

# Granger Causality, Exogeneity, Cointegration, and Economic Policy Analysis

Halbert White

Department of Economics  
University of California, San Diego

Davide Pettenuzzo

Department of Economics and International Business School  
Brandeis University

This version: June 2010

## Abstract

Policy analysis has 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 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.

**Acknowledgements:** The authors thank Peter Boswijk, Jim Hamilton, David Hendry, Peter Phillips, Sophocles Mavroeidis, and the participants of the 2010 Nottingham Sir Clive Granger Memorial Conference for helpful comments and suggestions. We thank Meng Huang of Bates White, LLC for excellent research assistance. Any errors are solely the authors's responsibility.

## 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.

Our starting point is 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. Policy effectiveness is ensured by another explicit causality requirement. 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 ( $\Delta Z_t$ ).

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. Our causal notions enable us to formulate

causal restrictions, essential for reliable policy analysis, that obviate the Lucas critique and that ensure policy effectiveness. Section 3 discusses a conditional form of exogeneity that serves to identify causal effects of interest and that forges links between structural causality and Granger causality. Section 4 reviews properties of cointegrated systems relevant here, with particular attention to their structural and causal 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 discusses econometric considerations that arise in empirically analyzing potentially cointegrated controlled systems and offers useful practical procedures and diagnostics. Section 8 illustrates these methods with an application to a simple model of the U.S. macroeconomy, and Section 9 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

### 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:

$$X_t \equiv \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \delta_0 + A_1 X_{t-1} + A_2 X_{t-2} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (1)$$

where  $Y_t$  represents observable "target" or "non-policy" variables<sup>1</sup> 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,  $N = N_1 + N_2$ . The vector  $\delta_0 \equiv (\delta'_{10}, \delta'_{20})'$  includes intercepts and any deterministic trend components. (See EHM, eq.(4).) We partition

---

<sup>1</sup>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 other aspect of  $Y_t$  or its distribution. When, for convenience, we refer simply to "targets" we always mean "target variables."

the nonrandom coefficient matrices  $A_1$  and  $A_2$  as

$$A_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} = \begin{bmatrix} A_{111} & A_{112} \\ A_{211} & A_{212} \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} = \begin{bmatrix} A_{121} & A_{122} \\ A_{221} & A_{222} \end{bmatrix}.$$

As econometricians, we do not know the  $A$ 's, nor do we observe the random "shocks"  $\varepsilon_t = (\varepsilon'_{1t}, \varepsilon'_{2t})'$ . Although  $\delta_0$ ,  $A_1$ , and  $A_2$  may depend deterministically on  $t$ , we leave this implicit to avoid further complicating the notation. We allow  $\delta_0$ ,  $A_1$ , 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  $x_{t-1} \rightarrow x_{t-1}^*$  and defined as the pair  $(x_{t-1}, x_{t-1}^*)$ . Then the *direct effect on*  $Y_t$  of the intervention  $x_{t-1} \rightarrow x_{t-1}^*$  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  $x_{jt-1} \rightarrow x_{jt-1} + 1$ . 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.

Using this notion of causality, we can say that if  $A_{112} = 0$ , then  $Z_{t-1}$  *does not structurally cause*<sup>2</sup>  $Y_t$ . Otherwise,  $Z_{t-1}$  *structurally causes*  $Y_t$ . If  $A_{112} = 0$  and  $A_{122} = 0$  then  $Z_{t-2}^{t-1} \equiv (Z_{t-2}, Z_{t-1})$  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 comments below, when we relate structural causality to Granger causality, using the framework of WL. Here, structural causality is the operating prerequisite.

## 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  $Z_t$ . To be sufficiently

---

<sup>2</sup>This causality is *direct* causality, but we will leave this implicit.

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

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

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  $v_{1t}$ . The meaning of  $\tilde{Z}_t$  is similar. These responses are determined in isolation, without permitting full equilibrium<sup>3</sup>, 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  $B_j \rightarrow B_j^*$  ( $j = 1$  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*,  $B_2 \rightarrow B_2^*$ , 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 fundamental requirement of mutual consistency, necessary for equilibrium. In equilibrium, the structure (2) satisfies

$$\begin{aligned}\tilde{Y}_t &= b_1 + B_{10}\tilde{Z}_t + B_{11}X_{t-1} + B_{12}X_{t-2} + v_{1t} \\ \tilde{Z}_t &= b_2 + B_{20}\tilde{Y}_t + B_{21}X_{t-1} + B_{22}X_{t-2} + v_{2t}, \quad t = 1, 2, \dots\end{aligned}\quad (3)$$

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 rela-

---

<sup>3</sup>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.

tions 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:

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

i.e., eq.(1), where

$$\begin{aligned} \delta_0 &= \Delta \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} & \Delta &\equiv \begin{bmatrix} I & -B_{10} \\ -B_{20} & I \end{bmatrix}^{-1} \\ A_1 &= \Delta \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} & A_2 &= \Delta \begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix}, \quad \text{and} \\ \varepsilon_t &= \Delta \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}. \end{aligned} \tag{4}$$

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 changes to the policymaker's subsystem of the DGP. A crucial requirement for traditional policy analysis is that the full equilibrium structural VAR for  $Y_t$  is invariant to the policy intervention. Otherwise, even without an explicit rational expectations framework, the Lucas critique (Lucas, 1976) operates with full force, with the implication that policies ignoring strategic behavior by the public are doomed to fail. In the EHM framework, superexogeneity ensures this invariance. EHM (section 3) also give compelling evidence that this invariance does hold in practice.

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  $B_2 \rightarrow B_2^*$  generally leads to a structural shift  $(\delta_{10}, A_{11}, A_{12}) \rightarrow (\delta_{10}^*, A_{11}^*, A_{12}^*)$  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,

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

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 sense formally given by the next result.

**Proposition 2.1** *Suppose  $B_{10}$  and  $B_{20}$  are such that  $\Delta$  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 to be ignorant of the policymaker's response function and shock. Instead, the public only has to know its own optimal response coefficients,  $B_1$ . Although  $v_{1t}$  may include components known only to the public, it may also contain an "implementation error" or "tremble" that the public has no control over or immediate knowledge of.

Recursivity is also behaviorally plausible. Indeed, experimental evidence in economics does not support the hypothesis that interacting agents arrive at fully rational Nash equilibria. Instead, the evidence supports a "level- $k$ " hierarchy of agents, who adopt strategies of varying sophistication (Stahl and Wilson, 1994; see also 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 supported by the fact that 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  $B_2$ .

Thus, there are both informational and behavioral factors supporting recursivity. 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 specifies structural causality as a property of the DGP. Particular causal properties of the DGP ensure necessary invariance and policy effectiveness conditions for policy

analysis. So far, 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 concepts, conditional exogeneity and Granger causality, that bear directly on policy analysis and that are also properties solely of the DGP.

### 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  $\delta_0$ ,  $A_1$ , and  $A_2$ .

WL give results implying that structural coefficients  $(\delta_0, A_1, A_2)$  can be identified when data are generated as in (1), provided that  $(X_{t-1}, X_{t-2})$  is independent of  $\varepsilon_t$  given covariates<sup>4</sup>  $W_t$ , 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, Cain, and Goldberger, 1980).

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

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

In this case,  $W_t$  has zero dimension. For example, (7) holds for the structure in (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 in this context.

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  $\varepsilon_t$ , including not only current and lagged values of variables that may also be driven by  $\varepsilon_t$  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.

---

<sup>4</sup>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  $Y_t$ .



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

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

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  $\delta_0$ ) 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  $\delta_0$ ,  $A_1$ ,  $A_2$ , 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,  $\delta_0$ ,  $A_1$ , 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  $\delta_0$ ,  $A_1$ , and  $A_2$ . For example, interest may attach just to  $A_{112}$  and  $A_{122}$  in

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

as  $A_{112}$  and  $A_{122}$  determine whether policy is effective or not. A milder exogeneity condition identifying just  $A_{112}$  and  $A_{122}$  is

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

Now we only require a correctly specified model for the sequence  $\{E(Y_t | X_{t-1}, X_{t-2}, W_t)\}$ ; 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  $Z^{t-1} \equiv (Z_{t-1}, Z_{t-2}, \dots)$  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  $X_t = (Y_t', Z_t)'$ ):

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

and suppose that  $\Delta X_t$  is  $I(0)$  and that there exist  $r < N$  cointegrating relations such that  $\beta' X_t$  is also  $I(0)$ , where  $\beta$  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  $\alpha$  with full column rank such that

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

See, e.g., Johansen (1988). Letting  $\Gamma \equiv -A_2$  then gives the standard error-correction cointegrating representation

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

where  $\Delta X_t \equiv X_t - X_{t-1}$  and  $\Psi_{t-1} \equiv \beta' X_{t-1}$ . This also has a structural interpretation, representing the causal relation holding between the response  $\Delta X_t$  and any admissible settings of RHS variables  $\Psi_{t-1}$ ,  $\Delta X_{t-1}$ , and  $\varepsilon_t$ . Thus, the matrices  $\alpha$  and  $\Gamma$  embody the effects on  $\Delta X_t$  of interventions to  $\Psi_{t-1}$  (long-run equilibrium departures) and  $\Delta X_{t-1}$ , respectively. On the other hand,  $\beta$  does not embody causal effects between elements of  $X_t$ .

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  $\beta$  in a first stage by least squares (Stock, 1987), forming an estimate  $\hat{\Psi}_{t-1}$  of  $\Psi_{t-1}$ , and then regressing  $\Delta X_t$  on an intercept and  $\hat{\Psi}_{t-1}$ ,  $\Delta X_{t-1}$ ,  $W_t$  to obtain standard estimators of  $\delta_0$ ,  $\alpha$ ,  $\Gamma$ , 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  $\Delta X_{t-1}$ .

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

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

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

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

It is plausible that this conditional heteroskedasticity can be exploited to yield a relatively efficient GLS-like estimator, based on a suitable specification for  $E(\eta_t \eta_t' | W_t)$ . 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  $\varepsilon_t$ , we should generally expect it to predict  $\eta_t \eta_t'$  as well, affording the opportunity for possible efficiency gains.

