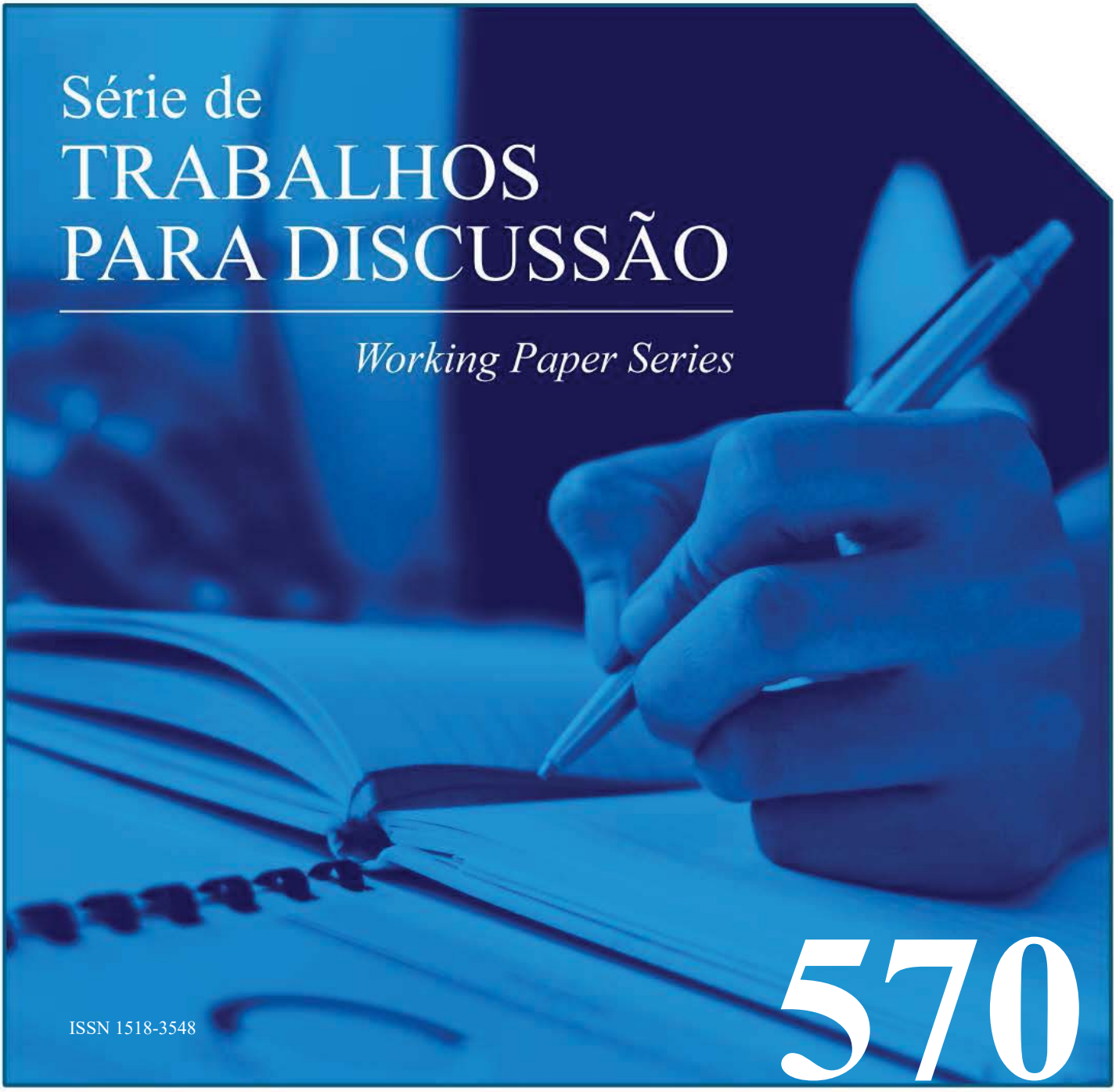

ISSN 1518-3548

September 2022

Causal Impulse Responses for Time Series

Leonardo Marinho

BANCO CENTRAL
DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 570	Setembro	2022	p. 3-57

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

The Banco Central do Brasil Working Papers are evaluated in double-blind referee process.

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-technical summary

The theory of causal inference, comprising Structural Causal Models and graphical models, advocated by Judea Pearl and various researchers, is relatively overlooked by the econometrics literature. The present work is intended to show that this fact is unjustifiable.

For this purpose, objects of interest for macroeconomic policy evaluation are defined, and applications of the developed techniques on simulated time series data are presented in order to show their effectiveness. Then, the same techniques are applied to monetary policy evaluation using real data from Brazil. The main lesson to learn is that much can be done with less assumptions than traditional econometric models.

The distinction between causation and correlation is, on the one hand, apparently obvious. On the other hand, subtle traps can cause lots of confusion among researchers. The strict distinction between causation and correlation, including notational ones, is of central concern in the theory of causal inference adopted here. As shown in the literature, this theory embraces the celebrated Neyman-Rubin potential outcomes model and the ubiquitous Structural Equation Model as particular cases.

Here the effects of policy interventions are evaluated using the central concept of counterfactual. In words, the causal effect over a system arising from a change in some policy variable is the difference between the response of this system after the intervention and the response of the system without the intervention, *all other things being equal*. It is the famous *ceteris paribus*. The hypothetical intervention scenario is called a counterfactual.

At first glance, it is not possible to evaluate these effects without going back in time. But, under some statistical regularity conditions, this is achievable in some cases. The most famous example is the randomized controlled trial – considered the gold standard for the evaluation of causal effects.

Unfortunately, in social sciences in general, and economics in particular, randomized controlled trials are usually impossible, infeasible or unethical. However, the techniques applied here are designed for observational data, i.e., data generated by nature without controlled experiments.

The contribution of this work is threefold: first, impulse response functions are defined in a causal fashion and broad sense, being applicable over a myriad of situations, including non-linear systems. Second, an estimator for the effect of a sequence of policy interventions, called sequential plan, is presented for the linear case. Third, impulse responses of well-known macroeconomic aggregates after a monetary policy intervention are presented for Brazilian data. These latter results have been shown to be more precise, coherent and used less assumptions than the specialized literature. Issues such as “price puzzles” and the real exchange rate response are addressed and clarified. The hope is that the useful tools of causal calculus gain more acceptance among econometricians, turning its marginal role in econometrics closer to the center stage.

Sumário não técnico

A teoria da inferência causal, que abarca os Modelos Causais Estruturais e os modelos gráficos, defendida por Judea Pearl e vários pesquisadores, é relativamente negligenciada pela literatura em econometria. O presente trabalho tem a intenção de mostrar que este fato é injustificável.

Para isto, objetos de interesse para a avaliação de políticas macroeconômicas são definidos, e aplicações das técnicas desenvolvidas a dados de séries temporais simuladas são apresentadas com o intuito de mostrar sua efetividade. Em seguida, as mesmas técnicas são aplicadas à avaliação da política monetária usando dados do Brasil. A principal lição a se aprender é que muito pode ser feito com menos premissas que os modelos econométricos tradicionais.

A distinção entre causalidade e correlação é, por um lado, aparentemente óbvia. Por outro lado, armadilhas sutis podem causar muita confusão entre os pesquisadores. A estrita distinção entre causalidade e correlação, incluindo a distinção de notação, é de central importância para a teoria da inferência causal adotada aqui. Como mostrado na literatura, esta teoria abarca o célebre modelo de resultados potenciais de Neyman-Rubin e o onipresente Modelo de Equações Estruturais como casos particulares.

Aqui, os efeitos de intervenções de política são avaliados através do conceito central de contrafactual. Em palavras, o efeito causal sobre um sistema decorrente de uma mudança em alguma variável de política é a diferença entre a resposta do sistema após a intervenção e a resposta do sistema sem a intervenção, *mantendo tudo o mais constante*. Este é o famoso *ceteris paribus*. O cenário de intervenção hipotética é chamado de contrafactual.

À primeira vista, é impossível avaliar estes efeitos sem voltar no tempo. Porém, sob algumas condições de regularidade estatística, isto é possível em alguns casos. O exemplo mais famoso é o estudo randomizado controlado, considerado o padrão-ouro da avaliação de efeitos causais.

Infelizmente, em ciências sociais em geral, e economia em particular, estudos randomizados controlados são normalmente impossíveis, inviáveis ou antiéticos. Porém, as técnicas aplicadas aqui são projetadas para dados observacionais, isto é, dados gerados de forma natural sem experimentos controlados.

As contribuições deste trabalho são três: primeiro, funções de resposta ao impulso são definidas de forma causal e em sentido geral, sendo aplicáveis a uma miríade de situações, incluindo sistemas não lineares. Segundo, um estimador para o efeito de uma sequência de intervenções de política, chamado plano sequencial, é apresentado para o caso linear. Terceiro, respostas ao impulso de agregados macroeconômicos bem conhecidos após uma intervenção de política monetária são apresentadas para dados brasileiros. Estes últimos resultados se mostraram mais precisos, coerentes e usaram menos premissas que a literatura especializada. Questões como "price puzzles" e a resposta da taxa de câmbio real são analisadas e esclarecidas. A esperança é que estas ferramentas tão úteis ganhem mais aceitação entre os econometristas, saindo da sua posição marginal em econometria para se aproximar do centro do palco.

Causal Impulse Responses for Time Series

Leonardo Marinho*†

Abstract

I develop the concept of impulse response in a causal fashion, defining analytical tools suitable for different policy analysis. Applications of techniques presented to models containing features like confounders or nonlinearities through Monte Carlo experiments are given. I also apply some of these techniques to practical macroeconomic problems, computing impulse responses of GDP, interest rate, inflation and real exchange rate to monetary policy decisions of Banco Central do Brasil, the Brazilian Central Bank.

Keywords: causal analysis, structural causal model, DAG, time series, impulse response, monetary policy.

JEL Classification: C14, C22, E52.

1 Introduction

In this paper I deal with peculiarities raised by time sampled data on causal analysis and provide some examples of causal inference through Monte Carlo experiments and empirical applications. New concepts of impulse responses, more meaningful for policy analysis, are presented. These new definitions also address complications raised by non-linear dynamics.

Examples of identification through causal reasoning based on techniques presented in Pearl [2009], such as do-calculus, counterfactuals and Structural Causal Models (SCM) are given. Also, following the identification stage, causal effects are estimated fitting models suitable for each particular problem. The effects of future policy shocks are discussed, and a formula to approximate the effects of causal plans, understood as a sequence of future actions looking for some policy objective, is presented. Finally, some techniques developed are applied to real data to analyze monetary policy effects in Brazil.

In the economic literature, SCMs have not been widely adopted, but there are some examples in the literature. Moneta and Russo [2014] investigate causal interpretations in the context of economic modeling. Céspedes et al. [2008] identify SVARs with the help of Directed Acyclic Graphs (DAGs). Another part of the literature concentrates on conceptual discussions defending or dismissing SCMs as useful tools in economics, such as White and Chalak [2006], Chalak and White [2011], Pearl [2013], Heckman and Pinto [2015], Pearl [2015], Cunningham [2018], Hünermund and Bareinboim [2019], Imbens [2020], Kincaid [2021].

Here I present applications of the framework on simulated data, to show how they really work, and on real data, to explore its usefulness.

The paper is organized in the following way: Section 2 presents definitions of causal impulse responses. Section 3 applies these ideas to a linear system, exploring its advantages over structural vector autoregressions. Section 4 applies the techniques to a nonlinear system. Section 5 presents a formula for sequential plans evaluation. Section 6 applies the techniques developed on real data from Brazilian economy.

*Banco Central do Brasil (email: leonardo.marinho@bcb.gov.br).

†The Working Paper should not be reported as representing the views of the Banco Central do Brasil. The views expressed in the paper are those of the authors and do not necessarily reflect those of the Banco Central do Brasil.

2 Causal impulse response functions[1]

The main contributions of this paper are the following causal definitions of impulse response functions. These definitions are based on counterfactuals as defined in Structural Causal Models. For a brief introduction, see Appendix A.1.

Definition 2.1. The **counterfactual impulse response function (CIRF)** on horizon h of Y in a dynamic causal system between interventions $do(X_t = x_1)$ and $do(X_t = x_0)$ is given by

$$CIRF_Y^{x_1,x_0}(h, \omega_{t-1}, \epsilon_{t,h}) = Y_{t+h}^{x_1}(\omega_{t-1}, \epsilon_{t,h}) - Y_{t+h}^{x_0}(\omega_{t-1}, \epsilon_{t,h}), \tag{2.1}$$

where $Y_{t+h}^{x}(\omega_{t-1}, \epsilon_{t,h})$ is the potential response of Y_{t+h} given $X_t = x$, history ω_{t-1} and shocks realization $\epsilon_{t,h}$.

This definition uses counterfactuals as presented in Pearl [2009]. The "situation" or "state of the world" is given by the realization $(\Omega_{t-1}, E_{t,h}) = (\omega_{t-1}, \epsilon_{t,h})$, where Ω_{t-1} represents all variables in the dynamic system up to time $t-1$ and $E_{t,h}$ all exogenous disturbances on the system up to time $t+h$.

Definition 2.1 is a comparison between two hypothetical stories which differ only by the intervention performed in X at time t. The *base counterfactual*, $Y_{t+h}^{x_0}(\omega_{t-1}, \epsilon_{t,h})$, is the value Y_{t+h} would have had if an intervention at t given by $do(X_t = x_0)$ had been done with all other relevant endogenous and exogenous variables realization given by ω_{t-1} and $\epsilon_{t,h}$. The *main counterfactual*, $Y_{t+h}^{x_1}(\omega_{t-1}, \epsilon_{t,h})$, is the value Y_{t+h} would have had if an intervention at t given by $do(X_t = x_1)$ had been done with everything else being equal.

Obviously, this CIRF is not observable, unless the exact same situations occur and exogenous interventions are feasible. In practical time series analysis, both situations are impossible or almost that. But, conceptually, Definition 2.1 is the base for the construction of causal impulse responses that may be inferred from observational data.

How Definition 2.1 resemble counterfactuals as defined in Neyman-Rubin potential outcomes approach? Here, set $X_t = x_1$ is to give the treatment status to "individual" (or unit of treatment) $\{\omega_{t-1}, \epsilon_{t,h}\}$ and set $X_t = x_0$ is to give the control status to *the same* "individual". In structural causal models framework, an "individual" is a particular realization of the exogenous variables (see Pearl [2009]).

