsilon_{1t}, \tag{20}$$

$$\Delta Z_t = b_2 + \Gamma_{20} \Delta Y_t + D_2 Y_t + \upsilon_{2t} \quad t = 1, 2, ... \tag{21}$$

When $D_1 = \alpha_1 \beta_1'$ and $D_2 = \Gamma_{12}^{-1}[\tilde{\alpha}_1 \tilde{\beta}_1' - D_1]$, we have the system with open-loop cointegration and PI_1 control considered above. This is a special case of eq. (1) with possible cointegration in the open-loop target subsystem, and, consistent with PI_1 control, also possibly between Y_t and ΔZ_t. Here, we permit policy implementation to be noisy, as an implementation error υ_{2t} appears in the control subsystem. This does not affect any of the results of the previous section, as this noise just introduces a mean zero component $B_{12} \upsilon_{2,t-1}$ into the closed-loop target system.

The associated closed-loop target system is

$$\Delta Y_t = c_1 + \tilde{\alpha}_1 \tilde{\beta}_1' Y_{t-1} + C_{11} \Delta Y_{t-1} + \eta_{1t}, \tag{22}$$

where $c_1 \equiv b_1 + \Gamma_{12} b_2$, $\tilde{\alpha}_1 \tilde{\beta}_1' = D_1 + \Gamma_{12} D_2$, $C_{11} \equiv \Gamma_{11} + \Gamma_{12} \Gamma_{20}$, and $\eta_{1t} \equiv \upsilon_{1t} + \Gamma_{12} \upsilon_{2,t-1}$. Note that as (22) is a standard cointegrated system, its coefficients can be estimated using standard

21

methods, such as the Engle-Granger (1987) estimator or the methods of Johansen (1995).

An important feature of both the open-loop and closed-loop systems is that they might be subject to structural shifts. The target system (20) might be subject to "exogenous" shifts, that is, shifts arising outside the controlled system. The policy system (21) is subject to policy interventions associated with policy regime changes, tuning exercises, or exogenous shifts in the target system. Endogenous shifts in the target system (20) represent a failure of invariance, in which case the Lucas critique operates. Because the closed-loop target system contains coefficients from both the target and control structures, it follows that the coefficients of (22) can shift for any of the reasons just given.[6] To examine whether the closed-loop target system (22) is stable, we use the recursive log-likelihood test described in Juselius (2006). While there are several different tests that could be used,[7] we focus on the Juselius (2006) test as it accommodates cointegration and it permits us to examine both the short-run and long-run components of the DGP. The test statistic is computed recursively, starting from a baseline period and extending backward or forward in time by adding observations to the baseline. Here we apply the backward recursion. Let T_1 index the first observation in the baseline sample considered in the recursion. We set T_1 to December 2002, ensuring five years of the data in the baseline period, $T_1, ..., T$. The statistic for subsamples including observations $t_1, ..., T$, with $t_1 = T_1, T_1 - 1, ..., 1$ is

$$Q_T\left(t_1\right) = \frac{t_1}{T}\sqrt{\frac{T}{2p}}\left[\log\left|\widehat{\Omega}_{t_1}\right| - \log\left|\widehat{\Omega}_T\right| + \frac{1}{T}\left(0.5p\left(p+1\right) + r + p\left(k-1\right) + 1\right)\left(1 - \frac{t_1}{T}\right)\right],$$

where $p\ (=4)$ is the number of equations in the system and r is the cointegrating rank, as estimated by the Johansen (1995) procedure. Here, we find a single $(r=1)$ cointegrating relation. The estimated variance-covariance matrices for the sub-sample including observations $t_1, ..., T$ and the full sample are $\widehat{\Omega}_{t_1}$ and $\widehat{\Omega}_T$, respectively.

Under the null hypothesis of DGP stability, the 95% quantile for the test is 1.36. We display two versions of the test in Figure 1. The first, labelled $X(t)$, is based on the full model, whereas the second, $R1(t)$, is based on the long-run concentrated model, where the short-term variables have been concentrated out. This latter version is based on the model obtained after the first stage of the Johansen (1995) procedure.

---

[6]It is also in principle possible for the policy authority to undertake policy interventions that precisely offset exogenous structural changes to (20), leaving (22) unchanged. Nevertheless, this requires a sufficient degree of knowledge and flexibility that exogenous structural shifts in (20) are likely in practice to be reflected in structural shifts in (22).

[7]See for example the methods of Bai and Perron (1998) or Juselius (2006), or the indicator saturation methods of Hendry, Johansen, and Santos (2008) (HJS); see also Johansen and Nielsen (2009).

Recursively calculated log likelihood test
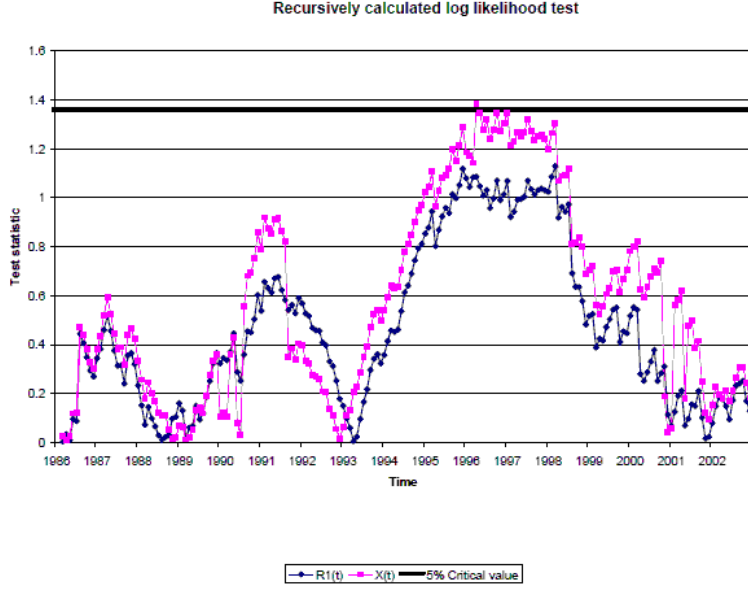
Recursive log likelihood function. The baseline sample is 2002:12 - 2007:12

As the graph shows, we do not reject stability for the closed-loop system using the $R1$ form
of the test. The statistic for the X version crosses the critical value in April, 1996, but by a
small amount. In line with our illustrative intent, we take these results as largely consistent
with stability for the closed-loop system and proceed with our analysis; but we keep in mind
that there could be some short-term instability.