For completeness, we record the normal quasi-maximum likelihood estimator (QMLE) as the solution to the problem

$$\begin{aligned} \max_{\theta \in \Theta} L_T(\theta) &\equiv T^{-1} \sum_{t=1}^T -0.5 \ln \det(\Sigma(W_t; \theta_2)) \\ &\quad -0.5 (X_t - \mu(X_{t-1}, X_{t-2}, W_t; \theta_1))' [\Sigma(W_t; \theta_2)]^{-1} (X_t - \mu(X_{t-1}, X_{t-2}, W_t; \theta_1)), \end{aligned}$$

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

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, and the present framework draws the needed distinction, based on work of WL, who show that  $G$ -causality and structural causality are equivalent, provided a suitable conditional form of exogeneity holds. In this sense,  $G$ -causality is a derivative requirement that may be useful for testing the structural causality of policy instruments, which is 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

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

Central to the EHM approach is the reparametrization of (11) in terms of conditional ( $\Delta Y_t | \Delta Z_t$ ) and marginal ( $\Delta Z_t$ ) distributions. Here this yields

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

$$\Delta Z_t = \delta_{02} + \Gamma_{21} \Delta Y_{t-1} + \Gamma_{22} \Delta Z_{t-1} + \varepsilon_{2t}, \quad (13)$$

with

$$\zeta_t \equiv \varepsilon_{1t} - \theta_1 \varepsilon_{2t} \sim N(0, \Omega),$$

where  $\Omega = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$  and

$$\begin{aligned} \theta_1 &= \Sigma_{12} \Sigma_{22}^{-1} & \theta_0 &= \delta_{01} - \theta_1 \delta_{02} \\ \theta_2 &= \Gamma_{11} - \theta_1 \Gamma_{21} & \theta_3 &= \Gamma_{21} - \theta_1 \Gamma_{22}. \end{aligned} \quad (14)$$

Without dynamic stability restrictions, the parameters of the conditional and marginal distributions,  $\lambda_1 = (\theta_0, \theta_1, \theta_2, \theta_3, \Omega)$  and  $\lambda_2 = (\delta_{02}, \Gamma_{21}, \Gamma_{22}, \Sigma_{22})$ , define a sequential cut between the conditional model (12) and the marginal model (13). The cross restrictions stemming from (14) impose no specific restrictions on the elements of either  $\lambda_1$  or  $\lambda_2$ . 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  $\lambda_1$  and  $\lambda_2$  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. (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 throughout.

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 in-



struments 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,

$$\begin{aligned}\tilde{Y}_t &= b_1 + B_{11}X_{t-1} + B_{12}X_{t-2} + v_{1t}, \\ \tilde{Z}_t &= b_2 + B_{20}Y_t + B_{21}X_{t-1} + B_{22}X_{t-2} + v_{2t} \quad t = 1, 2, \dots\end{aligned}$$

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

$$\begin{aligned}Y_t &= b_1 + B_{11}X_{t-1} + B_{12}X_{t-2} + v_{1t}, \\ Z_t &= b_2 + B_{20}Y_t + B_{21}X_{t-1} + B_{22}X_{t-2} + v_{2t} \quad t = 1, 2, \dots\end{aligned} \tag{15}$$

We explicitly rule out instantaneous causation by requiring that  $v_{1t}$  is 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 first consider a system of the particular form

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t}, \\ \Delta Z_t &= b_2 + B_{20} \Delta Y_t \quad t = 1, 2, \dots\end{aligned} \tag{16}$$

For simplicity, we assume here that the shocks  $\{v_{1t}\}$  are IID with mean zero. Also just for simplicity, we take  $v_{2t} = 0$ , so the policymaker is able to precisely implement their policy. Thus,  $v_{1t}$  is independent of  $(X_{t-1}, X_{t-2})$  and estimation is standard (e.g., Ahn and Reinsel, 1990).

Where convenient, we take  $N_1 \leq N_2$ ; this helps ensure that there are enough policy variables to control all the targets. In the appendix we explicitly treat the case where  $N_2 < N_1$ .

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

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

When  $\beta_1 = 0$ , we have an integrated open-loop system for  $Y_t$  without cointegration. Otherwise, with  $0 < r \equiv \text{rk}(\beta_1) < N_1$ , 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  $\Delta Z_t$  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 just note that under suitable conditions, one feasible optimal policy for targeting  $\gamma_o$  has the form above,

$$\Delta Z_t = b_2 + B_{20} \Delta Y_t,$$

where  $b_2$  and  $B_{20}$  are properly chosen. The famous Taylor rule (Taylor, 1993) is of this form.

In the engineering literature, this is known as proportional (P) control of  $\Delta Y_t$ . More sophisticated control methods are also common, especially the class of proportional-integral-derivative (PID) methods, which can also be expressed as a constrained version of eq.(15). We first consider P control, as this keeps the analysis relatively simple and delivers useful insights; later in this section, we consider a form of proportional-integral (PI) control. In discussing our empirical example, we discuss a form of PID control. Throughout, our analysis is fairly elementary. We refer the interested reader to JJ or Monti (2003) for deeper analysis.

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

$$\Delta Y_t = (b_1 + B_{12} b_2) + \alpha_1 \beta_1' Y_{t-1} + (B_{11} + B_{12} B_{20}) \Delta Y_{t-1} + v_{1t}.$$

As long as  $\Delta Z_{t-1}$  structurally causes  $\Delta Y_t$  ( $B_{12} \neq 0$ ), the dynamics of the open-loop and closed-loop systems can be very different. Because of partial equilibrium recursivity, the open-loop

deep parameters  $(b_1, B_{11}, B_{12})$  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  $(b_2, B_{20})$  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. We develop our analysis by starting with simple cases and considering progressively more complex possibilities.

## 6.1 P Control without Open-Loop Cointegration

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

$$\Delta Y_t = (b_1 + B_{12}b_2) + (B_{11} + B_{12}B_{20})\Delta Y_{t-1} + v_{1t}. \quad (17)$$

Closed-loop target system stability is determined by the roots of the characteristic equation

$$\det[Iz - (B_{11} + B_{12}B_{20})] = 0,$$

where  $z$  is a complex number. The closed-loop target system is stable if these roots lie inside the unit circle. To attain  $E(\Delta Y_t) = \gamma_o$ , we require that  $\Delta Y_t$  is  $I(0)$ , so we assume that the policymaker chooses  $B_{20}$  such that the roots of the characteristic equation lie inside the unit circle. Interestingly, the open-loop target system can have roots on or outside the unit circle; proper choice of  $B_{20}$  can therefore change the order of integration between open- and closed-loop systems. On the other hand,

**Proposition 6.1** *With P control, when the open-loop target system is integrated but not cointegrated, no choice of  $(b_2, B_{20})$  can initiate cointegration in the closed-loop target system.*

For brevity in this section, we leave implicit the other conditions imposed above.

Further, with P control,  $\Delta Z_t$  is an affine function of an  $I(0)$  process, so  $\Delta Z_t$  is  $I(0)$  and  $Z_t$  is generally  $I(1)$ . When  $Z_t$  is  $I(1)$ , we can ask whether  $Z_t$  can be cointegrated. For this, we seek non-zero  $\theta$  such that  $\theta'Z_t$  is  $I(0)$ . We have

$$\theta'Z_t = \theta'Z_{t-1} + \theta'b_2 + \theta'B_{20}\Delta Y_t.$$

This is  $I(0)$  if and only if  $\theta'b_2 + \theta'B_{20}\Delta Y_t = 0$ , in which case  $\theta'Z_t$  is constant. Thus,

**Proposition 6.2** *With P control, when the open-loop target system is integrated but not cointegrated, the control system exhibits cointegration if and only if there exists non-zero  $\theta$  in the null space of  $B'_{20}$  such that  $\theta'b_2 = 0$ .*

We thus view cointegration within the P control system as possible but unusual. To support this, we now examine the choice of  $(b_2, B_{20})$  in more detail. As noted above, we choose  $(b_2, B_{20})$  to achieve the policy objective,  $E(Y_t) = \gamma_o$ . To find a solution to this policy problem, we assume dynamic equilibrium at  $\gamma_o$  and take expectations on both sides of (17) to obtain

$$\gamma_o = (b_1 + B_{12}b_2) + (B_{11} + B_{12}B_{20})\gamma_o.$$

Collecting terms and solving for  $b_2$  gives

$$b_2 = B_{12}^{-1}\{(I - [B_{11} + B_{12}B_{20}])\gamma_o - b_1\}, \quad (18)$$

where  $B_{12}^{-1}$  denotes the right inverse,  $B_{12}^{-1} \equiv B'_{12}(B_{12}B'_{12})^{-1}$ , provided  $B_{12}$  has full row rank ("full structural causality" of  $\Delta Z_{t-1}$ ). If  $B_{12}$  has less than full rank, then either  $\gamma_o$  cannot be attained or there are multiple solutions. Here, we take  $B_{12}$  to be of full rank, ensuring a unique solution. The appendix treats the case where  $B_{12}^{-1}$  need not exist.

From (18), we see that with this long-run objective the policymaker has a fair degree of latitude, as both  $b_2$  and  $B_{20}$  are subject to policymaker choice. Eq.(18) gives the required value for  $b_2$  given any choice for  $B_{20}$ , even a choice with deficient rank or with  $B_{20} = 0$ , the no feedback rule. In the latter case, however, we require  $b_2 \neq 0$  for control to be present, in which case cointegration in the control system cannot hold.

We also note that the closer the roots of the characteristic equation are to the unit circle, the slower is convergence to dynamic equilibrium; the closer to zero, the quicker. As can be seen from eq.(17), the policy goal can be attained as quickly as possible by choosing, if feasible,  $B_{20}$  to fully offset the open-loop dynamics, such that

$$B_{20} = -B_{12}^{-1} B_{11}.$$

Another consequence of (18) is that when  $\gamma_o$  is fixed, a structural shift in the open-loop target process generally results in a policy intervention  $(b_2, B_{20}) \rightarrow (b_2^*, B_{20}^*)$ . Although the closed-loop process may then undergo a further structural shift, the long-run behavior remains unchanged as long as the policy goal  $\gamma_o$  remains constant. We can thus distinguish several distinct kinds of policy interventions: changes to the policy goal  $\gamma_o$  (a "policy regime change," resulting, for example, from a change in political regime); changes to the proportional gain  $B_{20}$ , resulting from "tuning exercises" undertaken by the policymaker to change the speed of convergence to equilibrium, with consequent adjustment to  $b_2$ ; and changes to  $b_2$  and  $B_{20}$  due to policymaker responses to structural shifts in the open-loop target system.

So far, we have seen that with P control and in the absence of open-loop target system cointegration, there can be no cointegration within the closed-loop target system and that cointegration within the control system is possible but very special. As  $Y_t$  and  $Z_t$  are generally both

$I(1)$ , it also makes sense to ask whether there is or can be cointegration between the closed-loop target and control systems.