From Definition 2.1, we follow Koop et al. [1996] and let the background variables vary, treating them as random variables. So, we may ask about (2.1) for arbitrary realizations ω_{t-1} of Ω_{t-1} and $\epsilon_{t,h}$ of $E_{t,h}$, which leads to Definition 2.2:

Definition 2.2. The **generalized counterfactual impulse response function (GCIRF)** on horizon h of Y in a dynamic causal system between interventions $do(X_t = x_1)$ and $do(X_t = x_0)$ is given by

$$GCIRF_Y^{x_1,x_0}(h, \Omega_{t-1}, E_{t,h}) = Y_{t+h}^{x_1}(\Omega_{t-1}, E_{t,h}) - Y_{t+h}^{x_0}(\Omega_{t-1}, E_{t,h}), \tag{2.2}$$

where $Y_{t+h}^{x}(\Omega_{t-1}, E_{t,h})$ is the random variable Y_{t+h} given the random variable history Ω_{t-1}, the random variable shocks $E_{t,h}$ and intervention $do(X_t = x)$.

Of course, CIRFs are particular realizations of GCIRFs. From Definition 2.2 we see that GCIRF is a random variable, and one may define different responses of interest based on properties of this distribution. In particular, we define the *expected generalized counterfactual impulse response function*:

Definition 2.3. The **expected generalized counterfactual impulse response function (EGCIRF)** on horizon h of a dynamic causal system between interventions $do(X_t = x_1)$ and $do(X_t = x_0)$ is given by

$$\begin{aligned} EGCIRF_Y^{x_1,x_0}(h) \equiv & E\left[GCIRF_Y^{x_1,x_0}(h, \Omega_{t-1}, E_{t,h})\right] = \\ & E\left[Y_{t+h}^{x_1}(\Omega_{t-1}, E_{t,h}) - Y_{t+h}^{x_0}(\Omega_{t-1}, E_{t,h})\right], \end{aligned} \tag{2.3}$$

where the expectation on GCIRF is taken relative to Ω_{t-1} and $E_{t,h}$.

[1] A note on notation: in what follows, counterfactuals are defined with "treatment variables" as superscript to avoid confusion with time indexes in subscript. Also, I use summation notation even for continuous random variables for convenience. These must be interpreted as the suitable integral over densities when necessary.

Notice that the EGCIRF is history independent and has a clear counterfactual interpretation. Indeed, we may see it as the expected value of the difference of Y_{t+h} between two conceptual experiments: setting $X_t = x_1$ and $X_t = x_0$. In other words, it represents the average difference of Y_{t+h} after a manipulation of the system setting $X_t = x_1$ against setting $X_t = x_0$, averaged over all relevant variables.

EGCIRF is the difference between two hypothetical interventions. Besides its usefulness, especially in comparing different potential policy decisions, policy makers are regularly interested on the direct effects of their actions compared with no action. One possible impact measure of interest is what would be the response of Y to an intervention in X compared with no intervention, regardless the state of all variables involved. We may achieve this goal comparing the effect of an intervention with what would have been the trajectory of the system without the intervention. This is easily done through GCIRF not setting x_0, which leads to the following definition:

Definition 2.4. The **causal generalized impulse response (CGI)** on horizon h of a dynamic causal system after intervention $do(X_t = x)$ is given by

$$CGI_Y^x(h, \Omega_{t-1}, E_{t,h}) = Y_{t+h}^x(\Omega_{t-1}, E_{t,h}) - Y_{t+h}(\Omega_{t-1}, E_{t,h}), \tag{2.4}$$

where $Y_{t+h}(\Omega_{t-1}, E_{t,h})$ is the value of Y_{t+h} given history Ω_{t-1} and shocks $E_{t,h}$.

As before, this is a random variable. Again we may define its expected value:

Definition 2.5. The **expected causal generalized impulse response (ECGI)** on horizon h of a dynamic causal system after intervention $do(X_t = x)$ is given by

$$ECGI_Y^x(h, \Omega_{t-1}, E_{t,h}) = E\left[Y_{t+h}^x(\Omega_{t-1}, E_{t,h}) - Y_{t+h}(\Omega_{t-1}, E_{t,h})\right], \tag{2.5}$$

where $Y_{t+h}(\Omega_{t-1}, E_{t,h})$ is the value of Y_{t+h} given history Ω_{t-1} and shocks $E_{t,h}$.

Each of these definitions has its own interest and potential applications. In next sections we apply these concepts together with causal calculus to make inferences based on Monte Carlo simulations of various systems.

3 Impulse responses and causal interventions in linear systems

To make things clear, lets begin with linear models. In this class of models, impulse responses are *history independent*, that is, independent of ω_{t-1}. Suppose we want the response of Y_{t+h} for various horizons h to a unitary shock in X_t in the model given by equations (3.1):

$$\begin{aligned} c_t &= 0.4c_{t-1} + 0.3c_{t-2} + 0.2y_{t-1} + \epsilon_t^c \\ x_t &= 0.4z_t - 0.3y_{t-1} + \epsilon_t^x \\ y_t &= 0.5z_t + 0.5x_t + \epsilon_t^y \\ z_t &= 0.6z_{t-1} + 0.5c_t + \epsilon_t^z, \end{aligned} \tag{3.1}$$

where $\epsilon_t^k \sim N(0,1)$ for all $k \in \{c, x, y, z\}$ are serially independent and also jointly independent. In the language of SCMs, if we assume (3.1) as the data generating process, the diagram in Figure 1 represents its causal relations.

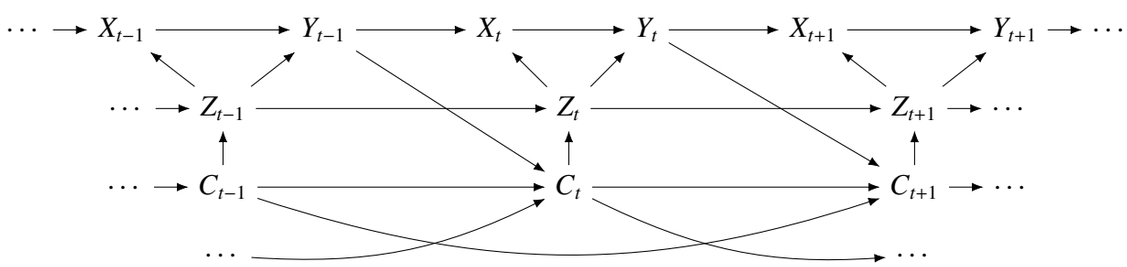

Figure 1: Dynamic causal network representing the dynamic model in (3.1).

This network is an infinite DAG. Looking for equations (3.1), it is clear that it is a SCM as defined in Pearl [2009] whose functions are time invariant and sampling is made across time. Considerations about ergodicity are necessary to show the existence of probability distributions

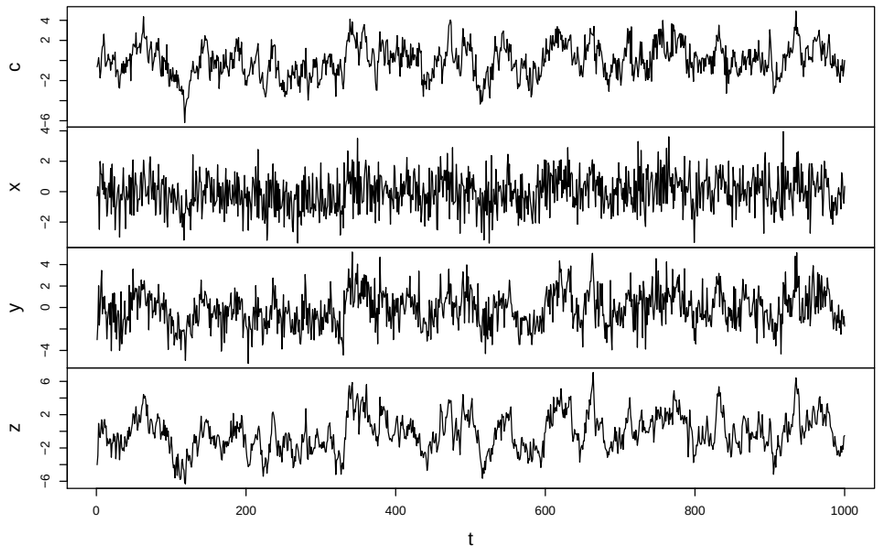

Figure 2: Sample size 1,000 simulation of process (3.1). Burn-in period is 100 time samples.

assumed in the examples, but I will not address these issues here. Figure 2 shows a sample simulation of (3.1).

Lets compute the EGCIRF for this system between interventions $do(X_t = 1)$ and $do(X_t = 0)$. From equation (2.3), and remembering the relation between counterfactuals and the "do" operator shown in Pearl et al. [2016], we get:

$$\begin{aligned} EGCIRF_Y^{1,0}(h) =& E\left[Y_{t+h}^1(\Omega_{t-1}, E_{t,h}) - Y_{t+h}^0(\Omega_{t-1}, E_{t,h})\right] = \\ & E[Y_{t+h}|do(X_t = 1)] - E[Y_{t+h}|do(X_t = 0)]. \end{aligned} \tag{3.2}$$

So, by the back-door criterion applied to Figure 1 for each t, if we adjust for Y_{t-1} and Z_t we can compute the causal effect from X_t to any horizon Y_{t+h}, $h \in \mathbb{N}$. Actually, we may write:

$$P[Y_{t+h}|do(X_t = \delta)] = \sum_{Y_{t-1}, Z_t} P(Y_{t+h}|X_t = \delta, Y_{t-1}, Z_t)P(Y_{t-1}, Z_t), \ \forall t \in \mathbb{Z} \tag{3.3}$$

by the back-door formula. Multiplying (3.3) by Y_{t+h} and summing, gives:

$$E[Y_{t+h}|do(X_t = \delta)] = \sum_{Y_{t-1}, Z_t} E(Y_{t+h}|X_t = \delta, Y_{t-1}, Z_t)P(Y_{t-1}, Z_t). \tag{3.4}$$

Finally, take the difference of (3.4) for $\delta = 1$ and $\delta = 0$:

$$\begin{aligned} & E[Y_{t+h}|do(X_t = 1)] - E[Y_{t+h}|do(X_t = 0)] = \\ & \sum_{Y_{t-1}, Z_t} \left[E(Y_{t+h}|X_t = 1, Y_{t-1}, Z_t) - E(Y_{t+h}|X_t = 0, Y_{t-1}, Z_t)\right]P(Y_{t-1}, Z_t) \end{aligned} \tag{3.5}$$

which is the EGCIRF for this system. Since the model is linear, the conditional expectations may be written as

$$E(Y_{t+h}|X_t, Y_{t-1}, Z_t) = \alpha_h + \beta_h^{X_t} X_t + \beta_h^{Y_{t-1}} Y_{t-1} + \beta_h^{Z_t} Z_t. \tag{3.6}$$

Substituting (3.6) into (3.5) gives

$$E[Y_{t+h}|do(X_t = 1)] - E[Y_{t+h}|do(X_t = 0)] = \beta_h^{X_t}. \tag{3.7}$$

Hence, the causal effect on the expected value of Y_{t+h} by setting $X_t = 1$ against setting $X_t = 0$ is given by (3.7), which can be estimated by OLS from (3.6). This approach is reminiscent of impulse responses estimation by local projections presented in Jordà [2005], but here we explicitly address the causal nature of impulse responses.

Figure 3 shows the theoretical IRF of (3.1) along with estimates using (3.7) from simulations of (3.1). Results show a nice fit to the theoretical curve.

From a more traditional econometrics viewpoint, this model can be put in SVAR (Structural Vector Autoregression) form:

$$A\psi_t = B_1\psi_{t-1} + B_2\psi_{t-2} + \epsilon_t, \tag{3.8}$$

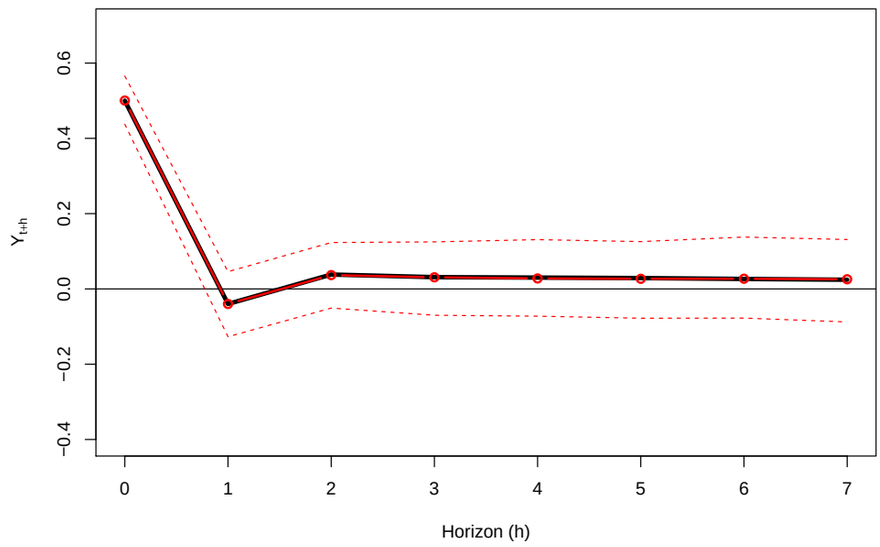

Figure 3: The thick black line represents theoretical IRF of Y_t in (3.1) for a unit shock in X_t. The red doted line represents the mean of 1000 estimates using (3.7) estimated from 1000 simulations of system (3.1). Each simulation generates series of length 1000 after a burn-in period of 100 time samples. Dashed red lines represent percentiles 2.5% and 97.5% of $\beta_h^{X_t}$ estimates.

where $\psi_t = \begin{pmatrix} c_t \\ x_t \\ y_t \\ z_t \end{pmatrix}$, $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -0.4 \\ 0 & -0.5 & 1 & -0.5 \\ -0.5 & 0 & 0 & 1 \end{pmatrix}$, $B_1 = \begin{pmatrix} 0.4 & 0 & 0.2 & 0 \\ 0 & 0 & -0.3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6 \end{pmatrix}$, $B_2 = \begin{pmatrix} 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ and $\epsilon_t = \begin{pmatrix} \epsilon_t^c \\ \epsilon_t^x \\ \epsilon_t^y \\ \epsilon_t^z \end{pmatrix}$.

The same results could be achieved estimating an unrestricted order 2 VAR and identifying the A matrix in (3.8) restricting all but the instantaneous coefficients. So, what is the advantage of the presented approach over SVARs?