We next proceed to test which type of proportional control best describes the data. We start
by noting that while (22) ensures that Y_t is $I(1)$, (20) and (21) allow for two main possibilities
for ΔZ_t. The first is that ΔZ_t is $I(0)$; the second is that ΔZ_t is $I(1)$. The first possibility
arises in either of two cases. It is easily checked that ΔZ_t is $I(0)$ with open loop cointegration
$(D_1 = \alpha_1\beta_1')$ and either (i) P control $(\tilde{\alpha}_1\tilde{\beta}_1' = \alpha_1\beta_1')$ or (ii) PI$_3$ control $(\tilde{\beta}_1 = \beta_1$, but $\tilde{\alpha}_1 \neq \alpha_1)$.
On the other hand, with either PI$_1$ or PI$_2$ control we have $\Delta Z_t \sim I(1)$, regardless of open loop
cointegration. Thus, a test of the null hypothesis that ΔZ_t is $I(1)$ vs. the $I(0)$ alternative is
a test for PI$_1$ or PI$_2$ control vs. P or PI$_3$ control, assuming correct specification of (20) and
(21). When we apply an augmented Dickey-Fuller test to Z_t, the Federal Funds rate variable,
we do not reject the unit root null, whereas the same test run on the first differences does reject
the unit root null at the 1% level. Taken at face value, these results suggest that $\Delta Z_t \sim I(0)$
and that we are in the world of either P control or PI$_3$ control. We next test the hypothesis
that P control is in effect, against the alternative of PI$_3$ control. Note that with either P or PI$_3$
control, (21) can be estimated by regressing ΔZ_t on a constant, ΔY_t, and $\tilde{\Psi}_t$, where $\tilde{\Psi}_t = \beta_1\widetilde{Y}_t$.

Under P control, the coefficients on $\tilde{\Psi}_t$ are zero, while under PI$_3$ control they are not. Given the previous finding of $\Delta Z_t \sim I(0)$, we can test the hypothesis of P control vs. PI$_3$ control by applying standard tests for zero values of these coefficients. Specifically, we find that the coefficient on $\tilde{\Psi}_t$ is significant ($p = 0.07$), so we reject the P control hypothesis.[8]

Having determined the form of proportional control that best fits the data, we are now almost ready to test Fed policy effectiveness, that is, to test if the coefficients Γ_{12} in the open loop control system are significantly different from zero. Nevertheless, we take two additional steps. First, we investigate whether either or both (20) or (22) are misspecified. Note that as $\Delta Z_t \sim I(0)$ implies $\tilde{\beta}_1 = \beta_1$, a specification test for (20) and (22) can be carried out by formally testing the hypothesis $\tilde{\beta}_1 = \beta_1$. In particular, a rejection of the null hypothesis $\tilde{\beta}_1 = \beta_1$ would indicate that either or both (20) or (22) are misspecified. A test for $\tilde{\beta}_1 = \beta_1$ can be conveniently performed by estimating a version of (20) modified by including $\tilde{\Psi}_{t-1}$ as well as Y_{t-1}, using the Johansen procedure to estimate and test the cointegrating rank. Under the null hypothesis, the cointegrating rank will be zero, as cointegration is already captured by $\tilde{\Psi}_{t-1}$. If the Johansen procedure rejects the null of no cointegration, one must reject $\tilde{\beta}_1 = \beta_1$. When we test whether the cointegrating rank in this augmented regression is zero using Johansen's (1995) trace statistic test, we find that we cannot reject the zero-rank hypothesis at the 10% level, so we conclude that there is no clear evidence of misspecification on this basis.[9]

Second, we more closely investigate the relation between ΔZ_t and $(Y_t, \Delta Y_t)$, to asses potential difficulties that might hinder the identification of Γ_{12}. An immediate simple diagnostic is the R^2 from the regression of ΔZ_t on a constant, Y_t and ΔY_t, i.e., (21). Ideally, we would like to find a good but not perfect fit. Too loose a fit suggests that the policy instrument is not actually being used to manipulate the supposed target or that some control rule other than P or PI is in use. A good fit suggests at least that the policymaker believes Γ_{12} is not zero. The R^2 of this regression is 0.153, so we conclude that collinearity of the policy instrument with ΔY_t and Y_t is not an issue. On the other hand, this somewhat low R^2 suggests that the control equation may not be fully capturing the Fed's behavior.

We are now ready to test whether the coefficients Γ_{12} in the open-loop control system are

---

[8]In the event where $\Delta Z_t \sim I(1)$, two possibilities for conducting inference on the B_{12} coefficients from the open loop target subsystem (20) suggest themselves. The first is to directly apply the results of Park and Phillips (1988, 1989). The resulting inference may be non-standard, however. An apparently simpler possibility is to use a two step procedure where first a modified version of (22) is invoked to estimate B_{12}, namely $\Delta Y_t = c_1 + \tilde{\alpha}_1 \tilde{\beta}_1' Y_{t-1} + C_{11}\Delta Y_{t-1} + B_{12}v_{2,t-1} + v_{1t}$, where $v_{2,t-1}$ is replaced by an estimate from (21), say $\hat{v}_{2t-1}$. Using the estimator for B_{12} from this two-stage procedure, say $\tilde{B}_{12}$, should only involve standard $\sqrt{T}$ inference, although adjustment for the effects of the first-stage estimation may be required. Note also that estimating the policy equation (21) involves a regression of an $I(1)$ variable (ΔZ_t) on an $I(1)$ variable (Y_t) with cointegration between them, as in Stock (1987). In fact, one can identify and consistently estimate b_2 and B_{20} from (21) with a variety of standard procedures, plausibly with standard $\sqrt{T}$ asymptotics.

[9]Alternatively, one could test the rank of B_{12} by using the convenient methods recently given by Camba-Méndez and Kapetanios (2008).