For this, we seek  $\theta = (\theta'_1, \theta'_2)'$  such that  $\theta'_1 Y_t + \theta'_2 Z_t$  is  $I(0)$ . P control implies

$$\begin{aligned}\theta'_1 Y_t + \theta'_2 Z_t &= \theta'_1 Y_t + \theta'_2 (Z_{t-1} + b_2 + B_{20} \Delta Y_t) \\ &= (\theta'_1 Y_{t-1} + \theta'_2 Z_{t-1}) + (\theta'_1 + \theta'_2 B_{20}) \Delta Y_t + \theta'_2 b_2.\end{aligned}$$

As  $\Delta Y_t$  is  $I(0)$ , it is easily verified that cointegration holds (with  $\theta'_1 Y_t + \theta'_2 Z_t = \text{const}$ ) if and only if  $\theta'_2 b_2 = 0$ , with  $\theta'_1 = -\theta'_2 B_{20}$ .

**Proposition 6.3** *With P control, when the open-loop target system is integrated but not cointegrated, there is cointegration between the control and the controlled target if and only if there exists non-zero  $\theta_2$  such that*

$$\theta'_2 B_{12}^{-1} \{ (I - [B_{11} + B_{12} B_{20}]) \gamma_o - b_1 \} = 0.$$

This places very stringent conditions on the relation between  $\gamma_o$  and  $B_{20}$ .

Thus, even though there is no cointegration in the open-loop target system, P control can induce cointegration between target and control. But this is a quite special circumstance that need not be consistent with policymaker behavior, as the policymaker may choose a speed of convergence to the long-run policy goal  $\gamma_o$  incompatible with this condition.

We also see that policy interventions may create or destroy cointegration between control and target. Specifically, interventions to  $\gamma_o$  but not  $B_{20}$  (policy regime change) or to  $(b_2, B_{20})$  but not  $\gamma_o$  (tuning exercises) will destroy cointegration if it exists and may (but are not likely to) initiate cointegration if it does not. Interventions to  $(b_2, B_{20})$  resulting from structural shifts in the target subsystem may (but are not likely to) create or can easily destroy cointegration between target and control.

## 6.2 P and PI Control with Open-Loop Cointegration

### 6.2.1 P Control

Now consider a system that implements P control of a target exhibiting cointegration in the open-loop target dynamics:

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta'_1 Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t}, \\ \Delta Z_t &= b_2 + B_{20} \Delta Y_t \qquad \qquad \qquad t = 1, 2, \dots\end{aligned}\tag{19}$$

We again suppose the policymaker seeks to attain  $E(\Delta Y_t) = \gamma_o$ . Our discussion parallels that for the pure integrated case, but now we must take proper account of the cointegrating terms.

The closed-loop representation for the target system is now

$$\Delta Y_t = (b_1 + B_{12}b_2) + \alpha_1\beta_1'Y_{t-1} + (B_{11} + B_{12}B_{20})\Delta Y_{t-1} + v_{1t}. \quad (20)$$

From this, we see that

**Proposition 6.4** *With P control, when the open-loop target system is cointegrated, no choice of  $(b_2, B_{20})$  can remove cointegration from the closed-loop target system.*

Analysis for the control system identical to that for the pure integrated case gives

**Proposition 6.5** *With P control, when the open-loop target system is cointegrated, the control system exhibits cointegration if and only if there exists non-zero  $\theta$  in the null space of  $B_{20}'$  such that  $\theta'b_2 = 0$ .*

To see how policy can achieve a long-run policy goal, take expectations on both sides of (20) at the steady state, which gives

$$\gamma_o = (b_1 + B_{12}b_2) + \alpha_1\mu_1 + (B_{11} + B_{12}B_{20})\gamma_o, \quad (21)$$

where  $\mu_1 \equiv \beta_1'E(Y_{t-1})$ . To determine  $\mu_1$ , observe that

$$\beta_1'E(Y_t) = \beta_1'E(Y_{t-1}) + \beta_1'E(\Delta Y_t),$$

so that in the steady state

$$\beta_1'\gamma_o = 0.$$

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  $\Delta Y_t$  using a smaller set of policy control variables. We discuss this case in the appendix.

To solve for  $\mu_1$ , multiply both sides of (21) by  $\beta_1'$ . This gives

$$\begin{aligned} \beta_1'(b_1 + B_{12}b_2) + \beta_1'\alpha_1\mu_1 + \beta_1'(B_{11} + B_{12}B_{20})\gamma_o &= 0 \quad \text{so that} \\ \mu_1 &= -(\beta_1'\alpha_1)^{-1}\beta_1'[(B_{11} + B_{12}B_{20})\gamma_o + (b_1 + B_{12}b_2)], \end{aligned}$$

where we use the fact that  $\alpha_1$  and  $\beta_1$  have full column rank, ensuring that  $\beta_1'\alpha_1$  is nonsingular. Substituting this into (21) gives

$$\gamma_o = (b_1 + B_{12}b_2) - \alpha_1(\beta_1'\alpha_1)^{-1}\beta_1'[(B_{11} + B_{12}B_{20})\gamma_o + (b_1 + B_{12}b_2)] + (B_{11} + B_{12}B_{20})\gamma_o.$$

Collecting terms, we get

$$\{I - J_1(B_{11} + B_{12}B_{20})\}\gamma_o = J_1(b_1 + B_{12}b_2), \quad (22)$$

where  $J_1 \equiv I - \alpha_1(\beta'_1\alpha_1)^{-1}\beta'_1$  is the "long-run impact matrix" (see JJ, eq.(4)). Observe that  $J_1(B_{11} + B_{12}B_{20})$  plays the role of a "first order autocorrelation" in this system. This determines the speed of convergence to the policy steady state, provided the roots of the associated characteristic equation are inside the unit circle. With P control, the policymaker can influence this speed of convergence by choice of  $B_{20}$ .

Generally,  $J_1$  is singular, so when we express  $b_2$  in terms of  $B_{20}$ , we obtain

$$J_1B_{12}b_2 = [I - J_1(B_{11} + B_{12}B_{20})]\gamma_o - J_1b_1.$$

Compare this to (18) above: the pure integrated case results when  $J_1 = I$ . 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  $\gamma_o$ . One way to proceed in such cases is to minimize a convex function of  $b_2$  (for example,  $b'_2b_2$ ) subject to (22).

We emphasize that only choices for  $\gamma_o$  satisfying  $\beta'_1\gamma_o = 0$  give valid choices for  $B_{20}$  and  $b_2$ .

As for the pure integrated case, we ask whether cointegration can hold between the target and control subsystems. Now we seek  $\theta = (\theta'_1, \theta'_2)'$  such that  $\theta'_1Y_t + \theta'_2Z_t$  is  $I(0)$ . The analysis is essentially identical to the pure integrated case, as both involve P control. We obtain

$$\begin{aligned} \theta'_1Y_t + \theta'_2Z_t &= \theta'_1Y_t + \theta'_2(Z_{t-1} + b_2 + B_{20}\Delta Y_t) \\ &= (\theta'_1Y_{t-1} + \theta'_2Z_{t-1}) + (\theta'_1 + \theta'_2B_{20})(Y_t - Y_{t-1}) + \theta'_2b_2. \end{aligned}$$

As in the previous case, we obtain cointegration if and only if  $\theta'_2b_2 = 0$ , with  $\theta'_1 = -\theta'_2B_{20}$ . Thus,

**Proposition 6.6** *With P control, when the open-loop target system is cointegrated, there is cointegration between the control and the controlled target if and only if there exists non-zero  $\theta_2$  such that*

$$\begin{bmatrix} -J_1B_{12} & [I - J_1(B_{11} + B_{12}B_{20})] \\ \theta'_2 & 0 \\ 0 & \beta'_1 \end{bmatrix} \begin{pmatrix} b_2 \\ \gamma_o \end{pmatrix} = \begin{pmatrix} J_1b_1 \\ 0 \\ 0 \end{pmatrix}. \quad (23)$$

In contrast to the pure integrated case, there may be no such solution  $\gamma_o$ . Thus, this possibility is even more exceptional than the pure integrated case. Although P control can sometimes induce cointegration between target and control, this is a very special circumstance, depending on exact policymaker choices that may be incompatible with policymaker objectives.

### 6.2.2 PI Control

A potential drawback of P control is that it cannot achieve policies violating  $\beta_1' \gamma_o = 0$ . If the policymaker's desired  $\gamma_o$  is not compatible with this restriction, it would be useful to alter  $\beta_1$ , say to  $\tilde{\beta}_1$ , for which  $\tilde{\beta}_1' \gamma_o = 0$  does hold. This cannot be done with P control. Accordingly, consider a system with a cointegrated open-loop target as above, but with the control system augmented by the inclusion of a term depending on  $Y_t$ :

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t}, \\ \Delta Z_t &= b_2 + B_{20} \Delta Y_t + B_{12}^{-1} \alpha_1 \lambda_1' Y_t\end{aligned}\quad t = 1, 2, \dots,$$

where  $\lambda_1$  is  $N_1 \times r$ . This control system has a form known as proportional-integral (PI) control in the engineering literature. The presence of  $Y_t$  implements the "integral" aspect of the control. We denote this particular implementation  $PI_1$ .

The closed-loop target behavior for this system is given by

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} (b_2 + B_{20} \Delta Y_{t-1} + B_{12}^{-1} \alpha_1 \lambda_1' Y_{t-1}) + v_{1t} \\ &= b_1 + B_{12} b_2 + (\alpha_1 \beta_1' + \alpha_1 \lambda_1') Y_{t-1} + (B_{11} + B_{12} B_{20}) \Delta Y_{t-1} + v_{1t} \\ &= b_1 + B_{12} b_2 + \alpha_1 (\beta_1 + \lambda_1)' Y_{t-1} + (B_{11} + B_{12} B_{20}) \Delta Y_{t-1} + v_{1t}.\end{aligned}$$

We see that this system is cointegrated, but now with cointegration parameters  $(\alpha_1, \tilde{\beta}_1)$ , with  $\tilde{\beta}_1 \equiv \beta_1 + \lambda_1$ .  $PI_1$  control permits the policymaker to modify the cointegration-feasible policies. By suitable choice of  $\lambda_1$ , the policymaker can reduce the cointegrating rank or even remove cointegration entirely from the closed-loop target system (set  $\lambda_1 = -\beta_1$ ). We have

**Proposition 6.7** *When the open-loop target system is cointegrated,  $PI_1$  control can modify or remove the closed-loop target system cointegrating vector,  $\beta_1 + \lambda_1$ .*

We also see that with  $PI_1$  control, since  $Y_t$  is  $I(1)$ ,  $\Delta Z_t = b_2 + B_{20} \Delta Y_t + B_{12}^{-1} \alpha_1 \lambda_1' Y_t$  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. Although the target and control are generally of different orders, we can see immediately from the control equations that  $\Delta Z_t - B_{12}^{-1} \alpha_1 \lambda_1' Y_t$  is  $I(0)$ , so  $\Delta Z_t$  and  $Y_t$  are cointegrated. Interestingly, this cointegration is generated by the causal control relation between  $\Delta Z_t$  and  $Y_t$ , rather than simply reflecting aspects of system dynamics.