Consider this: what if C_t was an unobserved variable? Our results would be untouched, since we do not need C_t to control for confounding based on Figure 1. But the SVAR will not work. It may be argued that equations for x_t and y_t in (3.1) can be estimated by OLS, but the equation for z_t, which is essential to the dynamics of x_t and y_t, cannot. Some econometricians may suggest an instrumental variable approach, which depends on the existence of a suitable instrument. It is possible to infer from equations (3.1) that adjusting for y_{t-1} is sufficient to estimate the IRF, but this is not as obvious as is just looking to Figure 1. And this is the main reason why Pearl [2009] insists on a specific language for causality and the use of causal graphs: simplicity and intuitive appeal. Another advantage is to be a common language for identification of causal effects in general, which helps to connect ideas apparently unrelated.

To be fair, the SVAR approach is still useful in identifying the IRF from X_t to Y_t, especially if we choose the restrictions in the structural matrix based on causal information provided by Figure 1, showing the usefulness of causal models even for traditional econometric approaches. Consider again a SVAR, but now with only variables X_t, Y_t and Z_t. It is clear that contemporaneous restrictions over A are given by Figure 1 looking for a time slice t. Also, effects through the unobservable variable C_t come from longer lags than only 2, because we cannot control for C_t anymore. This sort of confounder is called a *dynamic confounder* by Blondel et al. [2017]. We can control for it with two more lags in the reduced form VAR and achieve the same results as before, but in a clearly more convoluted way.

Our linear example does not show the full power of the presented approach, since it is tractable by SVARs. But, as mentioned, the observability status of a confounder variable such as C_t may make the analysis based on automatic application of SVARs problematic, and causal reasoning based on DAGs may also help even in SVAR identification. Additionally, the non-linear case, beyond the reach of SVARs, is much more complicated, since the concept of traditional impulse response function itself loses its meaning. This will be discussed in next section.

4 Impulse responses and causal interventions in non-linear systems

Koop et al. [1996] introduce the concept of *generalized impulse response function*, which tries to overcome difficulties raised by non-linear systems. As seen in Section 2, we built our analysis over their ideas, although emphasizing the causal nature of IRFs.

The main difficulty raised by non-linear systems stems from its history and shock dependence. In Section 3, the specific values of X_t, Y_t, Z_t and C_t at the moment of the shock do not have any effect over the IRF. The size of the shock may also be rescaled to any size desired, simply multiplying the response by the same amount. These properties have important consequences for the interpretation of what *is* an IRF. For example, in Section 3 we treated the operation $do(X_t = 1)$ as if it was an intervention to set $X_t = 1$. But, in model (3.1), a unit shock in X_t is to set $\epsilon_t^x = 1$, according to prevalent econometric jargon. Nonetheless, both affirmations are the same in linear models, since the deterministic part of X_t, namely, $0.4z_t - 0.3y_{t-1}$, is already given. So, when we use Koop et al. [1996] definition of traditional impulse response function, the deterministic term disappears, leaving only the shock ϵ_t^x.

However, in non-linear models this is not true. Writing the equation for X_t in a more general form, like

$$X_t = f(pa(X_t), \epsilon_t^x), \tag{4.1}$$

for arbitrary f and exogenous (as in the linear case) ϵ_t^x, where $pa(X_t)$ are the parents of X_t. In this case, not necessarily the difference between system responses for $do(X_t = 1)$ and $do(X_t = 0)$ are the same as the difference setting $\epsilon_t^x = 1$ against $\epsilon_t^x = 0$. Furthermore, the additive shock interpretation used in econometrics becomes obscured because $f(pa(X_t), \epsilon_t^x)$ is now history dependent, and the quantity $E(X_t|pa(X_t), \epsilon_t^x = 1) - E(X_t|pa(X_t), \epsilon_t^x = 0)$ is a function of $pa(X_t)$, since its effects do not necessarily cancel as in the linear case. So, for different values of $pa(X_t)$, $\epsilon_t^x = x$ induces different values for X_t depending on the system history. Meanwhile, the "do" operation remains unambiguous, since it imposes the substitution of (4.1) by $X_t = x$, setting unequivocally X_t.

Of course, it may be of interest the effect of a manipulation for a given X_t, such as the effect of raising X_t by one unit. This situation is still complicated in (4.1) for the additive shock approach, since ϵ_t^x does not enter into this equation as an additive term. Indeed, Koop et al. [1996] define their system with additive shocks, probably to avoid this sort of complication. The "do" operator also cannot solve this problem, but counterfactuals deal with it naturally. For details, see Pearl [2009] and Pearl et al. [2016].

As an example of application of the concepts in non-linear case, consider a non-linear version of model (3.1) given by equations (4.2):

$$\begin{aligned} c_t &= 0.4c_{t-1} + 0.3c_{t-2} + 0.2y_{t-1} + \epsilon_t^c \\ x_t &= 0.4z_t - 0.3y_{t-1} + \epsilon_t^x \\ y_t &= 4\cos(0.5z_t + 0.5x_t) + \epsilon_t^y \\ z_t &= 0.6z_{t-1} + 0.5c_t + \epsilon_t^z \end{aligned} \tag{4.2}$$

where $\epsilon_t^k \sim N(0, 1)$ for all $k \in \{c, x, y, z\}$ are serially independent and also jointly independent, as before. The infinite DAG representing this system is the same as in Figure 1, but now the system is non-linear because of the cosine in the equation for Y_t. One way to simulate the GCIRF (2.2) for model (4.2) is the following:

1. Simulate the system for a set of random shocks ϵ^i and initial values ω_0^i setting $X_t = x_1$ for some sufficiently large t and collect Y_{t+h} for all h, here denoted $Y_{t+h}^{x_1}(i)$;

2. Repeat the simulation *for the same set of shocks and initial values* of item 1, but setting $X_t = x_0$ on the same t and collect Y_{t+h} for all h, now called $Y_{t+h}^{x_0}(i)$;

3. Collect the difference $Y_{t+h}^{x_1}(i) - Y_{t+h}^{x_0}(i)$. This represents one sample of the distribution $GCIRF_Y^{x_1,x_0}(h, \Omega_{t-1}, E_{t,h})$;

4. Repeat steps 1, 2 and 3 until collecting sufficient samples of $GCIRF_Y^{x_1,x_0}(h, \Omega_{t-1}, E_{t,h})$.

Figure 4 shows the mean and percentiles 2.5% and 97.5% of 10,000 samples of the simulated $GCIRF_Y^{x_1,x_0}(h, \Omega_{t-1}, E_{t,h})$ with $x_1 = 1$ and $x_0 = 0$ for process (4.2).

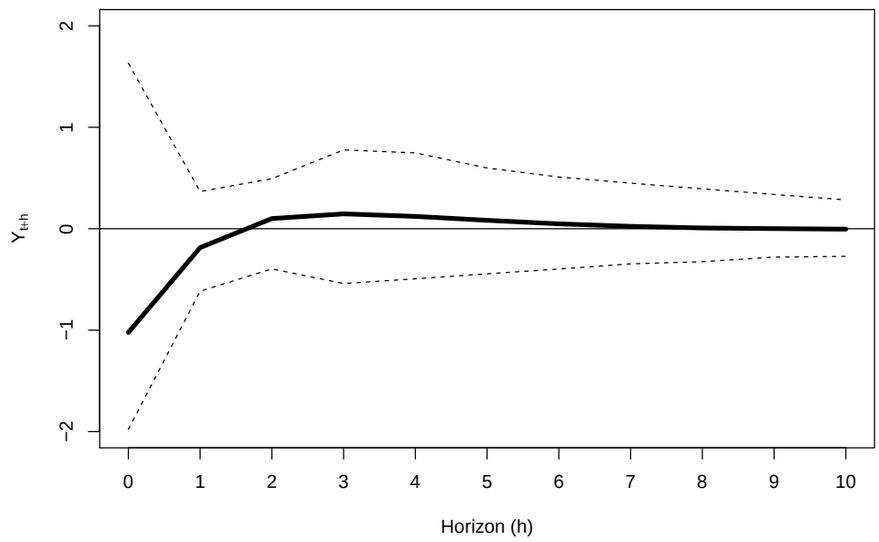

Figure 4: Generalized counterfactual impulse response function (GCIRF) for Y_{t+h} with $x_1 = 1$ and $x_0 = 0$. Thick solid line is the mean and dashed lines are percentiles 2.5% and 97.5% of 10,000 samples.

Notice that the thick solid line represents the expected generalized counterfactual impulse response function (EGCIRF) for this system. We may estimate it by the following procedure: from equation (2.3), we have

$$\begin{aligned} EGCIRF_Y^{x_1,x_0}(h) =& E\left[Y_{t+h}^{x_1}(\Omega_{t-1}, E_{t,h}) - Y_{t+h}^{x_0}(\Omega_{t-1}, E_{t,h})\right] = \\ & E[Y_{t+h}|do(X_t = x_1)] - E[Y_{t+h}|do(X_t = x_0)], \end{aligned} \tag{4.3}$$

with $x_1 = 1$ and $x_0 = 0$. By the back-door criterion,

$$\begin{aligned} & EGCIRF_Y^{1,0}(h) = \\ & \sum_{Y_{t-1},Z_t} [E(Y_{t+h}|X_t = 1, Y_{t-1}, Z_t) - E(Y_{t+h}|X_t = 0, Y_{t-1}, Z_t)]P(Y_{t-1}, Z_t). \end{aligned} \tag{4.4}$$

To estimate (4.4) we may first estimate the conditional expectation function $f(x, Y_{t-1}, Z_t) = E(Y_{t+h}|X_t = x, Y_{t-1}, Z_t)$. Then, we estimate the expected value of $g(Y_{t-1}, Z_t) = f(1, Y_{t-1}, Z_t) - f(0, Y_{t-1}, Z_t)$, substituting sample values for Y_{t-1} and Z_t and taking the average.

Of course this procedure is much simpler for linear f, leading to equation (3.7). This linear approach is shown in Figure 5. Notice that estimates are biased, especially for $h = 0$. This is because (4.2) is non-linear.

The non-linearity of the model may be inferred from data through standard diagnostics of regression estimates. This is important, because separating causal identification from estimation makes it easier know from data if we have a misspecification due to a bad causal model (which can be tested through conditional independence implications of the causal diagram) or to a bad statistical model (which can be tested through residual analysis and lots of other tools available from statistics).

As an example of these diagnostics, consider Figure 5. This plot shows a tentative estimate of the EGCIRF for Y_{t+h} in model (4.2) with $x_1 = 1$ and $x_0 = 0$ and $h = 0$ to 10. I assume here the correct causal inference based on Figure 1, because causal relations do not change from system (3.1) to (4.2). But the functional form is different, and this is just a statistical modeling problem, not a causal one.

In plotting Figure 5, I estimate the conditional expectation $E(Y_{t+h}|X_t, Y_{t-1}, Z_t) = \alpha_h + \beta_h^{X_t} X_t + \beta_h^{Y_{t-1}} Y_{t-1} + \beta_h^{Z_t} Z_t$ for each h with 1,000 simulations of (4.2) through linear regression, looking for the causal parameter $\beta_h^{X_t}$. Table 1 shows results for one of these regressions, namely, one of all 1,000 runs for $h = 0$.

At first sight this table is fine, pointing to a causal effect of -0.496 at $h = 0$ (for this sample process). But looking to residuals we see that this model is probably misspecified. Figure 6 shows its fitted values against residuals, a standard plot for functional form misspecification diagnostic. As can be seen, the model is badly specified, and this plot suggest a non-linear relationship among

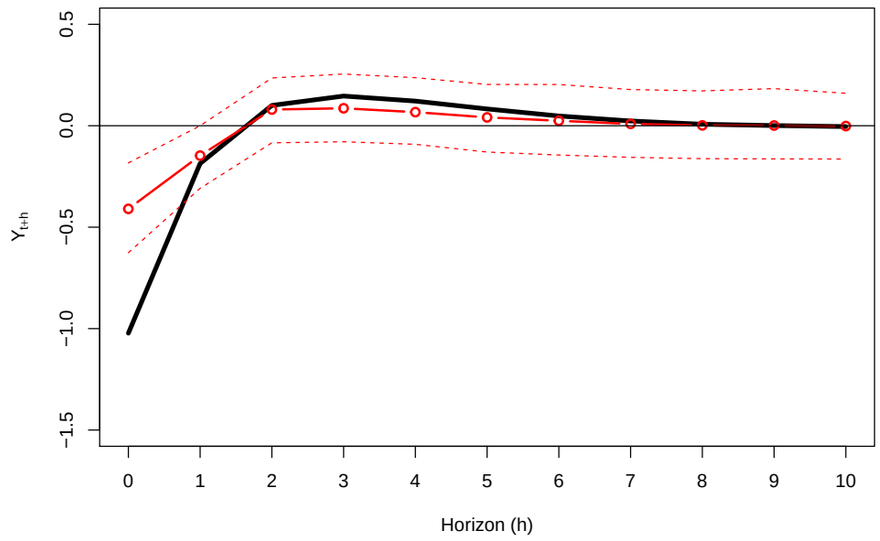

Figure 5: Thick solid line is the expected generalized counterfactual impulse response function (EGCIRF) for Y_{t+h} in model (4.2) with $x_1 = 1$ and $x_0 = 0$. Red dotted dashed line is the mean of 1,000 estimates of EGCIRF based on linear conditional expectation estimates of 1,000 simulations of (4.2). Red dashed lines are percentiles 2.5% and 97.5% for these estimates. Each simulation has 1000 data points after a burn-in period of 100.

	Dependent variable:
	Y_t
X_t	-0.496^{***}
	(0.070)
Z_t	-0.555^{***}
	(0.054)
Y_{t-1}	0.169^{***}
	(0.033)
Constant	1.886^{***}
	(0.099)
Observations	999
R^2	0.343
Adjusted R^2	0.341
Residual Std. Error	2.178 (df = 995)
F Statistic	173.224^{***} (df = 3; 995)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1: Regression results for $EGCIRF_Y^{1,0}(h)$ with $h = 0$ estimated from one simulation of (4.2). The causal effect of interest is the coefficient for X_t.

regressors and regressand, as expected. Looking for other samples of the process is not necessary. Actually, in practice the estimation is made over one sample alone, since simulations of the process are not available.

To deal with the non-linearity found, higher order terms on the regression may be added, for example. However, here I choose a different approach, through the use of generalized additive models (GAM) (see Hastie and Tibshirani [1987], Hastie [2017], Wood [2017]).