Target variable	Coeff	Std Err	t-stat	p-value
IPI total	0.005	0.002	2.928	0.004
Inflation	0.087	0.078	1.116	0.265
Unemployment	-0.033	0.008	-4.332	0.000
Cushing oil price	0.052	0.024	2.200	0.029

Table 1: Open loop estimates for the matrix B_{12}. Results are obtained by applying the Johansen (1995) procedure. The computed standard errors are White (1980) robust standard errors.

significantly different from zero. Noting that $\Delta Z_t \sim I(0)$ implies that the estimator of Γ_{12}, say $\hat{\Gamma}_{12}$, has standard properties, we can proceed by using standard inference. As Table 2 shows, the estimates of Γ_{12} are significantly different from zero for both the IPI and the unemployment variables, and are borderline significant for the oil price variable.[10] On the other hand, altering the Fed Funds rate in an attempt to *directly* affect the change in the inflation rate would appear to be ineffective, as the insignificant coefficient in the inflation equation implies. Nevertheless, *indirect* effects arise from the various feedforward channels, i.e., through unemployment and IPI.

We also examine the estimated residual first-order autocorrelations of the closed-loop system (22) and the open-loop target system (20), as a diagnostic for exogeneity of $(Y_{t-1}, \Delta Y_{t-1}, \Delta Z_{t-1})$. Overall, these do not indicate obvious problems. For the closed-loop system these are -0.0152 (IPI), 0.0659 (inflation), -0.0304 (unemployment), and -0.0062 (oil). For the open-loop target system, these are -0.0129 (IPI), 0.0676 (inflation), -0.0563 (unemployment), and -0.0042 (oil).

On the other hand, the estimated autocorrelation coefficient for the control equation is .284, suggesting some form of misspecification, dynamic or otherwise. To see if a simple autocorrelation adjustment can resolve matters, we apply the Cochrane-Orcutt technique to (21). This gives estimated autocorrelation of .462 and very different coefficient estimates. As there is no lagged dependent variable in the control equation, this suggests that more than simple autocorrelation may be at work. Plausibly, there may be one or more omitted variables.

To keep the scope of our example manageable, we just examine the possibility that the policymaker might be executing PI$_3$-derivative (PI$_3$D) control, in which case a previously omitted term $\Delta^2 Y_t$ should appear in the control equations[11], and (21) is modified to

$$\Delta Z_t = b_2 + \Gamma_{20}\Delta Y_t + \Gamma_{21}\Delta^2 Y_t + D_2 Y_t + \upsilon_{2t} \quad t = 1, 2, ... \tag{23}$$

When we estimate the PI$_3$D control equation, we find that the R^2 increases from .153 to .298. This marked increase is due mainly to $\Delta^2 Y_t$ terms associated with inflation and unemployment; at face value, these terms have a clear role to play. On the other hand, residual autocorrelation drops to .241. This is a move in the right direction, but clearly PI$_3$D control is not the whole story. Applying Cochrane-Orcutt to (23), we find an estimated autocorrelation coefficient of

---

[10] Tables 2 and 3 present White's (1908) robust standard errors.

[11] As can be readily verified, the addition of D control has no impact on the cointegration results of section 6.

.296. Encouragingly, the signs and magnitudes of the estimated coefficients in this equation, while differing somewhat from the OLS estimates, do not change nearly as much as they did when considering only PI_3 control. In particular, we reject the PD control hypothesis in favor of PI_3D. In line with our illustrative intent, we proceed, assuming PI_3D control with autocorrelation. We further investigate this equation below, however.

PI_3D control also modifies the closed-loop system (22) to include a $\Delta^2 Y_t$ term. We refer to this as the "closed-loop PID system." Note that our prior omission of the $\Delta^2 Y_t$ term could explain the apparent short-term instability earlier found by the Juselius test. Indeed, when we re-run the Juselius test on the closed-loop PID system, we find no evidence at all of instability; the maximum value for the X version of the test is only about 1.06. That for the $R1$ form is smaller. We also find a very similar value for the cointegrating vector.

Next, we perform tests to explore whether (20), (23), or the closed-loop PID system are linear or whether there may be neglected nonlinearity. There is an extensive literature on testing for neglected nonlinearity in regression analysis, ranging from Ramsey's (1969) classic RESET procedure to modern neural network or random field tests. (See, for example, Lee, White and Granger, 1993; Hamilton, 2001; and Dahl and Gonzalez-Rivera, 2003.) The methods of WL for testing linearity (CI test regression 1 and 2) are quite convenient. One can also test for encompassing (e.g., Hendry and Mizon, 1982), the information matrix equality (White, 1982) and other indicia of misspecification, as detailed, for example, in White (1990). These methods can be straightforwardly applied to (20), (21), or (22). Here we test for neglected nonlinearity by adopting a neural network-based method. The idea is to augment the regressors in a given equation with neural network terms, as in White's (2006b) QuickNet procedure, and then test whether the coefficients of the neural network terms are all zero. This class of tests has been found to have good power to detect neglected nonlinearity. More specifically:

- for eq. (20), we use a Wald statistic for each equation $h = $ IPI, inflation, unemployment, oil price, to test the joint hypothesis $\zeta_{11h} = ... = \zeta_{1kh} = 0$ in:

$$
\begin{aligned}
\Delta Y_{th} &= b_{1h} + D_{1h}^* \beta_1' Y_{t-1} + \Gamma_{11h} \Delta Y_{t-1} + \Gamma_{12h} \Delta Z_{t-1} \\
&\quad + \sum_{j=1}^{k} G\left(\gamma_{10hj} + \gamma_{11hj}\beta_1' Y_{t-1} + \gamma_{12hj}\Delta Y_{t-1} + \gamma_{13hj}\Delta Z_{t-1}\right)\zeta_{1jh} + \varepsilon_{1th}
\end{aligned}
\tag{24}
$$