With  $PI_1$ , the policymaker can attain  $\gamma_o$  for which  $\tilde{\beta}_1' \gamma_o = 0$ , choosing  $b_2$  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  $\tilde{J}_1 \equiv I - \alpha_1(\tilde{\beta}'_1\alpha_1)^{-1}\tilde{\beta}'_1$ .

PI<sub>1</sub> control is limited in the sense that it cannot initiate cointegration in the closed loop target system if cointegration is absent in the open-loop target system. Also, PI<sub>1</sub> control does not allow the policymaker to alter the effect  $\alpha_1$  of equilibrium departures on the target variable. Nevertheless, a straightforward elaboration of PI<sub>1</sub> does permit these possibilities. Specifically, consider PI<sub>2</sub> control, where, with  $\kappa_1$  an  $N_1 \times r$  matrix,

$$\Delta Z_t = b_2 + B_{20}\Delta Y_t + B_{12}^{-1}[\alpha_1\lambda'_1 + \kappa_1(\beta_1 + \lambda_1)']Y_t \quad t = 1, 2, \dots$$

A little algebra shows that the closed-loop target system has the form

$$\Delta Y_t = b_1 + B_{12}b_2 + \tilde{\alpha}'_1\tilde{\beta}'_1Y_{t-1} + (B_{11} + B_{12}B_{20})\Delta Y_{t-1} + v_{1t}.$$

Here, the cointegration vector is  $\tilde{\beta}_1 \equiv \beta_1 + \lambda_1$  and the effect of equilibrium departures is  $\tilde{\alpha}_1 \equiv \alpha_1 + \kappa_1$ . Not only does PI<sub>2</sub> control permit the policymaker to modify the cointegration-feasible policies, it permits the policymaker to adjust the response to equilibrium departures or even introduce cointegration into the closed-loop target system, despite its absence in the open-loop target system. We have

**Proposition 6.8** *When the open-loop target system is integrated or cointegrated, PI<sub>2</sub> control can modify the closed-loop target system cointegrating coefficients  $(\alpha_1 + \kappa_1)$ ,  $(\beta_1 + \lambda_1)$ .*

As for PI<sub>1</sub> control,  $Z_t$  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, as for PI<sub>1</sub>, 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  $\Delta Z_t$  and  $Y_t$  are again causally cointegrated.

Parallel to PI<sub>1</sub>, with PI<sub>2</sub> the policymaker can attain  $\gamma_o$  for which  $\tilde{\beta}'_1\gamma_o = 0$ , where  $\tilde{\beta}_1 = \beta_1 + \lambda_1$ . Again 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  $\tilde{J}_1 \equiv I - \tilde{\alpha}_1(\tilde{\beta}'_1\tilde{\alpha}_1)^{-1}\tilde{\beta}'_1$ .

Another case of PI control fixes  $\lambda_1$  at zero, with  $\kappa_1 \neq 0$ . We call this PI<sub>3</sub> control. This is the special case of PI<sub>2</sub> where the policymaker only modifies the speed of error correction. Its properties can be inferred from our discussion of PI<sub>2</sub>, so we do not discuss this further here.

## 7 Estimating and Testing Controlled Systems

We now consider some econometric issues that arise in estimating and testing controlled systems. Our discussion here is mainly pragmatic, in the spirit Granger (2009) compellingly advocated.

Thus, we will not be concerned so much with regularity conditions or asymptotic properties, but rather with describing practical methods that can offer useful insight while preventing us from going too far astray with any particular line of investigation.

## 7.1 A Controlled System for Estimation and Testing

For concreteness, we consider a DGP accommodating PI control of the target first differences,

$$\Delta Y_t = b_1 + D_1 Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t}, \quad (24)$$

$$\Delta Z_t = b_2 + B_{20} \Delta Y_t + D_2 Y_t + v_{2t} \quad t = 1, 2, \dots \quad (25)$$

When  $D_1 = \alpha_1 \beta_1'$  and  $D_2 = B_{12}^{-1}[\tilde{\alpha}_1 \tilde{\beta}_1' - D_1]$ , we have the system with open-loop cointegration and  $\text{PI}_2$  control considered in the last section. This system is a special case of eq.(1) with possible cointegration in the open-loop target subsystem, and, consistent with this form of PI control, also possibly between  $Y_t$  and  $\Delta Z_t$ . Observe that we now permit policy implementation to be noisy, as an implementation error  $v_{2t}$  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  $B_{12}v_{2,t-1}$  into the closed-loop target system.

In what follows, it is often convenient to take  $N_2 \geq N_1$ , but we do not always require this. The appendix provides further discussion of the  $N_2 < N_1$  case.

The associated closed-loop target system is

$$\Delta Y_t = c_1 + \tilde{\alpha}_1 \tilde{\beta}_1' Y_{t-1} + C_{11} \Delta Y_{t-1} + \eta_{1t}, \quad (26)$$

where  $c_1 \equiv b_1 + B_{12}b_2$ ,  $\tilde{\alpha}_1 \tilde{\beta}_1' = D_1 + B_{12}D_2$ ,  $C_{11} \equiv B_{11} + B_{12}B_{20}$ , and  $\eta_{1t} \equiv v_{1t} + B_{12}v_{2,t-1}$ . Observe that even if  $D_1$  has full rank, the closed-loop system can be cointegrated. That is,  $\text{PI}_2$  control can induce a cointegrated closed-loop target system, even when the open-loop system is not cointegrated. For what follows, we assume that  $D_1 + B_{12}D_2$  has non-zero rank  $\tilde{r} < N_1$ , so that the closed-loop system is in fact cointegrated.

An important aspect of these structures is that they are subject to structural shifts. The target system (24) is subject to "exogenous" shifts, that is, shifts arising outside the controlled system. The policy system (25) 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 (24) represent a failure of invariance, in which case the Lucas critique operates. For now, we assume recursivity, ruling out endogenous shifts. Below, we discuss testing this.

Because the closed-loop target system contains coefficients from both the target and control structures, the coefficients of (26) can shift for any of the reasons just given. It is also in principle possible for the policy authority to undertake policy interventions that precisely offset exogenous

structural changes to (24), leaving (26) unchanged. Nevertheless, this requires a sufficient degree of knowledge and flexibility that exogenous structural shifts in (24) are likely in practice to be reflected in structural shifts in (26).

It is of course possible that the target subsystem experiences no exogenous shifts. This is the simplest and most favorable case for estimation and inference, so we begin with this. Correspondingly, we suppose for now that there are no policy regime changes or tuning exercises in the control system. It follows that the closed-loop target system coefficients also do not shift.

We first focus on whether policy can be effective; that is, we wish to know whether  $B_{12} = 0$  in (24), or, when relevant, whether  $B_{12}$  has full row rank. For simplicity, we take  $\{v'_{1t}, v'_{2t}\}$  to be IID, with  $(Y_1, \Delta Y_0, \Delta Z_0)$  independent of  $(v'_{11}, v'_{21})$ . This ensures strict exogeneity and makes estimation relatively straightforward.

Although we maintain the assumption of structural stability, to avoid going astray at the outset, it is helpful to begin by estimating the closed-loop system (26) and examining its stability. As (26) is a standard cointegrated system, one can apply standard methods, such as the Engle-Granger (1987) estimator or the methods of Johansen (1995). One can test for stability using 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). If one finds evidence of a stable sample or subsample, one can proceed by analyzing that data, as this evidence is consistent with stability of both the target and control systems. For now, we suppose we have such a subsample. Estimating  $c_1$ ,  $C_{11}$ ,  $\tilde{\alpha}_1$ , and  $\tilde{\beta}_1$  in (26) provides complete information about the closed-loop target dynamics under the governing policy regime in the absence of exogenous structural change in (24) and even when policy is ineffective.

Examining (24) and (25) with a view to estimating (24) and noting that (26) ensures that  $Y_t$  is  $I(1)$ , we see that there are two main possibilities for  $\Delta Z_t$ . The first is that  $\Delta Z_t$  is  $I(0)$ ; the second is that  $\Delta Z_t$  is  $I(1)$ . The first possibility arises in either of two cases. It is easily checked that  $\Delta Z_t$  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<sub>3</sub> control ( $\tilde{\beta}_1 = \beta_1$ , but  $\tilde{\alpha}_1 \neq \alpha_1$ ). On the other hand, regardless of open loop cointegration, we have  $\Delta Z_t \sim I(1)$  with either PI<sub>1</sub> or PI<sub>2</sub> control. Thus, a test of the null hypothesis that  $\Delta Z_t$  is  $I(1)$  vs. the  $I(0)$  alternative is a test for PI<sub>1</sub> or PI<sub>2</sub> control vs. P or PI<sub>3</sub> control, assuming correct specification of (24) and (25).

Different considerations arise in estimating (24), depending on whether  $\Delta Z_t$  is  $I(1)$  or  $I(0)$ . We take these up below. In either case, however, one must check whether  $\Delta Z_{t-1}$  is perfectly collinear with  $Y_{t-1}, \Delta Y_{t-1}$ . The possibility of this collinearity underlies claims that Granger causality testing (here, *structural* causality testing) is useless for policy analysis (Sargent, 1976; Buiter, 1984). As Granger (1988) pointed out, however, as long as there is any noise in the policy

rule, this objection falls. We saw in Section 6 that noiseless control is effective for achieving policy goals; here we see that this can hinder policy analysis by making it impossible to identify  $B_{12}$ . Thus, noisy control is helpful for learning about  $B_{12}$ , and, in this regard, the noisier the better, as near multi-collinearity is almost as much a prohibitive obstacle as is perfect collinearity.

To assess potential difficulties in identifying  $B_{12}$ , that is, to see how great a problem collinearity may be, one should investigate the relation between  $\Delta Z_t$  and  $(Y_t, \Delta Y_t)$  before estimating (24). An immediate simple diagnostic is the  $R^2$  from the regression of  $\Delta Z_t$  on a constant,  $Y_t$  and  $\Delta Y_t$ , i.e., (25). We would ideally 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  $B_{12}$  is not zero. We return to this regression below, but for now we suppose that we are in the typical situation in which the fit is reasonably good but not perfect.