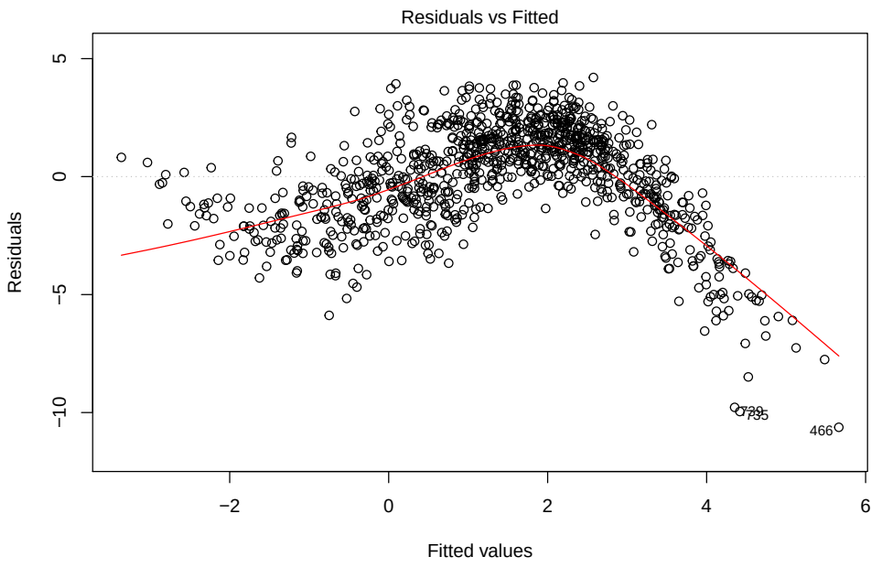

Figure 6: Scatter plot of fitted values vs residuals for model in Table 1. Clearly there is a non-linear relationship between variables.

I choose GAMs for three main reasons. First, it is flexible and powerful. Second, its application is straightforward in R with package *mgcv* (R Core Team [2019], Wood [2017]). Third, with GAMs I show how different techniques may be used to estimate conditional expectations whose causal interpretation is already known. For a brief presentation of GAMs and some details on the model used here, see Appendix A.4.

Figure 7 shows the same EGCIRF 4.4 with $g(Y_{t-1}, Z_t) = f(1, Y_{t-1}, Z_t) - f(0, Y_{t-1}, Z_t)$ estimated by GAMs. The fit is much better, almost without bias. Notice that in the analysis no knowledge about the DGP beyond that in Figure 1 was used. This knowledge is the same for the linear and non-linear versions, because causal links are abstracted from functional forms. Furthermore, I treat C_t as a non-observable confounder.

Now lets see the CGI, presented in Definition 2.4, for this model. CGI is also a random variable. It may be sampled in a way similar to what was done for the GCIRF. For that it is enough to just let x_0 free to vary. Figure 8 shows the mean and percentiles 2.5% and 97.5% of 10,000 samples of $CGI_Y^x(h, \Omega_{t-1}, E_{t,h})$ with $x = -1$ for process (4.2).

Notice again that the thick black line in Figure 8 represents the ECGI for system (4.2). We may estimate it using GAMs, as before. The result for 1,000 simulations is shown in Figure 9.

These results are remarkable because of the assumptions substantiating them. Basically, only a *qualitative* causal model is necessary, given by a causal graph like that of Figure 1. Even with unobserved confounders and non-linear dynamics, I was able to estimate with high precision some impulse responses, here defined as exogenous interventions of interest. Also, the assumptions are weaker than those normally presumed in econometric evaluation of policy interventions in time series.

All previously presented estimating procedures concentrated on average impulse responses defined in general models because of their simplicity. It is also possible to extend the analysis for estimation of other moments or full distributions such as GCIRF or CGI, for example. In principle, distributional problems are conceptually the same regarding causal identification, and distribution estimation is feasible through dozens of statistical and machine learning procedures available in the specialized literature. Another possible extension with great interest for policy analysis is the evaluation of conditional forecasts, for example.

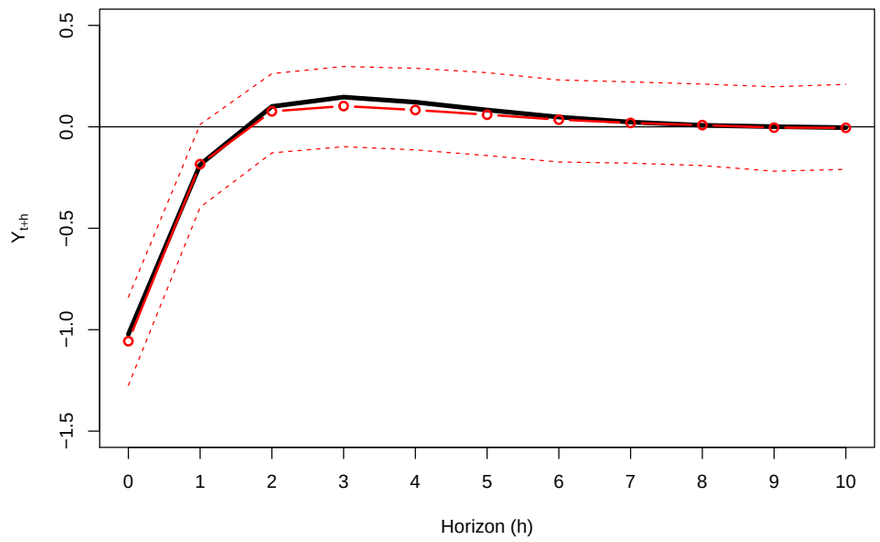

Figure 7: Thick solid line is the expected generalized counterfactual impulse response function (EGCIRF) for Y_{t+h} in model (4.2) with $x_1 = 1$ and $x_0 = 0$. Red dotted dashed line is the mean of 1,000 estimates of EGCIRF based on generalized additive models conditional expectation estimates of 1,000 simulations of (4.2). Red dashed lines are percentiles 2.5% and 97.5% for these estimates. Each simulation has 1,000 data points after a burn-in period of 100.

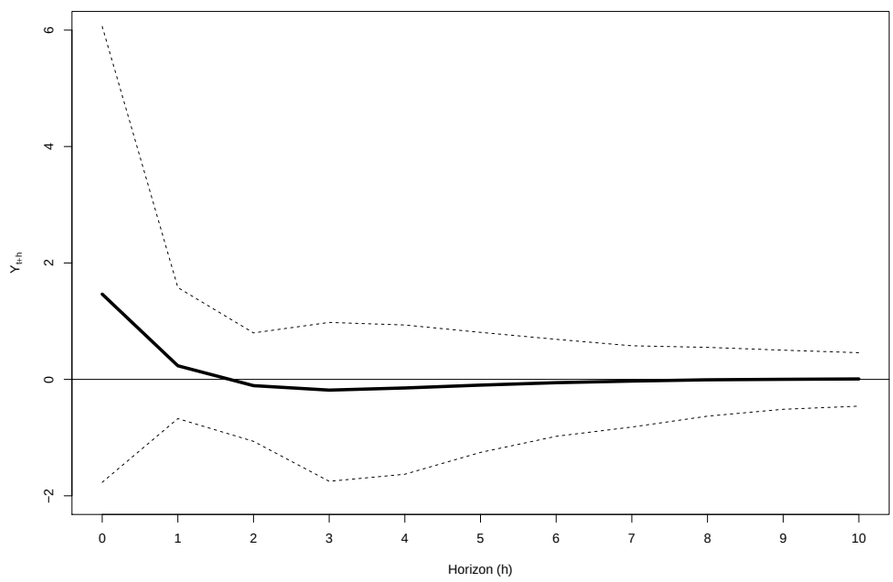

Figure 8: Causal generalized impulse response (CGI) for Y_{t+h} with $x = -1$. Thick solid line is the mean and dashed lines are percentiles 2.5% and 97.5% of 10,000 samples.

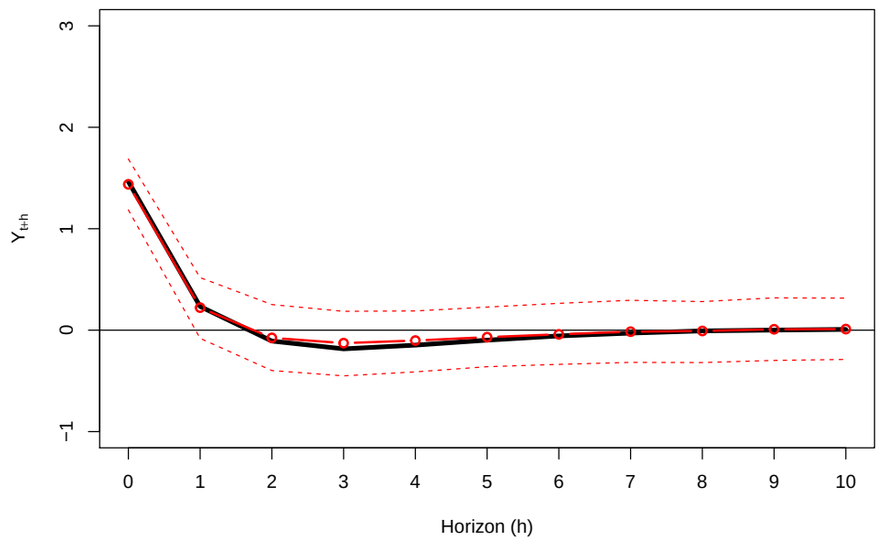

Figure 9: Thick solid line is the expected causal generalized impulse response (ECGI) for Y_{t+h} in model (4.2) with $x = -1$. Red dotted dashed line is the mean of 1,000 estimates of ECGI based on generalized additive models conditional expectation estimates of 1,000 simulations of (4.2). Red dashed lines are percentiles 2.5% and 97.5% for these estimates. Each simulation has 1,000 data points after a burn-in period of 100.

5 Sequential plan response

In previous sections we presented alternative definitions which not only address the lack of causal meaning on previous popular IRF definitions in the literature, but also face difficulties raised by non-linear models.

Here I raise a discussion about future shocks. Observe that in Koop et al. [1996] definition of traditional IRFs, future shocks are set to zero. In linear models with zero expectations exogenous shocks this is almost immaterial. Actually, since our interest is on inference about counterfactual interventions, namely, a comparison between two hypothetical experiments leaving *all else equal*, including future shocks, in models where shocks are additive, the difference among two hypothetical histories would cancel all future shocks. As was observed in Section 4, this is not true in general.

Also, in the presented framework, interventions are made directly on policy variables. So, instead of talking about future shocks, I talk about *future interventions*. The methods developed in earlier sections predict the response of the system to an exogenous manipulation of the policy variable at t, leaving this policy variable free to vary endogenously after the intervention. Now we are interested in the effect of a sequence of exogenous interventions through time.

A very useful concept in discussing future interventions is G-identification, which is a sufficient criterium for causal identification (Pearl and Robins [1995], Robins [1997], Gill and Robins [2001], Pearl [2009]). Below, $G_{\underline{X}_k,\overline{X}_{k+1},\ldots,\overline{X}_n}$ is a causal graph G excluding all arrows pointing out of X_k and all arrows pointing into $X_{k+1}, \ldots, X_n$:

Theorem 5.1 (Pearl [2009])**.** *The probability $P(y|\hat{x}_1, \ldots, \hat{x}_n)$ is G-identifiable in a causal graph G if and only if the following condition holds for every $1 \leq k \leq n$:*

$$(Y \perp\!\!\!\perp X_k | X_1, \ldots, X_{k-1}, W_1, \ldots, W_k)_{G_{\underline{X}_k,\overline{X}_{k+1},\ldots,\overline{X}_n}} \tag{5.1}$$

where W_k is the set of all covariates in G that are both non-descendants of $\{X_k, X_{k+1}, \ldots, X_n\}$ and have either Y or X_k as descendants in $G_{\underline{X}_k,\overline{X}_{k+1},\ldots,\overline{X}_n}$. Moreover, if this condition is satisfied, the plan evaluates as

$$P(y|\hat{x}_1, \ldots, \hat{x}_n) = \sum_{w_1,\ldots,w_n} \left[P(y|w_1, \ldots, w_n, x_1, \ldots, x_n) \times \prod_{k=1}^{n} P(w_k|w_1, \ldots, w_{k-1}, x_1, \ldots, x_{k-1}) \right]. \tag{5.2}$$

A plan consisting of a sequence of interventions is exactly what happens when we set not just the shock at time t, but also at $t+1, t+2, \ldots$. Consider that each w_i in Theorem 5.1 consists of a set of variables $\{w_{i,1}, w_{i,2}, \ldots, w_{i,l_i}\}$. Then, under conditions outlined in Theorem 5.1, the non-parametric causal effect on horizon h of a sequence of interventions $\{\hat{x}_1, \ldots, \hat{x}_h\}$ is:

$$P(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \sum_{w_{1,1},\ldots,w_{h,l_h}} P(y_h|w_{1,1},\ldots,w_{h,l_h},x_1,\ldots,x_h) \times \prod_{k=1}^{h} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}). \tag{5.3}$$

Multiplying both sides by y_h and summing through this variable, we arrive at the non-parametric expected causal effect:

$$E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \sum_{w_{1,1},\ldots,w_{h,l_h}} E(y_h|w_{1,1},\ldots,w_{h,l_h},x_1,\ldots,x_h) \times \prod_{k=1}^{h} P(w_{k,1},\ldots,w_{k,l_h}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}). \tag{5.4}$$

The conditional expectations in (5.4) may be approximated by linear functions, such that non-biased estimates of these approximations through OLS are obtainable. Moreover, this procedure can be iterated in such a way that an estimator for (5.4) is acquired. In fact, this is the content of the following proposition:

Proposition 5.2. *The expected causal effect in* (5.4) *may be approximated by the equation*

$$\begin{aligned} E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = {} & \alpha^{y_h} + \sum_{m=2}^{h}\sum_{n=1}^{l_m} \beta_{m,n}^{y_h}\alpha(m-1)^{m,n} \\ & + \sum_{j=1}^{l_1}\left[\beta_{1,j}^{y_h} + \sum_{m=2}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\beta(m-1)_{1,j}^{m,n}\right]E(w_{1,j}) \\ & + \sum_{i=1}^{h-1}\left[\gamma_i^{y_h} + \sum_{m=i+1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\gamma(m-1)_i^{m,n}\right]x_i + \gamma_h^{y_h}x_h, \end{aligned} \tag{5.5}$$

where α^{y_h}, $\beta_{m,n}^{y_h}$ and $\gamma_m^{y_h}$ are given by

$$E(y_h|w_{1,1},\ldots,w_{h,l_h},x_1,\ldots,x_h) = \alpha^{y_h} + \sum_{m=1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}w_{m,n} + \sum_{m=1}^{h}\gamma_m^{y_h}x_m \tag{5.6}$$

and $\alpha(m-1)^{m,n}$, $\beta(m-1)_{1,j}^{m,n}$ and $\gamma(m-1)_i^{m,n}$ are recursively calculated by the following equations:

$$\begin{aligned} \alpha(q+1)^{m,n} &= \alpha(q)^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\alpha(1)^{m-q,p} \\ \beta(q+1)_{i,j}^{m,n} &= \beta(q)_{i,j}^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\beta(1)_{i,j}^{m-q,p} \\ \gamma(q+1)_i^{m,n} &= \gamma(q)_i^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\gamma(1)_i^{m-q,p}, \end{aligned} \tag{5.7}$$

where $1 \leq q \leq m-1$ and $\alpha(1)^{m,n}$, $\beta(1)_{i,j}^{m,n}$ and $\gamma(1)_i^{m,n}$ are given by

$$\begin{aligned} & E(w_{m,n}|w_{1,1},\ldots,w_{m-1,l_{m-1}},x_1,\ldots,x_{m-1}) = \\ & \alpha(1)^{m,n} + \sum_{i=1}^{m-1}\sum_{j=1}^{l_i}\beta(1)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(1)_i^{m,n}x_i. \end{aligned} \tag{5.8}$$

Proof. See appendix A.2. □

So, OLS regressions may be used to estimate (5.6) and (5.8) and plug these results into (5.7) and (5.5). Notice that in (5.5), the control variable is a vector $x = (x_1,\ldots,x_h)$. Therefore, it is meaningless to ask about the average response to a unit variation in x. We may ask about the response to a *plan* $(x_1,\ldots,x_h)$. For example, we may ask about the response to a unitary shock in x_1 and zero shocks from x_2 to x_h. Another interesting possibility is to ask about the sequence $(x_1,\ldots,x_h)$ which maximizes (or minimizes) response variation. One more possibility is to find an

optimal plan, for some optimality criterion. Finally, we supposed x continuous, but proposition 5.2 may be easily adapted to discrete case.