- for eq. (23), we use a Wald statistic to test the joint hypothesis $\zeta_{21} = ... = \zeta_{2k} = 0$ in:

$$
\begin{aligned}
\Delta Z_t &= b_2 + \Gamma_{20}\Delta Y_t + \Gamma_{21}\Delta^2 Y_t + D_2^* \tilde{\beta}_1' Y_t + \rho\hat{\varepsilon}_{2,t-1} \\
&\quad + \sum_{j=1}^{k} G\left(\gamma_{20j} + \gamma_{21j}\Delta Y_t + \gamma_{22j}\Delta^2 Y_t + \gamma_{23j}\tilde{\beta}_1' Y_t + \gamma_{24j}\hat{\varepsilon}_{2,t-1}\right)\zeta_{2j} + \varepsilon_{2t}
\end{aligned}
\tag{25}
$$

k	1	2	3	4	5	Row BH
Δ IPI$_t$	0.1294	0.0227	0.0233	0.0004	0.0001	0.0004
Δ Inflation$_t$	0.1280	0.1084	0.0689	0.0404	0.0307	0.1280
Δ Unemployment$_t$	0.3170	0.0500	0.0497	0.0688	0.0794	0.1588
Δ Oil price$_t$	0.2221	0.2528	0.1327	0.1173	0.0752	0.2528
Col BH	0.3170	0.0906	0.0932	0.0016	0.0003	0.0016

Table 2: Misspecification tests for the target subsystem

k	1	2	3	4	5	Row BH
ΔZ_t	0.0880	0.0215	0.0004	0.0002	0.0000	0.0002

Table 3: Misspecification tests for the PID control equation

- for the closed-loop PID system, we use a Wald statistic for each equation $h =$ IPI, inflation, unemployment, oil price, to test the joint hypothesis $\zeta_{31h} = ... = \zeta_{3kh} = 0$ in:

$$\Delta Y_{th} = c_{1h} + \tilde{\alpha}_{1h}\tilde{\beta}_1' Y_{t-1} + C_{11h}\Delta Y_{t-1} + C_{12h}\Delta^2 Y_{t-1} \qquad (26)$$
$$+ \sum_{j=1}^{k} G\left(\gamma_{30j} + \gamma_{31j}\tilde{\beta}_1' Y_{t-1} + \gamma_{32j}\Delta Y_{t-1} + \gamma_{33j}\Delta^2 Y_{t-1}\right)\zeta_{3jh} + \eta_{1th}.$$

In these regressions, β_1 and $\tilde{\beta}_1$ are replaced by their estimates. Estimated lagged errors $\hat{\varepsilon}_{2,t-1}$ are included in (25) to accommodate the autocorrelation of the control equation errors.

The function G is an activation function from the class of generically comprehensively revealing (GCR) functions (see Stinchcombe and White, 1998). We use a ridgelet function, $G(x) = \left(-x^5 + 10x^3 - 15x\right)\exp\left(-x^2/2\right)$. Other examples of GCR functions include logistic cdf, normal pdf, etc. We call terms involving G "hidden unit" terms, consistent with the artificial neural network literature. The integer k controls the allowed degree of nonlinearity. We choose γ's from a set of candidates, constructed as in Huang and White (2009).

Tables 2-4 show the Wald statistic $p-$values for each equation and each k. BH denotes the Bonferroni–Hochberg adjusted $p-$values (Hochberg 1988). The right lower corner element is the BH $p-$value for the panel as a whole.

The message from these tests is that the control system equation is misspecified, while three

k	1	2	3	4	5	Row BH
Δ IPI$_t$	0.0364	0.0122	0.0000	0.0000	0.0000	0.0001
Δ Inflation$_t$	0.0143	0.0005	0.0002	0.0000	0.0001	0.0002
Δ Unemployment$_t$	0.0147	0.0015	0.0011	0.0007	0.0003	0.0014
Δ Oil price$_t$	0.0476	0.0433	0.0309	0.0436	0.0376	0.0476
Col BH	0.0441	0.0020	0.0002	0.0001	0.0001	0.0003

Table 4: Misspecification tests for the closed-loop PID system

of the four equations in the target subsystem (inflation, unemployment, and oil price) appear to be well specified. However, the IPI equation appears to be misspecified, as indicated by the Bonferroni–Hochberg adjusted $p-$value (0.0004). Not surprisingly, then, the closed-loop PID equations are also found to be misspecified. Clearly, simple linear structures do not adequately capture important features of the data. These findings are in line with previous studies of the behavior of Fed monetary policy. For example, our findings of neglected nonlinearities in the control system equations could in part be explained by the presence of asymmetries in the Fed's preferences, which in turn are related to the state of the business cycle. See, for example, the work of Cukierman (2000), Gerlach (2000), Bec et al. (2002), Nobay and Peel (2003), and Ruge-Murcia (2002, 2004). Also, Mizen et al. (2005) and Tillman (2011) show that if the central bank follows a min–max strategy to formulate its monetary policy, then the interest rate reaction function can be well described by nonlinear functions, as the interest rate reacts more strongly to inflation when inflation is further away from target. In other words, as inflation increases, the worst-case perception of the Phillips curve slope becomes larger, thus inducing the Fed to implement a larger interest rate adjustment. Mizen et al. (2005) use quantile regressions to document these nonlinearities, while Tillman (2011) finds evidence to support these nonlinearities for post-1982 U.S. data. Future research can be productively directed toward examining the adequacy and implications of more flexible specifications, such as the above neural network specifications.

One interpretation of our findings is that although the Juselius (2006) tests accord with structural stability for our system with PI_3D control, it is possible that the nonlinearity tests are detecting shifts in the short-run structure, against which the X form of the Juselius test does not have power. Similarly, if one were to test for and find structural shifts in the linear PI_3D system using other methods, e.g., those of Bai and Perron (1998) or HJS, one could well be detecting neglected nonlinearities. Thus, only after disentangling these possibilities does it make sense to conduct tests for recursivity or invariance, as discussed above. Neverthless, we provide some suggestions for ways in which one could test for recursivity or invariance within this framework. One possibility is to apply methods of Engle and Hendry (1993), who describe testing invariance without imposing weak exogeneity.[12] Another possibility, which relates to the methods of Hendry and Mizon (1998) and Krolzig and Toro (2002), is to use data spanning at least one policy intervention $(b_2, \Gamma_{20}) \rightarrow (b_2^*, \Gamma_{20}^*)$ to estimate a version of (20) augmented by including a vector of dummy variables, say d_{0t}, whose $i-$th element d_{0it} is zero prior to policy intervention i and is one thereafter. The dates of policy interventions can be specified a priori on theoretical or institutional grounds or can be determined empirically from estimation of (21), using, for example, methods of Bai and Perron (1998). This augmented structure has the form

$$\Delta Y_t = b_1 + D_1 Y_{t-1} + \Gamma_{11}\Delta Y_{t-1} + \Gamma_{12}\Delta Z_{t-1} + \Pi_1 d_{0t} + \upsilon_{1t}.$$

---

[12]See also Hendry (1988) and Hendry and Santos (2009). Hendry and Massman (2006) survey and extend the concept of co-breaking, directly relevant here.