## 7.2 The $\Delta Z_t \sim I(0)$ Case

Suppose now that  $\Delta Z_t$  is  $I(0)$ . Then we can estimate (24) by standard methods for cointegrated systems. In fact,  $\Delta Z_t \sim I(0)$  implies  $\tilde{\beta}_1 = \beta_1$ , so we can estimate (24) by least-squares regression of  $\Delta Y_t$  on a constant,  $\tilde{\Psi}_{t-1} \equiv \tilde{\beta}_1' Y_{t-1} = \Psi_{t-1}$ ,  $\Delta Y_{t-1}$ , and  $\Delta Z_{t-1}$ , where  $\tilde{\Psi}_{t-1}$  is estimated using the estimate of  $\tilde{\beta}_1$  obtained from (26). When  $\Delta Z_t$  is  $I(0)$ , the estimator of  $B_{12}$ , say  $\hat{B}_{12}$ , has standard properties, so we can immediately check policy effectiveness by testing  $B_{12} = 0$ . One can also use convenient methods recently given by Camba-Méndez and Kapetanios (2008) to estimate and test the rank of  $B_{12}$ .

If we do not impose the constraint that  $\tilde{\beta}_1 = \beta_1$  in estimating (24), we expect that the estimates obtained from (24) and (26) will be similar. A specification test can be performed by formally testing the hypothesis  $\tilde{\beta}_1 = \beta_1$ . If we reject, this indicates that either or both (24) or (26) are misspecified. This test can be conveniently performed by estimating a version of (24) 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  $\tilde{\beta}_1 = \beta_1$ .

The estimates of  $\alpha_1$  and  $\tilde{\alpha}_1$  may or may not differ, depending on whether the system is subject to P or  $PI_3$  control. By testing the hypothesis  $\tilde{\alpha}_1 = \alpha_1$ , we can test the null that P control is in effect, against the alternative of  $PI_3$  control. An easier test of the P control null can be accomplished using (25). With either P or  $PI_3$  control, this relation can be estimated by regressing  $\Delta Z_t$  on a constant,  $\Delta Y_t$ , and  $\tilde{\Psi}_t$ . Under P control, the coefficients on  $\tilde{\Psi}_t$  are zero; under  $PI_3$  control they are not. Thus, one can test P control vs.  $PI_3$  control by applying

standard tests for zero values of these coefficients.

Usually, we expect that the estimates of  $b_1$  and  $c_1$  will differ and that the estimates of  $B_{11}$  and  $C_{11}$  will differ. When  $N_1 = N_2$ , we can use estimates of  $b_1$ ,  $B_{11}$ ,  $B_{12}$ ,  $c_1$ , and  $C_{11}$  to estimate  $b_2$  and  $B_{20}$ , as

$$b_2 = -B_{12}^{-1}(c_1 - b_1) \quad B_{20} = -B_{12}^{-1}(C_{11} - B_{11}).$$

Generally, however, it is simplest to estimate  $b_2$  and  $B_{20}$  using a modified version of (25),

$$\Delta Z_t = b_2 + B_{20}\Delta Y_t + D_2^* \tilde{\Psi}_t + v_{2t} \quad t = 1, 2, \dots,$$

taking advantage of the P or PI<sub>3</sub> structure.

Even if  $\Delta Z_t$  and  $(Y_t, \Delta Y_t)$  are perfectly collinear, identification of  $B_{12}$  is not a lost cause, as policy interventions can also identify  $B_{12}$ . This can also improve the precision of estimation even in the absence of perfect collinearity. For this, we require a subsample in which (24) is stable and in which (25) exhibits one or more shifts. For simplicity, suppose there is a single policy intervention  $(b_2, B_{20}) \rightarrow (b_2^*, B_{20}^*)$ , and define  $c_1^* = b_1 + B_{12}b_2^*$  and  $C_{11}^* = B_{11} + B_{12}B_{20}^*$ . Then

$$\begin{bmatrix} b_1 & B_{11} & B_{12} \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & I & 0 & I \\ b_2 & B_{20} & b_2^* & B_{20}^* \end{bmatrix} = \begin{bmatrix} c_1 & C_{11} & c_1^* & C_{11}^* \end{bmatrix}.$$

Provided that

$$\mathcal{B} \equiv \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & I & 0 & I \\ b_2 & B_{20} & b_2^* & B_{20}^* \end{bmatrix}$$

has full row rank, it follows that

$$\begin{bmatrix} b_1 & B_{11} & B_{12} \end{bmatrix} = \begin{bmatrix} c_1 & C_{11} & c_1^* & C_{11}^* \end{bmatrix} \mathcal{B}'(\mathcal{B}\mathcal{B}')^{-1}.$$

We can identify and consistently estimate  $\begin{bmatrix} c_1 & C_{11} & c_1^* & C_{11}^* \end{bmatrix}$  from (26) and  $\mathcal{B}$  from (25) (if  $\mathcal{B}$  is unknown), permitting us to identify and consistently estimate  $\begin{bmatrix} b_1 & B_{11} & B_{12} \end{bmatrix}$ . Conducting inference is then straightforward, although computationally somewhat involved.

The methods just discussed suffice to identify and estimate all structural coefficients for the P or PI<sub>3</sub> cases, so we now turn our attention to the case in which  $\Delta Z_t \sim I(1)$ .

### 7.3 The $\Delta Z_t \sim I(1)$ Case

The  $\Delta Z_t \sim I(1)$  case arises with PI<sub>1</sub> or PI<sub>2</sub> control. From (25), we see that the  $I(1)$  variable  $Y_t$  appears on the RHS multiplied by  $D_2$ , which, in the case of PI<sub>1</sub> or PI<sub>2</sub> control, does not give an  $I(0)$  product  $D_2 Y_t$ . Since  $b_2 + B_{20}\Delta Y_{t-1} + v_{2t}$  is  $I(0)$ ,  $\Delta Z_t \sim I(1)$  follows.

Inspecting the target system

$$\Delta Y_t = b_1 + D_1 Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t},$$

we see that the LHS  $\Delta Y_t$  is  $I(0)$  (as ensured by (26)), but that two  $I(1)$  variables appear on the RHS,  $Y_{t-1}$  and  $\Delta Z_{t-1}$ . In the absence of other structure, this would prevent estimating  $D_1$  and  $B_{12}$ . But here we have that  $Y_{t-1}$  and  $\Delta Z_{t-1}$  are cointegrated by (25) and in just the right way to permit consistent estimation of  $(b_1, D_1, B_{11}, B_{12})$  by least squares. Note that  $D_1$  may or may not have full rank, but even when  $D_1 = \alpha_1 \beta_1'$ , standard estimation methods for cointegrated systems will not apply, as the relevant cointegrating vector here is  $\tilde{\beta}_1$ , not  $\beta_1$ .

Two possibilities for conducting inference about  $B_{12}$  suggest themselves. The first is to directly apply the results of Park and Phillips (1988, 1989). The resulting inference for  $B_{12}$  may be non-standard, however. An apparently simpler possibility is to use a modified version of (26) to estimate  $B_{12}$ , namely

$$\Delta Y_t = c_1 + \tilde{\alpha}_1 \tilde{\beta}_1' Y_{t-1} + C_{11} \Delta Y_{t-1} + B_{12} v_{2,t-1} + v_{1t}, \quad (27)$$

where we replace  $v_{2,t-1}$  with an estimate from (25), say  $\hat{v}_{2,t-1}$ . 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 that estimating the policy equation (25) involves a regression of an  $I(1)$  variable ( $\Delta Z_t$ ) 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 (25) with a variety of standard procedures, plausibly with standard  $\sqrt{T}$  asymptotics.

So far, the methods described for the  $\Delta Z_t \sim I(1)$  case cover identification and consistent estimation of all system parameters except  $\alpha_1$  and  $\beta_1$  when  $D_1 = \alpha_1 \beta_1'$ . But the singular value decomposition applies to represent  $D_1$  as

$$\begin{aligned} D_1 &= \mathcal{U} \Lambda \mathcal{V}' = \mathcal{U} \Lambda^{1/2} \Lambda^{1/2} \mathcal{V}' \\ &= \mathcal{U} \Lambda^{1/2} S S' \Lambda^{1/2} \mathcal{V}' \\ &\equiv \alpha_1 \beta_1', \end{aligned}$$

where  $\mathcal{U}$  and  $\mathcal{V}$  are  $N_1 \times N_1$  orthogonal matrices,  $\Lambda$  is an  $N_1 \times N_1$  diagonal matrix with  $r = \text{rk}(D_1)$  non-zero eigenvalues on the diagonal,  $S$  is the  $N_1 \times r$  selection matrix,  $S' \equiv [I_r, 0]$ , and  $\alpha_1 \equiv \mathcal{U} \Lambda^{1/2} S$ ,  $\beta_1' \equiv \mathcal{V} \Lambda^{1/2} S$ . When  $r < N_1$ , this decomposition permits us to identify and estimate the open-loop cointegrating parameters.

Because the details appear somewhat involved, we leave a formal analysis of inference for  $B_{12}$  based on the results of Park and Phillips (1988, 1989) to future research.

## 7.4 Model Diagnostics

Regardless of whether  $\Delta Z_t$  is  $I(0)$  or  $I(1)$ , there is a variety of model diagnostics that both the policymaker and the econometrician should examine. For the  $I(0)$  case, these can be used to conduct formal specification tests, as inference is standard; accordingly we omit the details. In the discussion to follow, any references to tests apply to just the  $I(0)$  case. Nevertheless, these diagnostics can be also be computed for the  $I(1)$  case and may be informative. They also plausibly can form the basis for formal specification tests, but this will require proper development of the relevant asymptotic distributions.

A particularly important diagnostic task is to examine the exogeneity of  $(Y_{t-1}, \Delta Y_{t-1}, \Delta Z_{t-1})$ . If this fails, the estimator of  $B_{12}$  is not informative about the effect of  $\Delta Z_{t-1}$  on  $Y_t$ . Here, exogeneity is ensured by the assumption that  $\{v'_{1t}, v'_{2t}\}$  is an independent sequence. A simple diagnostic for this can be based on estimates of  $\eta_{1t} \equiv v_{1t} + B_{12}v_{2,t-1}$  from (26), say  $\hat{\eta}_{1t}$ . One can form the analog of Durbin's (1970)  $h$ -test from the regression of  $\hat{\eta}_{1t}$  on  $\hat{\eta}_{1,t-1}$ . The estimated matrix  $\hat{\rho}$  of first-order autocorrelations can then be used to check or test whether the true autocorrelations are zero. This test will not have power against all possible alternatives, so one should also test autocorrelation using estimates of  $v_{1t}$  and  $v_{2t}$  from (24) and (25).