Now, consider the model:

$$\begin{aligned} x_t &= x_{t-1} + \epsilon_t^x \\ y_t &= \frac{1}{\sum_{i=0}^{8} \exp[-(-2+i/2)^2]} \sum_{i=0}^{8} e^{-(-2+i/2)^2} e_{t-i} \\ z_t &= x_t + y_t, \end{aligned} \tag{5.9}$$

where exogenous shocks ϵ_t^x are given by:

$$\epsilon_t^x \sim N\left(0, 0.2^2\right) \tag{5.10}$$

and policy shocks e_t are given by:

$$e_t \sim N\left[\frac{-1}{140}\sum_{i=1}^{7} i^2 \Delta z_{t+i-8}, \frac{1}{6}\sum_{i=1}^{7}\left(\Delta z_{t+i-8} - \frac{1}{7}\sum_{i=1}^{t}\Delta z_{t+i-8}\right)^2\right]. \tag{5.11}$$

Suppose that only z_t and the policy variable e_t are observable. Also, notice that e_t is endogenous and given by past values of z_t. Indeed, (5.11) is a *stochastic policy rule* that tries to smooth z_t, giving negative shocks to the system if z_t grows too fast and vice versa. A typical realization of this system is shown in Figure 10.

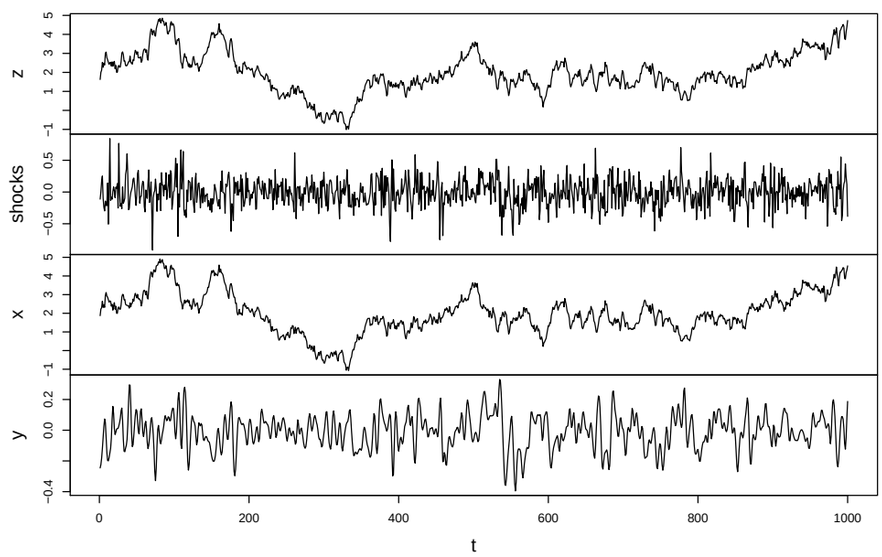

Figure 10: Size 1,000 sample simulation of process (5.9). Burn-in period is 100 time samples.

Notice that long-run movements are driven by random walk x_t while short run deviations are driven by stochastic policy e_t, called shocks in Figure 10. A DAG representing part of this model is given in Figure 11. Not all connections are represented because they are not relevant for the argument. The main point of Figure 11 is that setting $X_{k+i} = e_{t+i}$ and $W_t = Z_{t-1}$ in Theorem 5.1, the conditions for G-identifiability are fulfilled and Proposition 5.2 is directly applicable.

The expected generalized counterfactual impulse response function (EGCIRF) for this system between interventions $do(e_t = 1)$ and $do(e_t = 0)$ is given by

$$EGCIRF_Z^{1,0}(h) = E[Z_{t+h}|do(e_t = 1)] - E[Z_{t+h}|do(e_t = 0)], \tag{5.12}$$

which raises some problems. Actually, Z_t is clearly non-stationary, a fact that may be inferred from Equation 5.9, Figure 10 or unit roots tests. So, equation (5.12) does not make sense, since Z_t does not have a well defined expected value. A common approach to overcome this is to take the first difference of the series. But this raises other problems. Even if the identification of causal effects is not affected, the effect on the variation of Z over one single period will be measured, which is probably a weak effect very difficult to estimate without lots of data. Also, generally this is not the effect of interest for policy makers. A more meaningful measure in terms of policy and statistics is

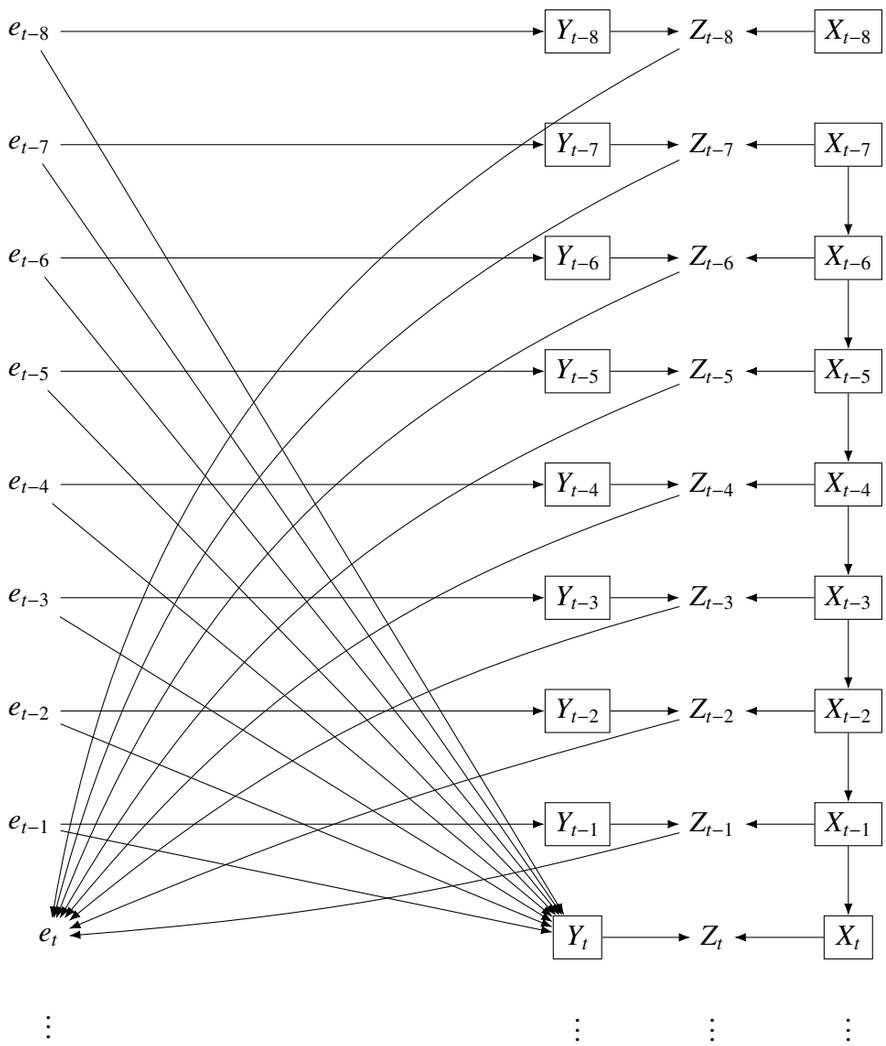

Figure 11: Dynamic causal network representing relevant aspects of the dynamic model in (5.9).

the total variation over horizon h, $\Delta_h Z_t = Z_{t+h} - Z_t$, which has a well defined expected value[2] for each t. Also, the presence of Z_t in the response variable does not affect the identification of causal effects in this case. So, the EGCIRF of interest is

$$EGCIRF^{1,0}_{\Delta_h Z}(h) = E[\Delta_h Z_t | do(e_t = 1)] - E[\Delta_h Z_t | do(e_t = 0)], \tag{5.13}$$

which can be simulated as before, and results are shown in Figure 12.

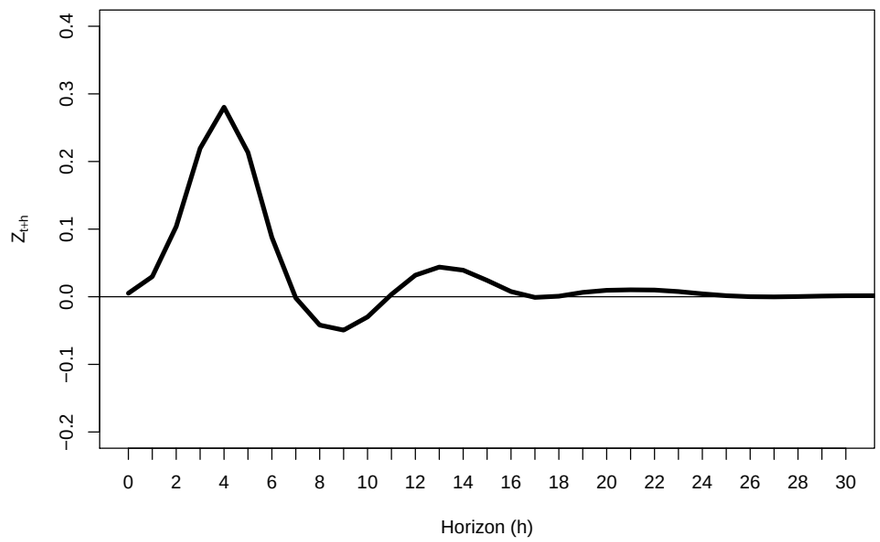

Figure 12: EGCIRF over $\Delta_h Z_t$ for model (5.9) with intervention $do(e_t = 1)$ against $do(e_t = 0)$.

A remarkable feature of Figure 12 is the oscillatory pattern shown by $\Delta_h Z_t$ after a unitary shock $do(e_t = 1)$ compared with a $do(e_t = 0)$ shock. This is due to the endogeneity of e_t. Since only interventions at t are considered here, future shocks are free to vary. It is somewhat like a shock in traditional econometric sense, such as a Taylor rule. The policy variable is shocked at t, but follows its own policy rule afterwards. In the case under analysis, a positive shock in e_t causes Z_t to go up gradually. But e_{t+k} responds negatively to increasing $Z_{t+k-1}, Z_{t+k-2}, \ldots$. This negative response in turn decelerates Z, which induces positive responses in e and so on, causing the oscillatory pattern.

Since this model is linear, we may estimate its EGCIRF as before, through equation (3.7). This is shown in Figure 13. The most salient feature here, besides the unbiasedness of the estimates, is the increasingly less precise estimation as the horizon h goes larger. This is so because the model is non-stationary, and the variance of $\Delta_h Z_t$ raises linearly with h.

In the counterfactual definitions of IRFs presented in Section 2, the comparison is made between two hypothetical interventions, leaving *all else equal*, including non-manipulated stochastic shocks. So, for linear models with *additive zero mean independent* shocks, setting future shocks to zero is immaterial, since all cancel out and, because of independence, a manipulation of a shock at time t does not affect any future shock. This is not the case for the model defined by equations (5.9), (5.10) and (5.11), where policy shocks are endogenous. Thus, set future shocks in this model has a non trivial effect over its response.

Lets evaluate the EGCIRF between interventions $do(X_t = 1, X_{t+1} = 0, \ldots, X_{t+h} = 0)$ and $do(X_t = 0, X_{t+1} = 0, \ldots, X_{t+h} = 0)$ for this model. I call its general form the expected counterfactual difference sequential plan response (ECDSP):

Definition 5.3. The **expected counterfactual difference sequential plan response (ECDSP)** of a dynamic causal system between interventions $\hat{x}^c = do(X_t = x_0^c, X_{t+1} = x_1^c, \ldots, X_{t+h} = x_h^c)$ and $\hat{x} = do(X_t = x_0, X_{t+1} = x_1, \ldots, X_{t+h} = x_h)$ is given by

$$ECDSP_Y^{x^c,x}(h) = E[Y^{x^c}_{t+h}(\Omega_{t-1}, E_{t,h}) - Y^{x}_{t+h}(\Omega_{t-1}, E_{t,h})]. \tag{5.14}$$

Figure 14 shows the ECDSP for $\Delta_h Z_t$ in model (5.9) for interventions $do(e_t = 1, e_{t+1} = 0, \ldots, e_{t+30} = 0)$ and $do(e_t = 0, e_{t+1} = 0, \ldots, e_{t+30} = 0)$. Notice that this figure does not show the oscillatory pattern present in Figure 12. This is so because setting e_{t+k} for all k from zero to 30 eliminates the autocorrelated structure of e_t, forcing all policy shocks to predetermined values.

[2]This can be easily validated through stationarity tests and also by the basic properties of random walks.