Under the null of recursivity, $\Pi_1 = 0$; evidence of departures from zero is evidence against recursivity. When ΔZ_t is $I(0)$, one can apply standard methods to test $\Pi_1 = 0$. Note that to adopt this approach, the assumption of target structure stability needs to be maintained. A drawback of this test is that unsuspected exogenous structural shifts in the target system could confound its results, leading to false rejections. A procedure not subject to this difficulty involves constructing a sequence of dummies $\{(d_{1t}, d_{2t})\}$ such that $d_{1t} = 1$ if there is a structural shift in period t in the target system (20) and $d_{1t} = 0$ otherwise; and $d_{2t} = 1$ if there is a structural shift in in period t in the policy system (21) and $d_{2t} = 0$ otherwise. One then regresses (d_{1t}, d_{2t}) on its lags and tests whether d_{1t} is structurally caused by lags of d_{2t}. Under the null of recursivity, there can be no such causality; otherwise, causality will be present. The challenge for this test is that it may require a relatively long data history with many breaks in order to have power.

We close this section with a brief examination of the impacts of Fed policy implied by taking the linear PI$_3$D system estimated here at face value. Specifically, we positively perturb $v_{2,t-1}$ in the PID analog of the closed-loop system (22) by three standard deviations for a single period.
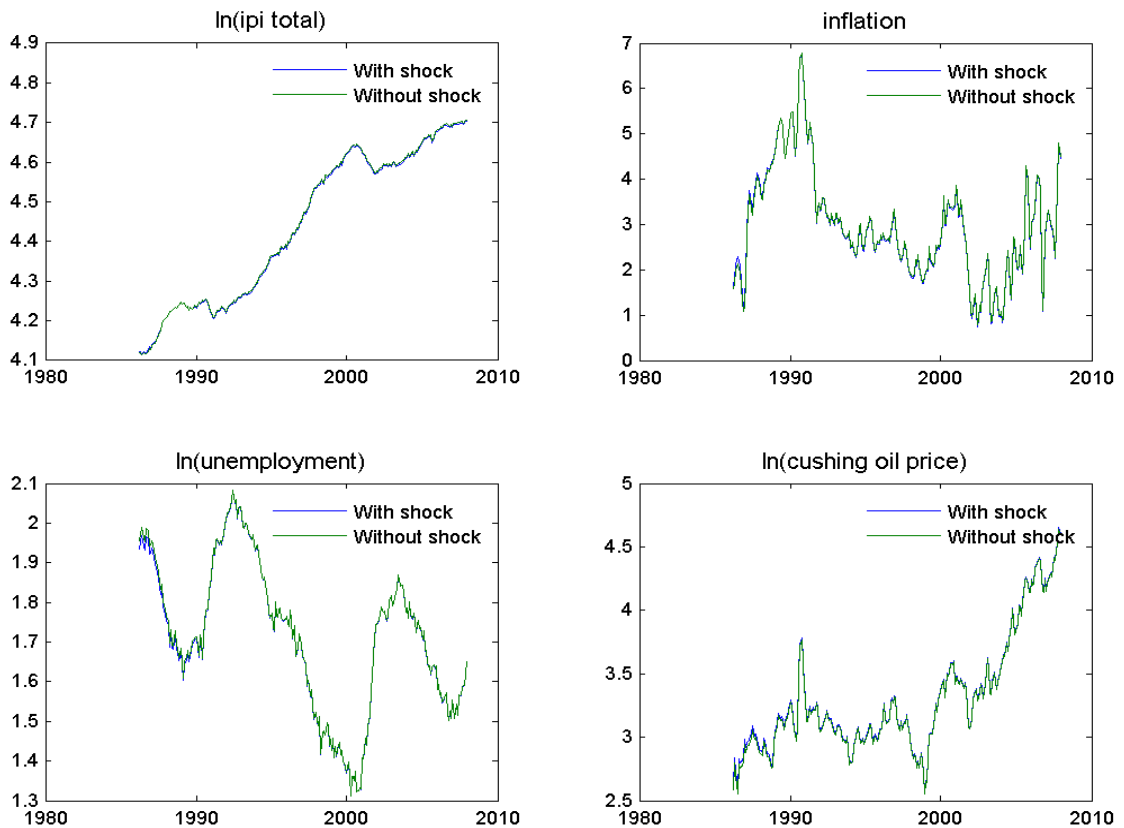
Simulation results for the closed-loop PID system

We choose a relatively large intervention in an attempt to make the effects visually apparent. This amounts to a policy intervention increasing the rate of change of the Fed Funds rate for a single period, somewhat more than doubling the intercept of the control equation for that period. Even though the effect (Γ_{12}) of the Fed Funds rate is statistically significant, Figure 7 shows that the impacts of even this large shock are barely visible, apart from some initial upward effects on unemployment. Inspection of the differences between the series with and without the intervention show that except for oil, each series experiences an initial upward impact, declining to an eventual small negative impact. For oil, the impact is initially negative, but becomes less so, converging to a small negative impact.