It is important to examine whether linear models based on (24) are correctly specified, that is, whether the DGP for  $\Delta Y_t$  is indeed the assumed linear structure analyzed here. If not, the estimator of  $B_{12}$  is not fully informative about the effect of  $\Delta Z_{t-1}$  on  $\Delta Y_t$ , and the policy rules discussed here may not be adequate for the desired control. 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 (24), (25), or (26). In our illustrative application in Section 8, we give details of a neural network-based method for testing neglected nonlinearity.

The policymaker and the econometrician must also check whether recursivity holds. If not, the needed invariance is absent, and traditional policy control is unworkable. For this, one can apply methods of Engle and Hendry (1993), who describe testing invariance without imposing weak exogeneity. See also Hendry (1988) and Hendry and Santos (2009). Hendry and Massman (2006) survey and extend the concept of co-breaking, directly relevant here. A particularly straightforward procedure for testing invariance, related to methods of Hendry and Mizon (1998)

and Krolzig and Toro (2002), can be implemented using data spanning at least one policy intervention  $(b_2, B_{20}) \rightarrow (b_2^*, B_{20}^*)$ , where one is willing to maintain the assumption of target structure stability. The idea is to estimate a version of (24) 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 (25), using, for example, methods of Bai and Perron (1998). This augmented structure has the form

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

Under the null of recursivity,  $\Pi_1 = 0$ ; evidence of departures from zero is evidence against recursivity. When  $\Delta Z_t$  is  $I(0)$ , one can apply standard methods to test  $\Pi_1 = 0$ .

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 (24) and  $d_{1t} = 0$  otherwise; and  $d_{2t} = 1$  if there is a structural shift in in period  $t$  in the policy system (25) 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.

As we further discuss below, it is important to rule out neglected nonlinearity when testing for recursivity, as neglected nonlinearity can manifest as one or more structural shifts.

## 7.5 Policy Implementation and Operation

Once the policy authority has determined that its policy instruments are indeed effective (recursivity holds and  $B_{12}$  is non-zero or, better, of full rank) and has gained reliable knowledge of  $b_1$ ,  $D_1$ ,  $B_{11}$ , and  $B_{12}$ , it can determine whether its desired policy goal  $\gamma_o$  is feasible for some set of policy parameters  $b_2$ ,  $B_{20}$ , and  $D_2$ . If there are multiple feasible implementations, the policymaker can select a preferred implementation and begin policy operations.

Once policy operations begin, a main activity for the policymaker, besides manipulating the policy instruments, is to monitor the target system to detect exogenous shifts that will require policy interventions to keep the system on track. This raises some important practical issues that have not been addressed here or, to our knowledge, elsewhere in the cointegration literature. The first of these is how the policymaker detects shifts. In our earlier discussion, we implicitly assumed the policymaker could detect these immediately. But this is unrealistic. More



realistically, the policymaker could apply statistical techniques for real-time structural change monitoring in either (24) or (26). Standard monitoring methods (e.g., Chu, Stinchcombe, and White, 1996; Hornik, Leisch, Kleiber, and Zeileis, 2005) do not necessarily allow detecting changes in a cointegrated structure; this is an interesting topic for future research. The second is the even more difficult question of how the policymaker learns the new structure, once a shift has been detected. In practice, this also will take some time; how much will depend on the nature of the shift. Meanwhile, the show must go on; policy must continue.

This suggests that a realistic framework for policy analysis is one with adaptive learning by the policy authority in a context that permits cointegration and exogenous structural breaks. Such a framework could possibly be based on the recursive learning framework of Chen and White (1998), although this would need modification to accommodate recursive learning of cointegrated structures. Handling exogenous breaks could be accommodated by keeping the learning rate constant or bounded away from zero, rather than declining to zero with the sample size, as Chen and White (1998) require. The form of policy rules emerging from such a framework could well be of the form (25), but with all the policy parameters adjusting through time, based on convenient recursive estimation strategies. Another possibility is that the policymaker behaves according to a recursive Bayesian procedure, such as that proposed by Pesaran, Pettenuzzo, and Timmermann (2006). As analyzing such frameworks is beyond the scope of the present analysis, we leave this for future research.

## 7.6 Questions for the Econometrician

Questions of interest to the econometrician but not the policymaker involve extracting information known only to the policymaker. For example, the econometrician should be interested in whether the system is in fact under PI or P control, or whether some other rule operates. We saw above that given correct specification, the econometrician can draw inferences about P or PI control from the integration properties of  $\Delta Z_t$ . We have also seen that the econometrician can identify and consistently estimate  $(\tilde{\alpha}_1, \tilde{\beta}_1)$ ,  $(b_1, B_{11}, B_{12}, \alpha_1, \beta_1)$ , and  $(b_2, B_{20}, D_2)$ .

The econometrician may also want to know  $\gamma_o$ . For this, (26) may provide sufficient information, as  $\gamma_o$  must satisfy

$$\mathcal{C} \gamma_o \equiv \begin{bmatrix} \tilde{\beta}'_1 \\ [I - \tilde{J}_1 C_{11}] \end{bmatrix} \gamma_o = \begin{bmatrix} 0 \\ \tilde{J}_1 c_1 \end{bmatrix}.$$

Provided that the  $(\tilde{r} + N_1) \times N_1$  matrix  $\mathcal{C}$  has full column rank, we can solve for  $\gamma_o$  as

$$\gamma_o = (\mathcal{C}'\mathcal{C})^{-1}\mathcal{C}' \begin{bmatrix} 0 \\ \tilde{J}_1 c_1 \end{bmatrix}.$$

To determine whether P or PI control correctly describes policymaker behavior, the econometrician can use the modified version of (25),

$$\Delta Z_t - D_2 Y_t = b_2 + B_{20} \Delta Y_t + v_{2t} \quad t = 1, 2, \dots,$$

where, for estimation, we replace unknown LHS coefficients with their (first-stage) estimates. One can then conduct specification tests on this regression. If misspecification is found, this indicates that the policy rule is not P or PI control of  $\Delta Y_t$  by  $\Delta Z_t$  or that the policy goal differs from long run targeting of  $E(\Delta Y_t) = \gamma_o$ , or both. Conducting these specification tests will require properly accounting for the joint distribution of the first-stage estimators of  $D_2$  and the second-stage estimators of  $b_2, B_{20}$ . It is plausible that the asymptotic distribution for the second-stage estimator will be affected by the distribution of the first-stage estimators of  $D_2$ . When  $\Delta Z_t$  is  $I(0)$ , adjusting for this is straightforward; the distribution of the estimator of  $\tilde{\beta}_1$  ( $= \beta_1$ ) will typically not play a role, due to its superconsistency.

Once one has estimated the system coefficients, one can investigate system impulse responses by perturbing  $v_{2,t-1}$  in (27) and simulating. One can also investigate the effects of policy interventions by studying the effects of changes to the elements of  $(b_2, B_{20}, D_2)$ .

## 8 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), as well as the references given there. Here, we apply the framework described above to illustrate how one could 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 arrive at a model sufficiently well specified to deliver definitive insights about Fed policy. Thus, we will pay attention to indicators of model shortcomings without necessarily resolving the issues identified. As will become apparent, resolving the issues uncovered will 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 natural logarithm of 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 begin by examining whether the closed loop target system (26) is stable. For this, we use the recursive log-likelihood test described in Juselius (2006). While there are several different tests that could be used, 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,  $T_1, \dots, T$ . The statistic for subsamples including observations  $t_1, \dots, T$ , with  $t_1 = T_1, T_1 - 1, \dots, 1$  is

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

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  $r = 1$  cointegrating relation. The estimated variance-covariance matrices for the sub-sample including observations  $t_1, \dots, T$  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.

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.

Next, 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

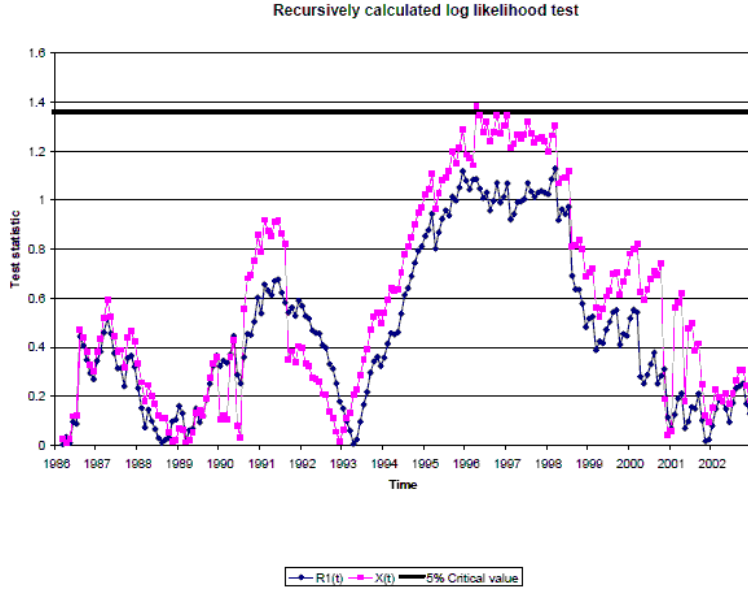


Figure 1: Recursive log likelihood function. The baseline sample is 2002:12 - 2007:12

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 now investigate whether  $\Delta Z_t$  is collinear with  $\Delta Y_t$  and  $Y_t$ . The  $R^2$  of this regression is 0.151, so we conclude that collinearity of the policy instrument with  $\Delta Y_t$  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. Below, we investigate this further. By running the  $\Delta Z_t$  regression with  $\tilde{\Psi}_t$  in place of  $Y_t$ , we can easily test P control vs.  $PI_3$  control. The coefficient on  $\tilde{\Psi}_t$  is strongly significant ( $p < .001$ ), so we reject the P control hypothesis.

Following the process described in section 7.2, we next test policy effectiveness by testing whether  $B_{12}$  is significantly different from zero. We proceed in two different ways. First, we estimate the open-loop target system (24) using the method of Johansen (1995). Table 2 reports the relevant results based on White (1980) robust standard errors. As the table shows, the estimates of  $B_{12}$  are significantly different from zero for both the IPI and the unemployment variables, and are borderline significant for the oil price variable. 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.

Target variable	Coeff	Std Err	t-stat	p-value
<b>IPI total</b>	0.025	0.006	4.279	0.000
<b>Inflation</b>	0.215	0.363	0.591	0.555
<b>Unemployment</b>	-0.147	0.029	-5.035	0.000
<b>Cushing oil price</b>	0.186	0.107	1.733	0.084

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.