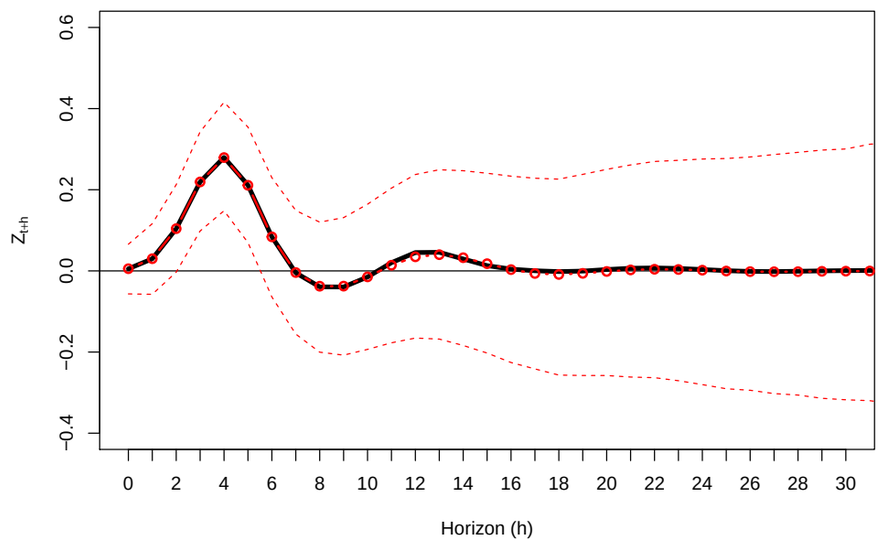

Figure 13: Thick black line is the EGCIRF for model (5.9) for $do(e_t = 1)$ against $do(e_t = 0)$. Red dotted dashed line is the mean of estimates based on Equation 5.13. Red dashed lines represent percentiles 2,5% and 97,5%. Estimates are based on 10,000 model simulations with sample size 1,000 each after a burn-in period of 100.

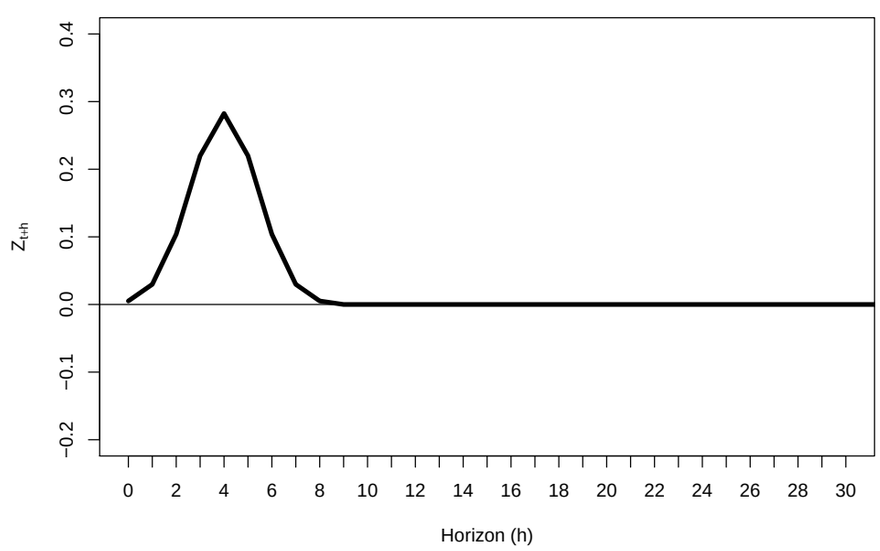

Figure 14: ECDSP over $\Delta_h Z_t$ for model (5.9) with intervention $do(e_t = 1, \ldots, e_{t+30} = 0)$ against $do(e_t = 0, \ldots, e_{t+30} = 0)$.

The estimation of this impulse response from observational data is more cumbersome than previous examples, but can be done with the help of Theorem 5.1 or, more generally, through the rules of do-calculus. Here we just apply Proposition 5.2, derived from Theorem 5.1, which fits nicely on linear models.

First, notice that, as before

$$ECDS\,P^{e^c,e}_{\Delta_h Z}(h) = E[\Delta_h Z_t | do(e^c)] - E[\Delta_h Z_t | do(e)], \tag{5.15}$$

where $e^c = (e_t = 1, e_{t+1} = 0, \ldots, e_{t+30} = 0)$ and $e = (e_t = 0, e_{t+1} = 0, \ldots, e_{t+30} = 0)$. The expected values on equation (5.15) can be estimated with the help of Proposition 5.2. It is easy to see from Figure 11 that conditions of Theorem 5.1 are fulfilled by all Z_{t+h} and e_{t+k}, and so the proposition applies. Results are shown in Figure 15.

The results are almost unbiased, the variance grows as before, and the oscillatory pattern is not present in the estimation, as expected.

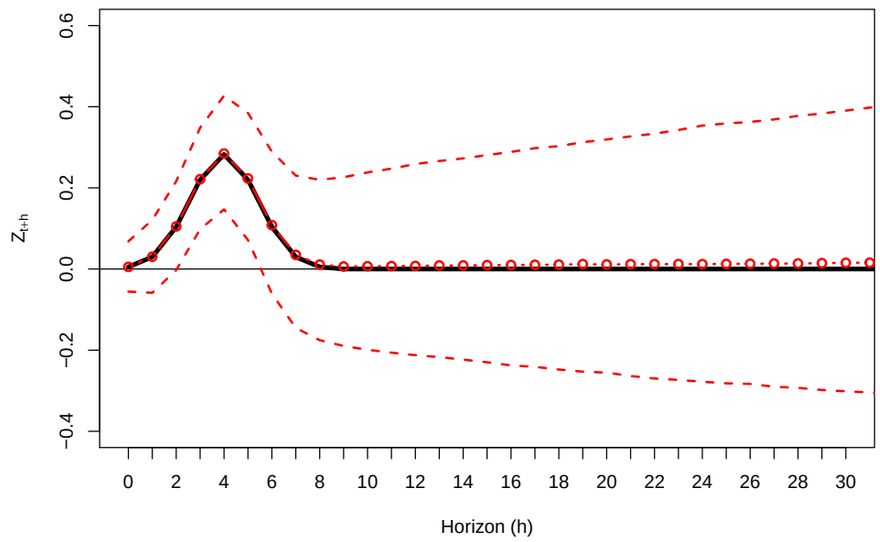

Figure 15: Thick black line is the ECDSP over $\Delta_h Z_t$ for model (5.9) with intervention $do(e_t = 1, \ldots, e_{t+30} = 0)$ against $do(e_t = 0, \ldots, e_{t+30} = 0)$. Red dotted dashed line is the mean of estimates based on Proposition 5.2. Red dashed lines represent percentiles 2,5% and 97,5%. Estimates are based on 10,000 model simulations with sample size 1,000 each after a burn-in period of 100.

6 Identifying causal effects of monetary policy in Brazil

Let's apply the concepts developed before in a real case. The analysis of monetary policy in Brazil, carried out by Banco Central do Brasil (BCB), the Brazilian Central Bank, is specially suited for the developed methods. This is so because the BCB adopted an inflation targeting regime with explicit target values for the consumer price inflation index (henceforth called IPCA) in mid-1999 (see Bogdanski et al. [2000]). The policy rate is the overnight interest rate (henceforth called Selic rate). The Selic target rate is chosen by the Monetary Policy Committee (henceforth called COPOM), comprised of the board of the BCB. The policy is carried out by open market operations, with market values of Selic rate maintained so close to the target Selic rate such that both may be considered the same. The inflation target with tolerance bands is determined by the Conselho Monetário Nacional (CMN), the National Monetary Council, for 3 years ahead, such that expectations are accommodated. So, an explicit manipulation of a variable (the overnight interest rate) is endured and I want to evaluate its causal effects over other variables.

To do so, I use the EGCIRF defined in Section 2. Control for confounding is essential in this context. Past inflation influences central bank decisions and also affects future inflation through inertial mechanisms, for example. Other potential confounders for inflation are expectations and real activity. Actually, several macroeconomic variables influence both, central bank decisions and future inflation, and are confounders for this problem.

Fortunately, central bankers are not omniscient beings, and must base their decisions on observable data and their own expert judgment. So, I take advantage of the inflation reports issued by the

BCB since the adoption of inflation targeting to take clues and data to control for direct causes of Selic rate changes, and so identify causal effects of these variations on economic aggregates.

An approach close to ours, in the spirit of causal analysis of monetary policy, is Angrist et al. [2018]. In this work the authors try to infer causal effects of monetary policy using US data, but since their goal is to evaluate non-linear effects, they use propensity score matching for inference. My approach resembles the one in Romer and Romer [2004]. Their method regresses intended changes on fed funds rate by the FOMC, inferred from the *Weekly Report of the Manager of Open Market Operations*, against internal forecasts of the Federal Reserve on inflation and real activity, extracted from the "Greenbook", and then take the residuals of this regression as a measure of monetary policy shocks. Their causal reasoning is clear: if we admit intended changes in fed funds rate as actions of FOMC and control these actions for confounders that are immediate causes of FOMC intentions, the residuals are, we hope, free of counfounding effects, being a suitable measure of exogenous monetary policy shocks.

Our goal is easier, since the Selic rate is strictly determined and its target publicly available, freeing us from inferring intended funds rate from reports, a potential source of inaccuracies. Moreover, we regress the response variable of interest, inflation for example, directly against the Selic rate and confounders composed of BCB forecasts and other relevant information available on inflation reports. These reports contain tables showing scenarios for inflation forecasts, and these values are used for inferences, as detailed in Tables 2 and 3, which contain a description of each regressor together with its data source.

The relation to Romer and Romer [2004] becomes clearer considering the following fact: regressing the response against residuals of the regression of Selic rate on confounders (these residuals are a measure of monetary policy shocks), like in Romer and Romer [2004], is the same as our approach (in linear models) by the Frisch-Waugh-Lovell theorem (see Frisch and Waugh [1933], Lovell [1963], Hayashi [2000]). Moreover, to do things in the way presented here facilitates regression diagnostics, confidence interval computations and also non-linear expansions in a more direct and general way than in the Romer and Romer [2004] approach.

The causal diagram representing the identification strategy is shown in Figure 16. The main idea is that whatever variables in economy affect response and central bank decisions, the latter is always done through some picture of reality mediated by observable data and possibly unobservable judgment. So, these are the parents of the Selic rate decision in our causal diagram. Assuming possible judgment bias remains constant on average along all sample period, conditioning on observable data used by central bankers for decision making is sufficient for causal effects identification by the back-door criterion. Also, regressions are done for each horizon h after meeting time t exactly like the various causal impulse responses presented in Section 2.

One word about judgment bias. They are represented by exogenous variables and/or functions f_i in the structural causal model (SCM) underlying the analysis. A systematic change in policy decision behavior represents a change on these mechanisms and is related to Lucas critique (Lucas [1976]). Here we will assume no systematic change in policy behavior both for simplicity and weak empirical evidence, but we recognize the possibility of distinct regimes.

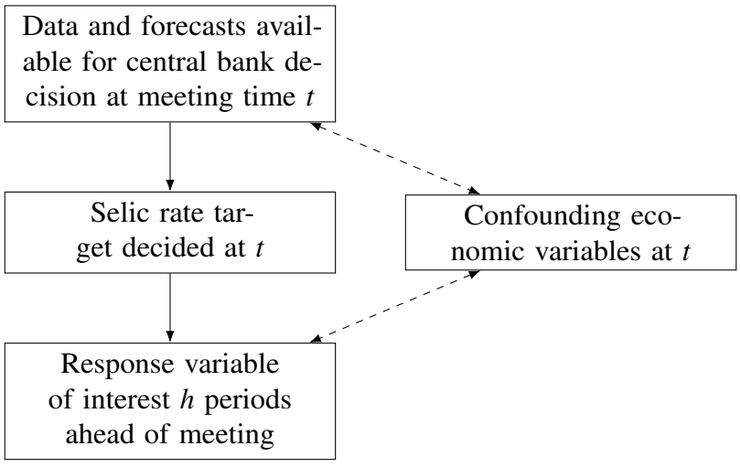

Figure 16: Graph representing manipulation of Selic rate by BCB.

6.1 Data

Before showing which series are used for inference, I explain how data is organized for estimations. The method regresses response variables h periods ahead of meeting at time t against policy instrument and available data for policy makers at t. Since meetings are not regularly spaced along time, some care must be taken. Suppose the response variable is Y_{t+h}, the policy variable at meeting time t is X_t, and data available for policy decision is C_t. Notice that C_t may contain variables realized before t, like the last available inflation index and its lags, for example. Then, for response h periods ahead, we regress Y_{t+h} against X_t and C_t. Notice that h is equally spaced over time, representing h months ahead, for example. It does not coincide, necessarily, with meeting frequency, which actually changes over time. Data were also interpolated to coincide with meeting days. More details in Appendix A.3.

That being said, I have chosen suitable variables extracted mainly from the BCB inflation reports, common macroeconomic aggregates and market projections collected by the Focus survey, which is regularly carried out by the BCB, compiling forecasts by market participants. Since I control for parents of Selic rate in Figure 16, all specifications are the same, independent of the response. So, after choosing which direct causes to control for, these same regressors may be used for the estimation of *any causal effect of Selic rate manipulation*, provided all premises of the causal model and adequate regularity conditions are fulfilled. The core specification for regressors is given in Tables 2 and 3, which show a symbol, a brief description and the data source for each regressor.

Variables in Tables 2 and 3 are not "cherry-picked" as might look at first glance. They are reminiscent of a selection process over a wider set of variables shown in Tables 7, 8 and 9, presented in Appendix A.5. I choose a set of meaningful observable determinants of the BCB decisions, compute the causal impulse responses of interest, and proceed to a process of variable exclusion in a way that minimally affects the causal impulse responses. The rationale behind this procedure is the strong correlations among regressors, because of redundant information carried by them. This correlation is not particularly important, since the only coefficient of interest is the one for policy variable, but less variables means more degrees of freedom and more precise or, at least, more reliable estimates. All coefficients but the policy variable are there just for conditioning, and almost all do not have any causal interpretation. This is natural, since it is fairly intuitive that the BCB internal forecasts, for example, do not have a *direct* causal impact on response variables such as inflation or GDP. Their causal impact is all mediated by the policy variable, as is clear from diagram in Figure 16[3].