8   Summary and Conclusion

One of Clive Granger's long-standing and central concerns was that econometric theory and practice should have direct value to policymakers. Here, we present a framework for economic policy analysis that provides a novel integration of several fundamental concepts at the heart of Granger's contributions to time-series analysis. We work with a dynamic structural system analyzed by WL with well defined causal meaning. The system contains target and control subsystems, with possibly integrated or cointegrated behavior. We ensure the invariance of the target subsystem to policy interventions and thus obviate the Lucas critique using an explicitly causal partial equilibrium recursivity condition, plausible on informational, behavioral, and empirical grounds. Policy effectiveness corresponds to another explicit causality condition. Identification of system coefficients holds given conditional exogeneity, an extension of strict exogeneity distinct from weak exogeneity or its extensions. As we discuss, given conditional exogeneity, Granger causality and structural causality are equivalent. Given structural *non*-causality, Granger causality and the *failure* of conditional exogeneity are equivalent. In this sense, Granger causality is not a fundamental system property requisite for reliable policy analysis, but an important consequence of necessary underlying structural properties.

By relying only on correct model specification and not weak exogeneity, our framework complements the policy analytic framework of Ericsson, Hendry, and Mizon (1998). As we show, our approach readily lends itself to analysis of the structural consequences of a variety of control rules that the policymaker may employ. Among other things, we find that proportional (P) control cannot modify the cointegrating properties of a target system, whereas proportional-integral (PI) control can. In fact, PI control can introduce, eliminate, or broadly modify the cointegrating properties of the uncontrolled target system. Whereas cointegration between target variables and policy instruments is possible but unusual with P control, PI control can easily induce causal cointegration between the target (Y_t) and the policy instruments (ΔZ_t). These properties are preserved under PID control.

The control mode also has interesting implications for estimation, inference, and specification testing in controlled systems. P, PI$_3$, or PI$_3$D control yield $\Delta Z_t \sim I(0)$, which results in standard inference. Other modes of PI or PID control yield $\Delta Z_t \sim I(1)$; the theory of Park and Phillips (1988,1989) applies to these cases.

One of the hallmarks of Clive Granger's work is that it has vigorously stimulated research, often in an astonishing number of different productive directions. Putting a positive spin on the fact that the analysis here leaves a potentially embarrassing number of questions asked but not answered, we are hopeful that, like Clive's work, these unanswered questions will stimulate interest in resolving them. In addition to suggesting the relevance of new theory for inference in partially nonstationary systems with covariates and conditional heteroskedasticity, the analysis here suggests, among other things, opportunities for developing specification tests distinguishing structural shifts and neglected nonlinearities, for studying control of nonlinear systems with cointegration using a misspecified model, for studying covariates in control, for developing methods useful for real-time monitoring of structural change in cointegrated systems, and for analyzing recursive methods of adaptive policy control, robustly able to operate in cointegrated systems subject to exogenous structural shifts. We hope, also, that the practical methods described and illustrated here will, as Clive would have desired, have direct value to policymakers.

References

Ahn, S.K. and G.C. Reinsel (1990), "Estimation for Partially Nonstationary Multivariate Autoregressive Models," *Journal of the American Statistical Association* 85, 813-823.

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

Barnow, B., G. Cain, and A. Goldberger (1980), "Issues in the Analysis of Selectivity Bias," *Evaluation Studies* 5, 42 - 59.

Bec F, Salem MB, Collard F. (2002), "Asymmetries in monetary policy reaction function: evidence for the U.S., French and German central banks", *Studies in Nonlinear Dynamics and Econometrics 6*, No. 2, Article 3.

Bernanke, B. and A. Blinder (1992), "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review* 82, 901-921.

Buiter, W.H. (1984), "Granger Causality and Policy Effectiveness," *Economica* 51, 151-162.

Camba-Méndez, G. and G. Kapetanios (2008), "Statistical Tests and Estimators of the Rank of a Matrix and their Applications in Econometric Modelling," European Central Bank Working Paper No. 850.

Chow, G. (1997). *Dynamic Economics*. Oxford: Oxford University Press.

Christiano, L., M. Eichenbaum, and C. Evans (1996), "The Effects of Monetary Policy

Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics* 78, 16-34.

Chu, C.-S., M. Stinchcombe, and H. White (1996), "Monitoring Structural Change," *Econometrica* 64, 1045-1066.

Crawford, V. and N. Iriberri (2007), "Fatal Attraction: Salience, Naïveté, and Sophistication in Experimental 'Hide-and-Seek' Games," *American Economic Review* 97, 1731-1750.

Cukierman A. (2000), "The inflation bias result revisited", Tel-Aviv University, Mimeo.

Dahl, C.M. and G. Gonzalez-Rivera (2003), "Identifying Nonlinear Components by Random Fields in the US GNP Growth: Implications for the Shape of the Business Cycle," *Studies in Nonlinear Dynamics and Econometrics*, 7, Article 2.

Dawid, P. (1979), "Conditional Independence in Statistical Theory," *Journal of the Royal Statistical Society* Series B 41, 1-31.

Durbin, J. (1970), "Testing for Serial Correlation in Least Squares Regression when Some of the Regressors are Lagged Dependent Variables," *Econometrica* 38, 410-421.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction," *Econometrica* 55, 251-276.

Engle, R.F. and D.F. Hendry (1993), "Testing Super Exogeneity and Invariance in Regression Models," *Journal of Econometrics* 56: 119- 139.

Engle, R.F., D.F. Hendry, and J.-F. Richard (1983), "Exogeneity," *Econometrica* 51, 277-304.

Ericsson, N.R., D.F. Hendry, and G. Mizon (1998), "Exogeneity, Cointegration, and Economic Policy Analysis," *Journal of Business and Economic Statistics* 16, 370-387.

Ermini, L. (1992), "Some Unresolved Issues in the Application of Control Theory to Economic Policy-Making," *Computers and Mathematics with Applications* 24, 125-136.

Fisher, A.M. (1993), "Weak Exogeneity and Dynamic Stability in Cointegrated VARs," *Economics Letters* 43, 167-170.

Gerlach S. (2000), "Asymmetric policy reactions and inflation", Bank of International Settlement working paper.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica* 37, 424–438.

Granger, C.W.J. (1973), "On the Properties of Forecasts Used in Economic Policy Decisions," *Journal of Public Economics* 2, 347-356.

Granger, C.W.J. (1988), "Causality, Cointegration, and Control," *Journal of Economic Dynamics and Control* 12, 551-559.