Target variable	Coeff	Std Err	t-stat	p-value
<b>IPI total</b>	0.024	0.006	4.275	0.000
<b>Inflation</b>	0.214	0.363	0.590	0.556
<b>Unemployment</b>	-0.147	0.029	-5.034	0.000
<b>Cushing oil price</b>	0.186	0.107	1.737	0.084

Table 2: Open loop estimates for the matrix  $B_{12}$ . Results are obtained by OLS regression of  $\Delta Y_t$  on a constant,  $\tilde{\Psi}_{t-1} \equiv \tilde{\beta}'_1 Y_{t-1} = \Psi_{t-1}$ ,  $\Delta Y_{t-1}$ , and  $\Delta Z_{t-1}$ , where  $\tilde{\Psi}_{t-1}$  is estimated using the estimate of  $\tilde{\beta}_1$  obtained from (26). The computed standard errors are White (1980) robust standard errors.

Alternatively, we estimate  $B_{12}$  by OLS regression of  $\Delta Y_t$  on a constant,  $\tilde{\Psi}_{t-1} \equiv \tilde{\beta}'_1 Y_{t-1} = \Psi_{t-1}$ ,  $\Delta Y_{t-1}$ , and  $\Delta Z_{t-1}$ , where  $\tilde{\Psi}_{t-1}$  is estimated using the estimate of  $\tilde{\beta}_1$  obtained from the closed loop equation (26). As is immediately apparent from the tables, the results from this second estimation strategy are substantially the same as before.

Next, we test  $\tilde{\beta}_1 = \beta_1$ . As described above, we can test this hypothesis by applying the Johansen (1995) procedure to reestimate the target system (24) with the RHS variables augmented to include  $\tilde{\Psi}_{t-1} \equiv \tilde{\beta}'_1 Y_{t-1}$ . We test whether the cointegrating rank in this augmented regression is zero using Johansen's (1995) trace statistic test. We cannot reject the zero-rank hypothesis at the 10% level, so there is no evidence of misspecification on this basis.

We also examine the estimated residual first-order autocorrelations of the closed-loop system (26) and the open-loop target system (24), as a diagnostic for exogeneity of  $(Y_{t-1}, \Delta Y_{t-1}, \Delta Z_{t-1})$ . Overall, these do not indicate any obvious problems. For the closed-loop system these coefficients are  $-0.0152$  (IPI),  $0.0659$  (inflation),  $-0.0304$  (unemployment), and  $-0.0062$  (oil). For the open-loop target system, these are  $-0.0018$  (IPI),  $0.0667$  (inflation),  $-0.0546$  (unemployment), and  $-0.0044$  (oil).

On the other hand, the estimated autocorrelation coefficient for the control equation is .340, suggesting some form of misspecification, dynamic or otherwise. To see whether a simple autocorrelation adjustment can resolve matters, we apply the Cochrane-Orcutt technique to (25). The result is an estimated autocorrelation coefficient of .527 and very different coefficient estimates. As there is no lagged dependent variable in the control equation, this outcome

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 instead of PI<sub>3</sub> control, the policymaker is executing PI<sub>3</sub>-derivative (PI<sub>3</sub>D) control, in which case a previously omitted term  $\Delta^2 Y_t$  appears in the control equations<sup>5</sup>, so that (25) is modified to

$$\Delta Z_t = b_2 + B_{20}\Delta Y_t + B_{21}\Delta^2 Y_t + D_2 Y_t + v_{2t} \quad t = 1, 2, \dots \quad (28)$$

When we estimate the PI<sub>3</sub>D control equation, we find that the  $R^2$  increases from .151 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 .289. This is a move in the right direction, but clearly PI<sub>3</sub>D control is not the whole story. Applying Cochrane-Orcutt to (28), we find an estimated autocorrelation coefficient of .380. 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<sub>3</sub> control. In particular, we reject the PD control hypothesis in favor of PI<sub>3</sub>D. In line with our illustrative intent, we proceed by assuming PI<sub>3</sub>D control with autocorrelated errors. We further investigate this equation below, however.

PI<sub>3</sub>D control also modifies the closed-loop system (26) 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 (24), (28), or the closed-loop PID system are linear or whether there may be neglected nonlinearity. These tests are essentially those described in WL, section 5. 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 equation (24), we construct a Wald statistic for each equation  $h = \text{IPI, inflation,}$

---

<sup>5</sup>As can be readily verified, the addition of D control has no impact on the cointegration results of the previous section.

unemployment, oil price, to test the joint hypothesis  $\zeta_{11h} = \zeta_{12h} = \dots = \zeta_{1kh} = 0$  in:

$$\begin{aligned} \Delta Y_{th} &= b_{1h} + D_{1h}^* \beta_1' Y_{t-1} + B_{11h} \Delta Y_{t-1} + B_{12h} \Delta Z_{t-1} \\ &+ \sum_{j=1}^k G(\gamma_{10hj} + \gamma_{11hj} \beta_1' Y_{t-1} + \gamma_{12hj} \Delta Y_{t-1} + \gamma_{13hj} \Delta Z_{t-1}) \zeta_{1jh} + \varepsilon_{1th} \end{aligned} \quad (29)$$

- for equation (28), we construct a Wald statistic to test the joint hypothesis  $\zeta_{21} = \zeta_{22} = \dots = \zeta_{2k} = 0$  in:

$$\begin{aligned} \Delta Z_t &= b_2 + B_{20} \Delta Y_t + B_{21} \Delta^2 Y_t + D_2^* \tilde{\beta}_1' Y_t + \rho \hat{\varepsilon}_{2,t-1} \\ &+ \sum_{j=1}^k G(\gamma_{20j} + \gamma_{21j} \Delta Y_t + \gamma_{22j} \Delta^2 Y_t + \gamma_{23j} \tilde{\beta}_1' Y_t + \gamma_{24j} \hat{\varepsilon}_{2,t-1}) \zeta_{2j} + \varepsilon_{2t} \end{aligned} \quad (30)$$

- for the closed-loop PID system, we construct a Wald statistic for each equation  $h = \text{IPI, inflation, unemployment, oil price}$ , to test the joint hypothesis  $\zeta_{31h} = \zeta_{32h} = \dots = \zeta_{3kh} = 0$  in:

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

In these regressions,  $\beta_1$  and  $\tilde{\beta}_1$  are replaced by their estimates. Estimated lagged errors  $\hat{\varepsilon}_{2,t-1}$  are included in (30) to accommodate the autocorrelation of the control equation error terms.

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,  $G(x) = (-x^5 + 10x^3 - 15x) \exp(-x^2/2)$ . 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  $k$  indicates the number of included hidden units and controls the allowed degree of nonlinearity. We choose  $\gamma$ 's from a set of candidates, constructed as in Huang and White (2009). The algorithm to select the  $\gamma$ 's follows the QuickNet procedure described in White (2006b).

Tables 3-5 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 stark message from these tests is that the target and control system equations are *all* misspecified, with the apparent exception of the unemployment target equation. Not surprisingly, then, the closed loop PID equations are also found to be misspecified. Clearly, simple

$k$	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Row BH</b>
$\Delta$ IPI $_t$	0.004823	0.000764	0.002111	0.003855	0.002843	0.003821
$\Delta$ Inflation $_t$	0.231894	0.262284	0.099672	0.091767	0.010387	0.051933
$\Delta$ Unemployment $_t$	0.395854	0.514657	0.222053	0.235633	0.15057	0.514657
$\Delta$ Oil price $_t$	0.084189	0.053384	0.059099	0.039732	0.012985	0.064924
<b>Col BH</b>	0.019293	0.003057	0.008445	0.015419	0.011372	0.015284

Table 3: Misspecification tests for the target subsystem

$k$	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Row BH</b>
$\Delta Z_t$	0.111215	0.004651	0.001902	0.001546	0.000225	0.001126

Table 4: Misspecification tests for the PID control equation

linear structures do not adequately capture important features of the data. Thus, research can be productively directed toward examining the adequacy and implications of more flexible specifications, such as the neural network specifications forming the basis for these tests. We leave this as a topic for future work.

Even though the Juselius (2006) tests accord with structural stability for our system with PI<sub>3</sub>D 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<sub>3</sub>D 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.

To distinguish structural shifts from neglected nonlinearities, one promising approach is to estimate neural networks of the form specified above using indicator saturation methods of HJS. If only neglected nonlinearity were the issue, no structural shifts would be found. If only structural shifts were the issue, then the  $\zeta$  coefficients would be essentially zero. If there were a mix of shifts and nonlinearities, this procedure would, in principle, permit their separate identification and estimation. We leave investigation of such procedures and their application

$k$	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>Row BH</b>
$\Delta$ IPI $_t$	0.061583	0.006365	5.99E-05	1.42E-05	7.86E-06	3.93E-05
$\Delta$ Inflation $_t$	0.009056	0.000503	0.000211	0.000478	3.94E-05	0.000197
$\Delta$ Unemployment $_t$	0.005372	0.006648	0.001038	0.000738	0.000939	0.003113
$\Delta$ Oil price $_t$	0.04323	0.046952	0.057542	0.030044	0.008714	0.043572
<b>Col BH</b>	0.021489	0.002011	0.000239	5.69E-05	3.14E-05	0.000157