Another important practical aspect is the stationarity of all variables involved. For a regression to represent an estimate of a conditional expectation, some regularity conditions must be fulfilled. And this is very important, because we may use OLS for almost any data we put in a computer, but if the conditional expectation does not exist, results are meaningless. Also, since I deal with time series data, ergodicity is a main concern (Hamilton [1994]), because this property guarantees temporal averages to converge to ensemble averages asymptotically in time dimension. So, non-stationary variables like GDP expectations have to be transformed, taking differences, for example. Taking only first differences monthly may not be sufficient, because policy makers may look for longer horizons, like annual, semiannual or quarterly variations. Also, annual variations may behave like non-stationary data even if it is stationary in the long run, because of short sample. So, I use some artifices on data differentiation and transformations to maintain stationarity without losing too much information in the process. These artifices are quite common in the literature, but may cause collinearity problems, and so I proceed to remove some collinear variables as described before. More details in Appendix A.5.

Finally, regression residuals are autocorrelated mainly because response variables are autocorrelated, a common feature in time series regression. This is unavoidable, since our regressions only have their causal meaning specified as we do. Certainly, for some response variable Y at $t + h$, if Y_{t+h-1} is controlled for, the regression fit would be better, but its causal meaning is lost. To see why, think about the causal path linking the policy variable S_t to Y_{t+h}. If this causal link exists, so probably a causal link from S_t to Y_{t+h-1} also exists, and controlling for this last variable would

[3]These observations highlight the main difference between the procedure presented and the traditional structural equations estimation carried over in econometrics. Here, regression equations are NOT structural. They represent *conditional expectations*, not mechanisms of variable determination. But, behind its justification there is a Structural Causal Model, pictured in a simplified way by the DAG in Figure 16. The economic mechanisms are hidden in the confounding economic variables at t, but I avoid modeling these mechanisms, because in the presented problem it is possible to block the back-door paths from policy variable to response variables controlling for direct causes of the policy variable. So, besides equations not being structural in the mechanistic sense, the coefficient of the policy variable *does have a causal meaning*, made clear by the SCM approach.

block the causal path $S_t \to Y_{t+h-1} \to Y_{t+h}$, invalidating our estimates of *total* causal effects. So, I use Newey-West HAC estimators to compute confidence intervals (see Newey and West [1986], Bierens [1996], Hamilton [1994]).

Regressor	Description	Source
S	Target Selic rate change. This is the policy variable, determined on each COPOM meeting. Its coefficient is the causal effect of interest.	BCB - Copom Meetings
R_1	12-month difference on target Selic rate	BCB - SGS
R_2	Last 6-month change of inflation index	BCB - SGS
R_3	Last 12-month change of inflation index	BCB - SGS
R_4	Median of Focus IPCA 12-month-ahead projections - seasonally adjusted	BCB - Market Expectations System
R_5	Difference between the reference scenario IPCA forecast by the BCB for meeting day and 12-month change in the IPCA	BCB - Inflation Reports
R_6	Difference between the reference scenario IPCA forecast by the BCB for 6 months ahead and the BCB forecast for 3 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_7	Difference between the reference scenario IPCA forecast by the BCB for 9 months ahead and the BCB forecast for 6 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_8	Difference between the reference scenario IPCA forecast by the BCB for 12 months ahead and the BCB forecast for 9 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_9	Monthly variation of median 12-month-ahead Focus Selic rate forecasts	BCB - Market Expectations System
R_{10}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for meeting day by Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{11}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for 3 quarters ahead by Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{12}	Real GDP 3-month log variation	BCB - SGS
R_{13}	Real exchange rate 6-month log variation - US dollar	BCB - SGS

Table 2: Regressors used in main estimates, part I.

Regressor	Description	Source
R_{14}	Focus end-of-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{15}	Focus 6-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{16}	Focus 12-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{17}	Monthly variation of Focus end-of-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{18}	Monthly variation of Focus 6-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{19}	Monthly variation of Focus 12-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{20}	Focus expected industrial production 6 months ahead minus 1 year industrial production variation	BCB - Market Expectations System
R_{21}	Focus expected industrial production 9-month ahead minus 1-year industrial production variation	BCB - Market Expectations System
R_{22}	Three-month difference of the reference scenario IPCA forecast by the BCB for meeting day	BCB - Inflation Reports
R_{23}	Three-month difference of the reference scenario IPCA forecast by the BCB for 3 months ahead	BCB - Inflation Reports
R_{24}	Three-month difference of the reference scenario IPCA forecast by the BCB for 12 months ahead	BCB - Inflation Reports
R_{25}	Difference between the market scenario IPCA forecast by the BCB for 12 months ahead and the same BCB forecast for 6 months ahead, both adjusted for inflation target	BCB - Inflation Reports

Table 3: Regressors used in main estimates, part II.

6.2 Main results

From Figure 16, it is clear that, when controlling for direct determinants of the Selic rate, its causal effects are identified by the back-door criterion. So, I proceed as in Section 3 and compute EGCIRFs for the following response variables: real GDP gap, 3-month Selic rate change (difference from 3 months ago), 12-month IPCA and real exchange rate Brazilian real per US dollar total percent variation from t to $t + h$, where h is the horizon of the EGCIRF. The policy variable is the Selic rate change (difference from the previous meeting), the counterfactual regimes are a 25-basis-point (0.25 *pp*) rise against no variation and the covariates chosen are those in Tables 2 and 3. The real GDP gap is computed by the HP filter in the usual way, and all variables are suitably transformed or interpolated to have values at meeting day t and monthly figures thereafter. I also add as regressors 4 lags of the policy variable and lags 2 to 4 of GDP gap for reasons that will become clear later.

Details on data transformations are in Appendix A.3. Regressions diagnostics and robustness checks are in Appendices A.5, A.6 and A.7. The EGCIRFs are shown in Figures 17, 18, 19 and 20.

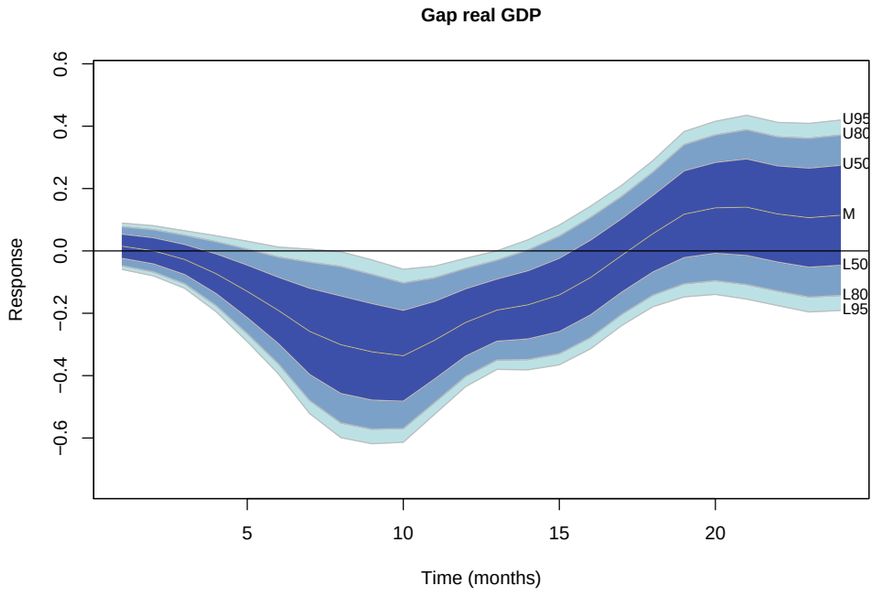

Figure 17: EGCIRF of real GDP gap for a 25-basis-point rise in the Selic rate against no variation. This response considers 4 lags of policy variable and lags 2 to 4 of GDP gap as confounders.

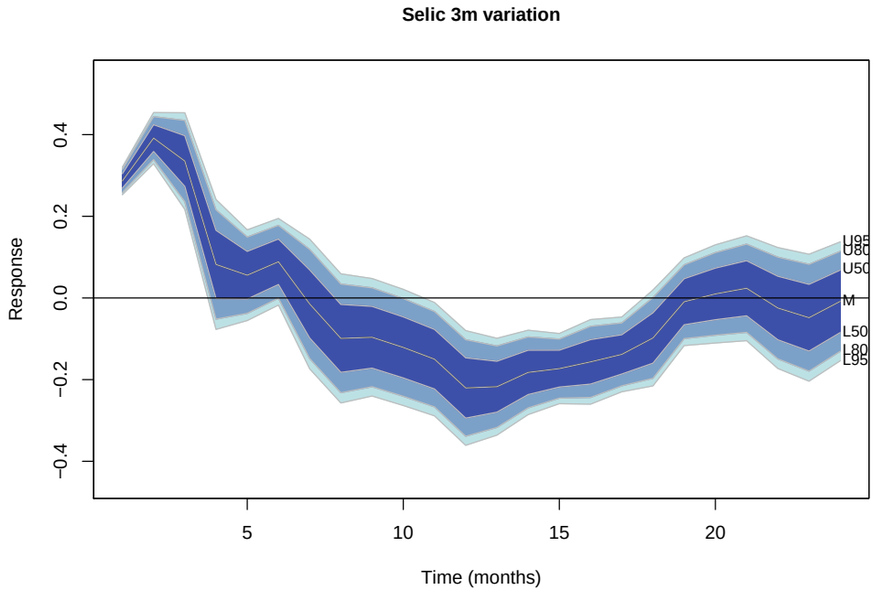

Figure 18: EGCIRF of a 3-month Selic rate change for a 25-basis-point rise in the Selic rate against no variation. This response considers 4 lags of policy variable and lags 2 to 4 of GDP gap as confounders.

These figures show the expected behavior from a theoretical viewpoint and also agrees to some degree with findings in the literature for Brazilian data. Notice that covariates are the same for

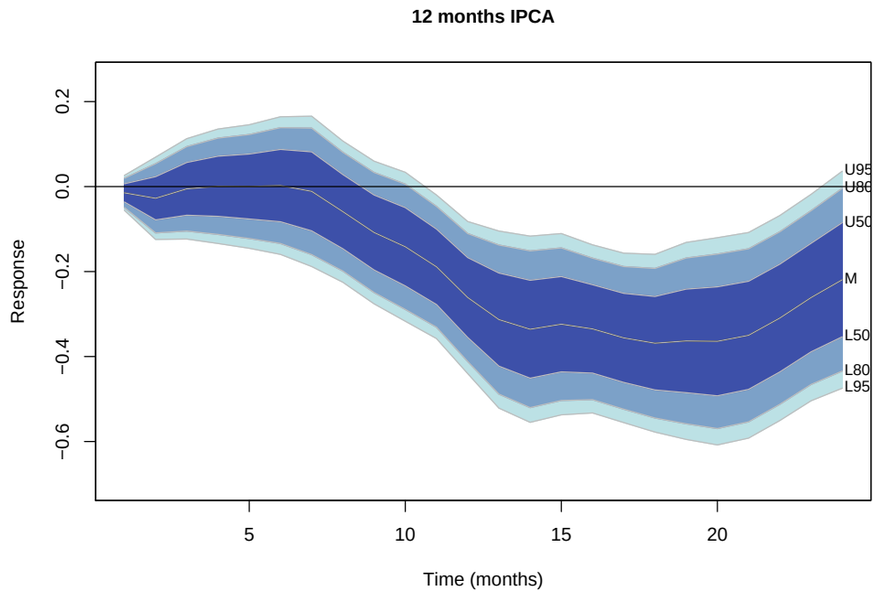

Figure 19: EGCIRF of 12-month IPCA change for a 25-basis-point rise in the Selic rate against no variation. This response considers 4 lags of policy variable and lags 2 to 4 of GDP gap as confounders.

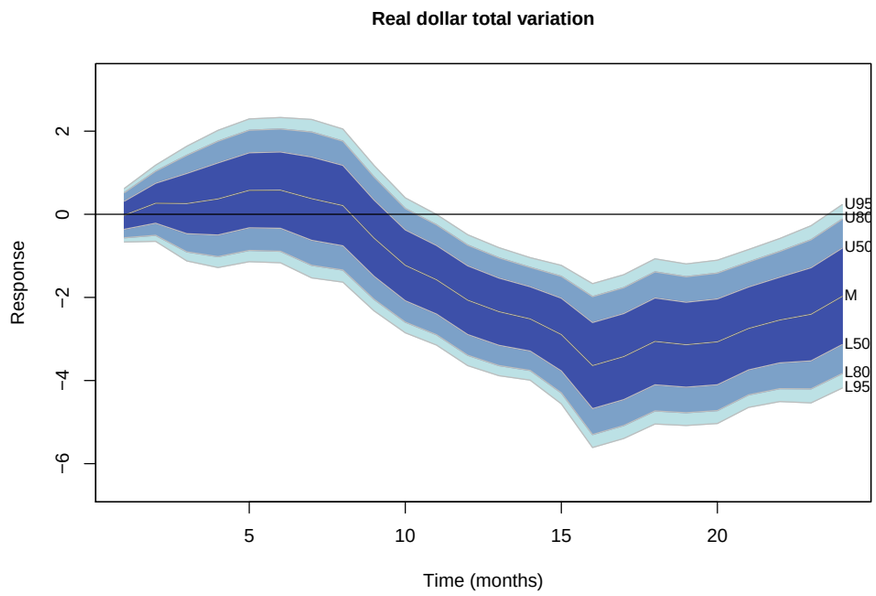

Figure 20: EGCIRF of real dollar total variation for a 25-basis-point rise in the Selic rate against no variation. This response considers 4 lags of policy variable and lags 2 to 4 of GDP gap as confounders.

all estimates, since only the response variable is changed. Lets analyze each of these EGCIRFs separately.

Real GDP gap

Figure 17 shows the response of output gap in percentage terms to a 25-basis-point increase in the Selic rate. As expected, this response is negative, reaching its minimum 10 months after the shock, with minimum expected value of -0.37%. Notice that the real GDP gap is hardly observable at t, especially the one computed with the HP filter, because this technique suffers from severe end point bias and various other problems (Hamilton [2018]). However, it is possible to add the gap at $t-2$ to $t-4$ as a proxy to the GDP gap at t, and so I did it in Figure 17. Without these covariates the EGCIRF becomes that of Figure 21.