Granger, C.W.J. (2009), "In Praise of Pragmatics in Econometrics," in J. Castle and N. Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry.* Oxford: Oxford University Press, pp. 59–87.

Granger, C.W.J. and M. Deutsch (1992), "Comments on the Evaluation of Policy Models," *Journal of Policy Modeling* 14, 497-516.

Granger, C.W.J. and P. Newbold (1986). *Forecasting Economic Time Series.* New York: Addison-Wesley (2nd ed.).

Hadley, G. (1961). *Linear Algebra.* Reading, MA: Addison-Wesley.

Hamilton, J.D. (2001), "A Parametric Approach to Flexible Nonlinear Inference," *Econometrica* 69, 537-573.

Hamilton, JD. (2003), "What is an oil shock?", *Journal of Econometrics*, 113: 363–398.

Hamilton, J.D. (2008), "Daily Monetary Policy Shocks and New Home Sales," Journal of Monetary Economics 55 1171-1190.

Hendry, D.F. (1988), "The Encompassing Implications of Feedback Versus Feed-forward Mechanisms in Econometrics," *Oxford Economic Papers* 40, 132-149.

Hendry, D.F. (1995). *Dynamic Econometrics.* Oxford: Oxford University Press.

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

Hendry, D.F. and M. Massmann (2006), "Co-breaking: Recent Advances and a Synopsis of the Literature," *Journal of Business and Economic Statistics* 25, 33-51.

Hendry, D.F. and G. Mizon (1982), "On the Formulation of Empirical Models in Dynamic Econometrics," *Journal of Econometrics* 20, 3-34.

Hendry, D.F. and G. Mizon (1998), "Exogeneity, Causality, and Co-breaking in Economic Policy Analysis of a Small Econometric Model of Money in the UK," *Empirical Economics* 23, 267–294.

Hendry D.F., M. Lu, and G. Mizon (2009), "Model Identification and Nonunique Structure," in J. Castle and N. Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry.* Oxford: Oxford University Press, pp. 343-364.

Hendry, D.F. and C. Santos (2009), "An Automatic Test of Superexogeneity," in T. Bollerslev, J. Russell, and M. Watson (eds.), *Volatility and Time Series: Essays in Honor of Robert F. Engle.* Oxford: Oxford University Press, pp. 164-193.

Hochberg, Y. (1988), "A Sharper Bonferroni Procedure for Multiple Tests of Significance," *Biometrika* 75, 800–802.

Hornik, K., F. Leisch, A. Zeileis, and C. Kleiber (2005) "Monitoring Structural Change in Dynamic Econometric Models," *Journal of Applied Econometrics* 20, 99-121.

Huang, M., and H. White. (2009), "A Flexible Test for Conditional Independence," UCSD Dept. of Economics Discussion Paper.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Auto-Regressive Models.* Oxford: Oxford University Press.

Johansen, S. and K. Juselius (2001), "Controlling Inflation in a Cointegrated Vector Autoregressive Model with an Application to US Data," European University Institute Working Paper ECO No. 2001/2.

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

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

Karunaratne, N.D. (1996), "Exchange Rate Intervention in Australia (December 1983 to May 1993)," *Journal of Policy Modeling* 18, 397-417.

Krolzig, H.-M., and J. Toro (2002), "Testing for Super-exogeneity in the Presence of Common Deterministic Shifts," *Annales d'Economie et de Statistique* 67/68, 41–71.

Lee, T.-H., H. White and C.W.J. Granger (1993), "Testing for Neglected Nonlinearity in Time-Series Models: a Comparison of Neural Network Methods and Alternative Methods," *Journal of Econometrics* 56, 269-290.

Leeper, E., C. Sims, and T. Zha (1996), "What Does Monetary Policy Do?" *Brookings Papers on Economic Activity* 2, 1-78.

Li, W.K., S. Ling, and H. Wong (2001), Estimation for Partially Nonstationary Multivariate Autoregressive models with Conditional Heteroskedasticity," *Biometrika* 88, 1135-1152.

Lucas, R. (1976), "Econometric Policy Evaluation: A Critique," (with discussion) in K. Brunner and A.H. Melzer (eds.), *The Phillips Curve and Labor Markets.* Amsterdam: North Holland, pp. 19-46.

Monti, F. (2003), "Implementing Optimal Control in Cointegrated I(1) Structural VAR Models," European Central Bank Working Paper No. 288.

Nobay RA, Peel DA. (2003), "Optimal discretionary monetary policy in a model of asymmetric central bank preferences", Economic Journal, 113: 657–665.

Pagan, A. (1997), "Whatever Happened to Optimal Control of Econometric Models," *Control Engineering Practice* 5, 527-533.

Park, J. and P. Phillips (1988), "Statistical Inference in Regressions with Integrated Processes: Part I," *Econometric Theory* 4, 468-497.

Park, J. and P. Phillips (1989), "Statistical Inference in Regressions with Integrated Processes: Part II," *Econometric Theory* 4, 95-131.

Pesaran, H., D. Pettenuzzo, and A. Timmermann (2006), "Forecasting Time Series Subject to Multiple Structural Breaks," *Review of Economic Studies* 73, 1057-1084.

Ramsey, J.B. (1969) "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis", *Journal of the Royal Statistical Society* Series B 31, 350–371.

Ruge-Murcia FJ. (2002), "A prudent central banker", IMF Staff Papers 49: 456–469.

Ruge-Murcia FJ. (2004), "The inflation bias when the central banker targets the natural rate of unemployment", *European Economic Review*, 48: 91–107.

Sargent, T. (1976), "A Classical Macroeconometric Model for the United States," *Journal of Political Economy* 84, 207-237.

Sin, C.-Y. (2004), "Estimation and Testing for Partially Nonstationary Vector Autoregressive Models with GARCH: WLS versus QMLE," Econometric Society 2004 North American Summer Meetings 476, Econometric Society.

Stahl, D. and P. Wilson (1994), "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior and Organization* 25, 309–27.

Stock, J. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica* 55, 1035-1056.