Table 5: Misspecification tests for the closed-loop PID system



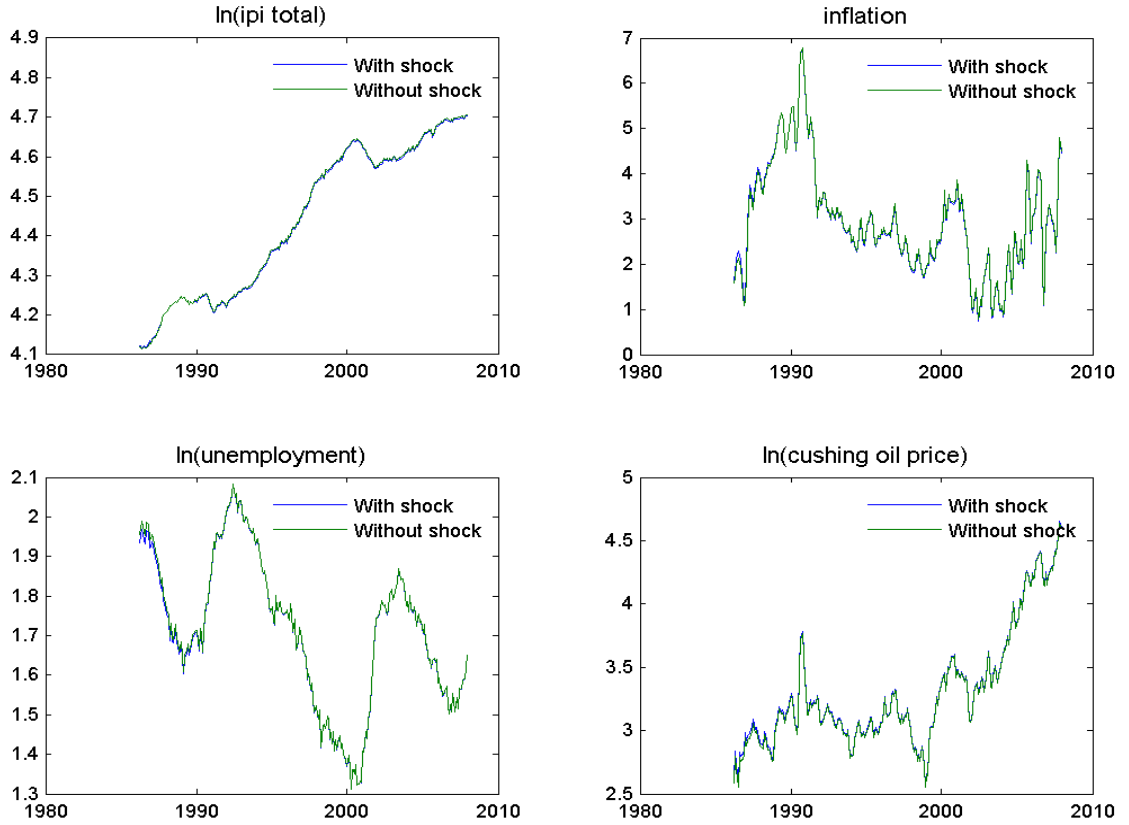


Figure 2: Simulation results for the closed-loop PID system

to future research.

We close this section with a brief examination of the impacts of Fed policy implied by taking the linear  $PI_3D$  system estimated here at face value. Specifically, we conduct the experiment of positively perturbing  $v_{2,t-1}$  in the PID analog of the closed-loop system (27) by three standard deviations for a single period. 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 ( $B_{12}$ ) of the Fed Funds rate is statistically significant, Figure 2 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.

## 9 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 ( $\Delta Z_t$ ). 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<sub>3</sub>, or PI<sub>3</sub>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 pursuing and 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.

## 10 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

$$B \equiv \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} \quad \Delta \equiv \begin{bmatrix} \Delta_1 \\ \Delta_2 \end{bmatrix}.$$

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  $\Delta$  ensures that  $\Delta_2$  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  $Y_{1,t} \Leftrightarrow Y_t$ ,  $\mathbf{Y}_{1,t-1} \Leftrightarrow (Y_{t-1}, Y_{t-2})$ ,  $\mathbf{Y}_{2,t-1} \Leftrightarrow (Z_{t-1}, Z_{t-2})$ ,  $\mathbf{X}_t \Leftrightarrow W_t$ , 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  $\varepsilon_{1t}$ , such as (1), direct structural non-causality ( $\mathbf{Y}_{2,t-1} \not\stackrel{d}{\perp}_{S(\mathbf{Y}_{1,t-1}, \mathbf{X}_t)} Y_{1,t}$  in WL's notation) is equivalent to  $A_{112} = 0$  and  $A_{122} = 0$  ( $\zeta_t(\mathbf{Y}_{t-1}, \mathbf{Z}_t) = \tilde{\zeta}_t(\mathbf{Y}_{1,t-1}, \mathbf{Z}_t)$ ). 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  $Y_t = \delta_{01} + A_{111}Y_{t-1} + A_{121}Y_{t-2} + \varepsilon_{1t}$ . 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  $\varepsilon_{1t} = Y_t - (\delta_{01} + A_{111}Y_{t-1} + A_{121}Y_{t-2})$ . 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). ■

### Attaining a composite policy objective with $N_1 > N_2$

Consider PI control of a system with open-loop cointegration when  $N_1 > N_2$ , i.e., there are more targets than instruments:

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12} \Delta Z_{t-1} + v_{1t} \\ \Delta Z_t &= b_2 + B_{20} \Delta Y_t + D_2 Y_t \quad t = 1, 2, \dots,\end{aligned}$$

where  $D_2$  is  $N_2 \times N_1$ . The closed-loop system is given by

$$\begin{aligned}\Delta Y_t &= b_1 + \alpha_1 \beta_1' Y_{t-1} + B_{11} \Delta Y_{t-1} + B_{12}(b_2 + B_{20} \Delta Y_{t-1} + D_2 Y_{t-1}) + v_{1t} \\ &= b_1 + B_{12} b_2 + (\alpha_1 \beta_1' + B_{12} D_2) Y_{t-1} + (B_{11} + B_{12} B_{20}) \Delta Y_{t-1} + v_{1t} \\ &= b_1 + B_{12} b_2 + \tilde{\alpha}_1 \tilde{\beta}_1' Y_{t-1} + (B_{11} + B_{12} B_{20}) \Delta Y_{t-1} + v_{1t},\end{aligned}$$

where

$$\begin{aligned}\tilde{\alpha}_1 \tilde{\beta}_1' &= \alpha_1 \beta_1' + B_{12} D_2 = \mathcal{U} \Lambda \mathcal{V}' \\ &= \mathcal{U} \Lambda^{1/2} S S' \Lambda^{1/2} \mathcal{V}',\end{aligned}$$

with  $\tilde{\alpha}_1 \equiv \mathcal{U} \Lambda^{1/2} S$  and  $\tilde{\beta}_1 \equiv \mathcal{V} \Lambda^{1/2} S'$ . Here, we apply the singular value decomposition as in the main text.

This system is cointegrated, provided  $\alpha_1 \beta_1' + B_{12} D_2$  has rank less than  $N_1$ . This is possible but not guaranteed, as  $B_{12} D_2$  has at most rank  $N_2 < N_1$ . The policymaker thus has some latitude to modify the cointegrating properties of the open-loop target system, but because  $N_2 < N_1$  there is less freedom than in the  $N_2 \geq N_1$  case. We proceed under the assumption that the policymaker can choose  $D_2$  to attain a cointegrated closed-loop system with cointegrating rank  $\tilde{r}$ . P control obtains as the special case where  $D_2 = 0$ .

The derivations of Section 6.2 apply directly to give  $\tilde{\beta}_1' \gamma_o = 0$  and

$$K \gamma_o = k_o, \quad (32)$$

where

$$K \equiv I - \tilde{J}_1 (B_{11} + B_{12} B_{20}) \quad \text{and} \quad k_o \equiv \tilde{J}_1 (b_1 + B_{12} b_2),$$

with  $\tilde{J}_1 \equiv I - \tilde{\alpha}_1(\tilde{\beta}'_1\tilde{\alpha}_1)^{-1}\tilde{\beta}'_1$ , as these derivations do not depend on the relative dimensions of  $N_1$  and  $N_2$ .

Now suppose that the policymaker seeks to attain a composite goal of the form

$$H E(\Delta Y_t) = H \gamma_o = h_o,$$

where  $H$  ( $N_2 \times N_1$ ) and  $h_o$  ( $N_2 \times 1$ ) are chosen by the policymaker. This objective may be feasible, as the goal has as many elements as there are policy instruments.

This approach represents the policymaker's willingness to trade off different components of its target goals. For example, it might be willing to accept higher unemployment to attain lower inflation or vice-versa, recognizing that the available policy instruments do not permit achieving both lower inflation and lower unemployment.

To see how this goal could be achieved, it is helpful to consider the rows  $H_i$  of  $H$  one at a time. First, note that whenever  $H_i$  belongs to  $\tilde{\mathbb{B}}_1 = \text{span}(\tilde{\beta}_1)$  (the set of all linear combinations of  $\tilde{\beta}_1$ , i.e.,  $\tilde{\beta}_1\xi$ , where  $\xi$  is  $\tilde{r} \times 1$ ), then the only feasible value for the corresponding  $h_{oi}$  is zero. If  $H_i\gamma_o = 0$  is indeed a goal for the policymaker, further control (beyond attaining  $\tilde{\beta}_1$ ) is unnecessary, as the system will always tend to this value. The policymaker can influence the speed of convergence to the goal by manipulating  $B_{20}$ , but here this is a secondary consideration.

Instead, the policymaker can focus attention on achieving policy goals represented by choices  $H_i$  belonging to  $\tilde{\mathbb{B}}_1^\perp$ , the subspace of  $\mathbb{R}^{N_1}$  containing vectors with a component orthogonal to  $\tilde{\beta}_1$ . With  $N_2$  instruments, the policymaker can specify  $N_2 \leq N_1 - \tilde{r}$  such choices. To be feasible, these must be consistent with (32), so  $H_i$  and  $h_{oi}$  must satisfy

$$H_i = \theta'_i K \quad \text{and} \quad h_{oi} = \theta'_i k_o$$

for some  $N_1 \times 1$  vector  $\theta_i$ . For such a  $\theta_i$  to exist, it is necessary and sufficient that

$$\text{rk}(K', H'_i) = \text{rk}(K'),$$

(e.g., Hadley, 1961), which is straightforward to check. For simplicity, suppose that for a given  $B_{20}$  there are  $N_2$  linearly independent such  $\theta_i$ 's. Stacking the rows  $H_i$  gives

$$H = \theta' K,$$

where  $\theta$  is an  $N_1 \times N_2$  matrix whose elements are functions of  $K$  and  $H$ . Because the policymaker can also adjust  $B_{20}$  (modifying  $K$ ), there is typically sufficient flexibility to ensure that this holds. If there is no such combination of  $\theta$  and  $B_{20}$ , then the policy goal is not feasible for the given choice of  $D_2$  (and the resulting value for  $\tilde{\beta}_1$ ). If the policymaker is willing to modify  $D_2$ , even this need not be an obstacle.

Given  $\theta$ , the policymaker can attain  $h_o$  by solving  $h_o = \theta' k_o$  for  $b_2$ . The solution is

$$b_2 = (\theta' \tilde{J}_1 B_{12})^{-1} (h_o - \theta' \tilde{J}_1 b_1).$$

Note that  $\tilde{J}_1$  has rank  $N_1 - \tilde{r}$ , but because  $N_2 \leq N_1 - \tilde{r}$ , the nonsingularity of  $\theta' \tilde{J}_1 B_{12}$  is plausible.

## 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.
- 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.
- Chen, X. and H. White (1998), "Nonparametric Adaptive Learning with Feedback," *Journal of Economic Theory* 82, 190-222.
- 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. Iriberry (2007), "Fatal Attraction: Salience, Naïveté, and Sophistication in Experimental 'Hide-and-Seek' Games," *American Economic Review* 97, 1731-1750.
- 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.

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, 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.

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.

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

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.