Strotz, R. and H. Wold (1960), "Recursive vs Nonrecursive Systems: An Attempt at Synthesis," *Econometrica* 28, 417-427.

Taylor, J.B. (1993), "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214

Tillman, P. (2011), "Parameter uncertainty and nonlinear monetary rules", *Macroeconomic Dynamics*, (2011), 15: 184-200

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1-25.

White, H. (1990), "A Consistent Model Selection Procedure Based on m-Testing," in C.W.J. Granger (ed.), *Modelling Economic Series: Readings in Econometric Methodology.* Oxford: Oxford University Press, pp. 369-403.

White, H. (1994). *Estimation, Inference, and Specification Analysis.* New York: Cambridge University Press.

White, H. (2006a), "Time-Series Estimation of the Effects of Natural Experiments," *Journal of Econometrics* 135, 527–566.

White, H. (2006b), "Approximate Nonlinear Forecasting Methods," in G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting.* New York, Elsevier, pp. 460–512.

White, H. and K. Chalak (2009), "Settable Systems: An Extension of Pearl's Causal Model with Optimization, Equilibrium, and Learning," *Journal of Machine Learning Research* 10, 759-799.

White, H., and P. Kennedy (2009), "Retrospective Estimation of Causal Effects Through Time," in J. Castle and N. Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry.* Oxford: Oxford University Press, pp. 59–87.

White, H., and X. Lu (2010), "Granger Causality and Dynamic Structural Systems," *Journal of Financial Econometrics* 8, 193–243.

Wright, P.G. (1915), "Review of Moore, 'Economic Cycles'," *Quarterly Journal of Economics* 29, 631-641.

Mathematical Appendix

Proof of Proposition 2.1 Sufficiency of $B_{10} = 0$ is immediate.

For necessity, let $\tilde{A}_1 \equiv [\delta_{10}, A_{11}, A_{12}]$, $\tilde{A}_2 \equiv [\delta_{20}, A_{21}, A_{22}]$, $\tilde{B}_1 \equiv [b_1, B_{11}, B_{12}]$, $\tilde{B}_2 \equiv [b_2, B_{21}, B_{22}]$, and

$$B \equiv \left[\begin{array}{c} \tilde{B}_1 \\ \tilde{B}_2 \end{array} \right] \qquad \Delta \equiv \left[\begin{array}{c} \Delta_1 \\ \Delta_2 \end{array} \right].$$

Given (4), we have

$$\tilde{A}_1 + B_{10}\tilde{A}_2 = \tilde{B}_1.$$

Then $\tilde{A}_1 = \tilde{B}_1$ for all $\tilde{B}_1$ implies that $\tilde{A}_1 = \tilde{B}_1$ for all $\tilde{B}_1$ such that $\tilde{B}$ has full row rank.

For all such $\tilde{B}_1$, $\tilde{A}_1 + B_{10}\tilde{A}_2 = \tilde{B}_1$ implies $B_{10}\tilde{A}_2 = 0$, or equivalently that $B_{10}\Delta_2\tilde{B} = 0$. Because $\tilde{B}$ has full row rank, it follows that $B_{10}\Delta_2 = 0$. The existence of Δ ensures that Δ_2 has full row rank. It follows that $B_{10} = 0$. ∎

Proof of Proposition 3.1: This is an immediate corollary to theorem 4.3 of WL, with the assignments $Y_{1,t} \Leftrightarrow Y_t$, $\boldsymbol{Y}_{1,t-1} \Leftrightarrow (Y_{t-1}, Y_{t-2})$, $\boldsymbol{Y}_{2,t-1} \Leftrightarrow (Z_{t-1}, Z_{t-2})$, $\boldsymbol{X}_t \Leftrightarrow W_t$, and $U_{1,t} \Leftrightarrow \varepsilon_{1t}$. The assumption (9), i.e., $(Z_{t-1}, Z_{t-2}) \perp \varepsilon_{2t} \mid (Y_{t-1}, Y_{t-2}, W_t)$, is Assumption C.2 of WL. We also use the fact that for structures separable in ε_{1t}, such as (1), direct structural non-causality ($\boldsymbol{Y}_{2,t-1} \overset{d}{\not\Rightarrow}_{\mathcal{S}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)} Y_{1,t}$ in WL's notation) is equivalent to $A_{112} = 0$ and $A_{122} = 0$ ($\zeta_t(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t) = \tilde{\zeta}_t(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t)$). See WL, p.219. ∎

Proof of Proposition 3.2: This is an immediate consequence of corollary 6.2 of WL. We give the proof for (i). That for (ii) is similar. If $(Z_{t-1}, Z_{t-2}) \perp \varepsilon_{1t} \mid (Y_{t-1}, Y_{t-2}, W_t)$, then $(Z_{t-1}, Z_{t-2}) \perp (\varepsilon_{1t}, Y_{t-1}, Y_{t-2}) \mid (Y_{t-1}, Y_{t-2}, W_t)$ by lemmas 4.1 and 4.2(i) of Dawid (1979) (D79). If $A_{112} = 0$ and $A_{212} = 0$, then $Y_t = \delta_{10} + A_{111}Y_{t-1} + A_{121}Y_{t-2} + \varepsilon_{1t}$. That $(Z_{t-1}, Z_{t-2}) \perp Y_t \mid (Y_{t-1}, Y_{t-2}, W_t)$ now follows from D79 lemma 4.2(i). If $(Z_{t-1}, Z_{t-2}) \perp Y_t \mid (Y_{t-1}, Y_{t-2}, W_t)$, then $(Z_{t-1}, Z_{t-2}) \perp (Y_t, Y_{t-1}, Y_{t-2}) \mid (Y_{t-1}, Y_{t-2}, W_t)$ by lemmas 4.1 and 4.2(i) of D79. If $A_{112} = 0$ and $A_{122} = 0$, then $\varepsilon_{1t} = Y_t - (\delta_{10} + A_{111}Y_{t-1} + A_{121}Y_{t-2})$. That $(Z_{t-1}, Z_{t-2}) \perp \varepsilon_{1t} \mid (Y_{t-1}, Y_{t-2}, W_t)$ now follows from D79 lemma 4.2(i). ∎