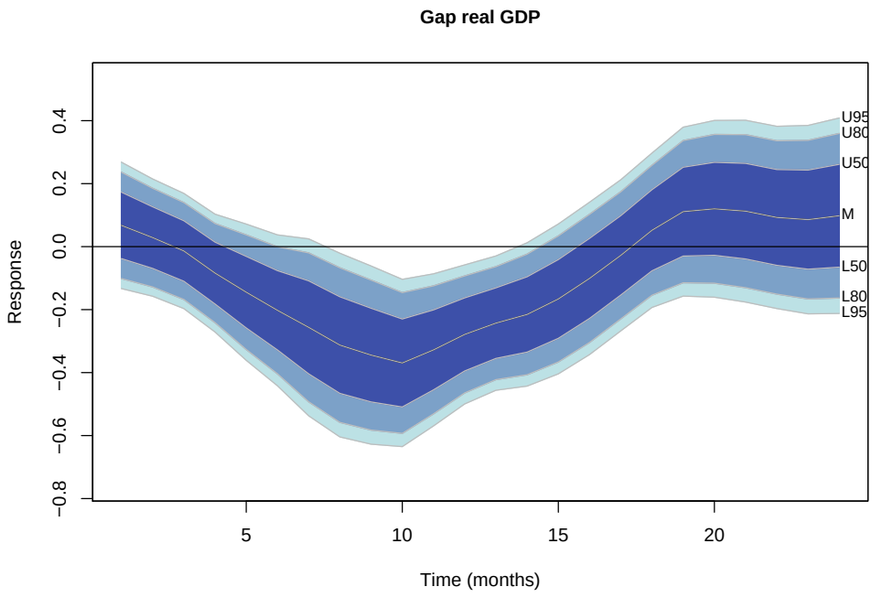

Figure 21: EGCIRF of real GDP gap for 25 basis points rise in Selic rate against no variation.

The effect of adding output gap as a covariate is to set closer to zero the effect at t and increase the estimate precision in some horizons close to the shock, having no visible effect on horizons beyond 6 months. Overall, the main conclusions remain intact. Also, the addition of output gap as a covariate had no effect in all other EGCIRFs presented.

The literature on GDP response to monetary policy shocks in Brazil is comprised mainly of estimations made through VARs, semi-structural models and DSGEs, as expected. Overall, results point to a negative effect over GDP by a contractionary shock, as our estimates, but the size of the effect and its timing vary.

Costa Filho [2017] uses monetary policy shocks inspired by Romer and Romer [2004] and includes these shocks in VARs to compute impulse responses. For GDP, his results point to a maximal effect of -0.5% reached before 5 months horizon, contrasting with our -0.37% in 10 months. But the majority of impulse responses presented in Costa Filho [2017] are not statistically significant, a common issue in VAR analysis, and other variables present some problems in some specifications, such as significant "price puzzles" on IPCA. The author argues that cost channels of monetary policy in Brazil may explain these findings. We recognize the possibility of these cost channels, but Figure 19 shows no price puzzle. Actually, all our specifications do not show significant price puzzles, but if some key confounders are not controlled for, price puzzles become strong and significant, showing that misspecification caused by inadequate control of confounding is a main concern on this issue.

Mendonça et al. [2010] used the SVAR with sign restrictions "agnostic" approach presented in Uhlig [2005] to estimate impulse responses. Their results point to 65% of probability of a negative response of GDP after a one-standard-deviation monetary shock, with 20% probability for negative variations between 0 and -0.5%. They also find a local minimum 10 months after the shock, but their confidence intervals are huge and this is the reason for their analysis to show results as probabilities, since they are not significantly different from zero. The agnostic approach is more free of assumptions, but also has its issues.

A semi-structural approach may be found in Minella and Souza-Sobrinho [2013]. Their results point to a maximum reduction in GDP around 3 quarters after a 25 basis points per quarter shock, and the size of the effect is around -0.2%. Overall, their results are close to ours, but the effects are weaker. There are not confidence bands for precision comparison.

A DSGE approach may be found in Castro et al. [2011]. This is the main DSGE model used by the BCB in monetary policy analysis, ingeniously named SAMBA (Stochastic Analytical Model with a Bayesian Approach). In the original form, their results point to a -0.25% minimal contraction of GDP after a monetary policy shock occurring after roughly 3 quarters, very close to our findings. More recently, in a revision of the SAMBA model (Special Study n.39/2019, issued in Inflation Report Boxes, March 2019, BCB), their GDP response is stronger, with minimum around -0.65% in a two-quarter horizon.

Results vary greatly, but the message is the same: monetary policy does cause real effects on GDP. We claim our approach is more easily interpretable, has less critical assumptions, and is more precise except for the DSGE approach, although this one is the most rigid in its assumptions, generating more debatable results. Anyway, I do not believe in silver bullets in science. For me, the approach presented is a clean way to access meaningful impulse responses for policy decisions and can be a reference for more complex models able to answer more complicated questions, although based on more stringent assumptions. In fact, causal impulse responses presented in this paper are a good way to validate these assumptions, contributing to the falsifiability of economic theories, something arguably essential for good science (see Popper [2005]).

3-month Selic rate change

The next response I analyze is the 3-month Selic rate change, the difference between current annualized Selic rate and the rate 3 months before, shown in Figure 18. It is very important for this work because here I am mainly focused on the methodology and not on economic analysis, and this EGCIRF may give some clues about the adequacy and interpretation of the results. A noticeable feature of Figure 18 is that, after a positive 25-basis-point hike in the Selic rate, subsequent shocks remain positive and after roughly seven months they change their sign. Remember from Section 2 that the EGCIRF is the response to an exogenous shock at t, but subsequent shocks are free to vary, different from the Sequential Plan Response ECDSP of Section 5, where we evaluate the response to a shock at t with subsequent shocks constrained to zero.

So, our results suggest a contractionary period of monetary policy held by the BCB is followed by a period of monetary expansion, much like the impulse response shown in Figure 12, where only an intervention at t is considered, leaving future shocks free to vary. This points to a conservative action of the BCB, where interest rates are raised excessively to drive inflation to its target, forcing a period of monetary expansion to accommodate the excess. This behavior predicts an oscillatory pattern in the Selic rate, which is indeed the case if we look for the historical Selic rate as shown in Figure 22.

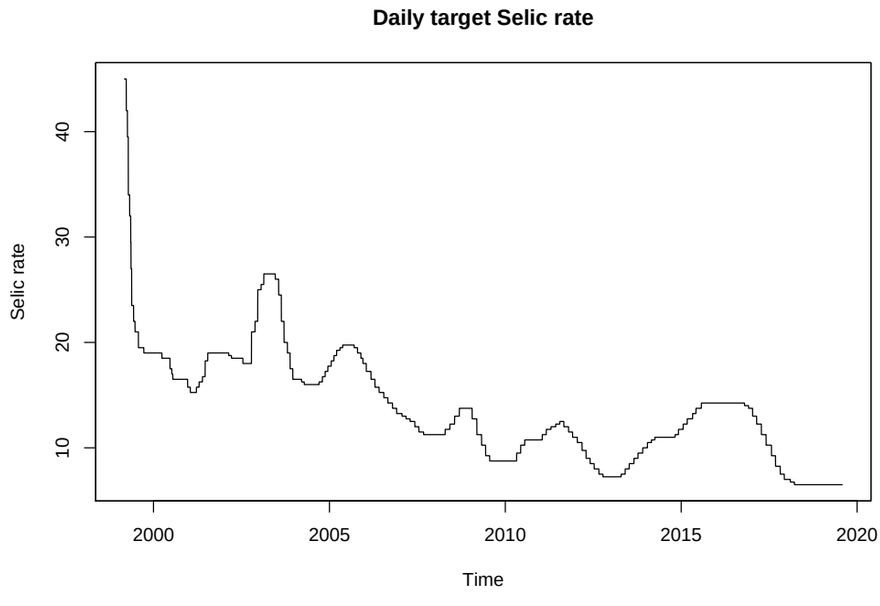

Figure 22: Daily target Selic rate.

Methodologically, an important issue is the response of the 3-month Selic change after 1 month. This value is roughly 25 basis points according to Figure 18, as expected. But if we do not use lags of the Selic rate change as regressors, the result is the one in Figure 23.

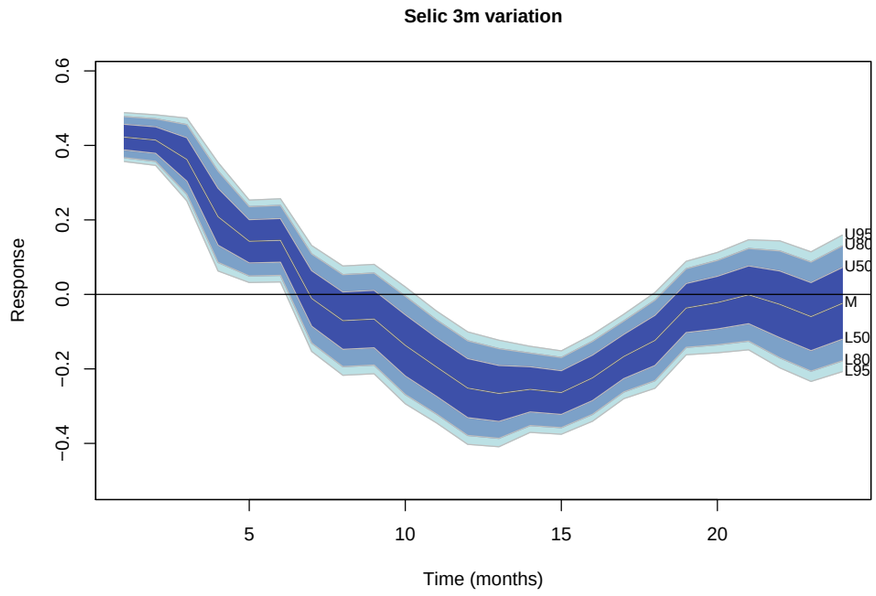

Figure 23: EGCIRF of 3-month Selic rate change for a 25-basis-point rise in Selic rate against no variation.

Notice that the response after 1 month is roughly 40 basis points in this case, with remaining periods very close to our main result. This suggests that correlations between policy decision and the already accumulated Selic rate before the meeting are being accounted for in the impulse response, and so it is not causal, because previous Selic rate variations are acting as confounders. Indeed, Selic rate variation between meetings is strongly autocorrelated. Part of this autocorrelation may be explained by the autocorrelation of variables that determine policy decisions, but the present analysis point to a residual "pure" autocorrelation that must be controlled for. We can achieve this by adding lags of the policy variable (Selic rate change). Adding 4 or more lags leads to identical results, and so we choose to add 4 lags.

This autocorrelation in the Selic rate change explainable by more than the autocorrelation of its other determinants suggests policy makers choose to slowly vary the Selic rate. Moreover, this slow variation is carried out for no other reason than to simply make things in a cautious way, acting in cycles of monetary contraction and expansion, clearly visible in Figure 22. This is reasonable and corroborates with human behavior under uncertainty. Comparing with results in the literature, the most noticeable feature is a similar behavior from the Selic response found in Minella and Souza-Sobrinho [2013] in terms of timing and intensity of the effect.

12-month IPCA change

Figure 19 shows the EGCIRF for 12-month IPCA change after a 25-basis-point rise in the Selic rate against no variation. Some interesting features of this result are: no significant price puzzle and inflation reduction is stronger *after* real GDP contraction, corroborating conventional wisdom. Also, the maximum effect occurs at roughly 15 to 20 months and has an expected value around -0.42%.

Costa Filho [2017] showed price puzzle for almost all specifications, something we attribute to poor control of confounding. Actually, the traditional price puzzle "solution", as explained by the author, is to control for commodity prices hoping to anticipate future inflation, because it is believed that the puzzle is caused by anticipation of future inflation by the central bank, which raises interest rates before inflation rises. Well, in terms of causal models this is typical confounding. What happens is that there are some variables which cause more inflation in the future, and looking for those variables, central bankers anticipate inflation hike raising interest rates. Controlling for those variables or its consequences for central bankers, like forecasts indicating more inflation, hopefully vanishes the puzzle. Results presented point to this conclusion.

Mendonça et al. [2010] point to a 35% probability of a -0,10% fall in IPCA six months after the shock. But, as before, their estimates lack precision, with most results statistically non-significant.

Minella and Souza-Sobrinho [2013] point to a roughly -0.6% IPCA fall 5 to 6 quarters after shock, and Castro et al. [2011] point to approximately -0.25% IPCA fall after 4 quarters. The 2019

updated SAMBA version, mentioned before, points to a roughly -0.5% IPCA fall after 4 quarters.

These results are less variable, and our figures are relatively close to Minella and Souza-Sobrinho [2013], with significant stronger effects over inflation on a longer horizon than other estimates.

Real exchange rate total variation

This is the response of real exchange rate variation from t to $t + h$ between currencies Brazilian real and US dollar, measured in reais per dollar adjusted for respective inflation indexes. It is shown in Figure 20. Notice that the appreciation effect is delayed months ahead, something probably related to the delayed effect of the Selic rate over IPCA. Indeed, their timing approximately coincide, as can be seen comparing Figures 19 and 20. However, some care must be taken in this comparison, because the IPCA is measured as 12-month variation, and the real exchange rate is total accumulated since shock. The maximum real appreciation of domestic currency occurs around 16 months after shock and corresponds to a US dollar real total depreciation of -3.9% on average.

The reaction of the real exchange rate to monetary contractions is a subject of intense debate in the literature, and puzzles are found for developed and emergent economies. Kohlscheen [2014] discusses some of these issues. Zettelmeyer [2004] studies the impact of monetary policy shocks on exchange rate of some small open economies, namely, Australia, Canada and New Zealand during the 1990s. His findings point to a 2-3% appreciation of the exchange rate after a 100 basis points hike in 3-month interest rates. Kohlscheen [2014] analyzes Brazil, Mexico and Chile, and looking for 1 day after policy shock the author does not find support for the predictions of standard small open economies models of an appreciation of the exchange rate. Our findings also do not find short-run effects over real exchange rate, and even show some depreciation effect, although not statistically significant at 5% level. But, on longer horizons the appreciation is clear and significant, as shown in Figure 20.

In more model dependent approaches, Costa Filho [2017] shows some weak evidence in favor of appreciation. Mendonça et al. [2010] found puzzling but non-significant effects. Minella and Souza-Sobrinho [2013] point to a maximum 1.75% appreciation in 3 quarters, while Castro et al. [2011] show immediate effects of about 1% real appreciation of domestic currency.

Results on real exchange rate are less consistent in the literature, and our approach may help shed some light on these questions. A point favoring our results is its consistency among different responses. Remember that regressors are the same for all estimates, and others results besides real exchange rate are less controversial and somewhat appealing. Of course that confounders relevant for the real exchange rate may be missing, since internal forecasts of the BCB used in our analysis seek only inflation. But, market expectations about the exchange rates are considered, and these are probably contemplated by the BCB policy makers. Indeed, the BCB forecasts contain scenarios considering the Focus survey expected exchange rate, and I control for these scenarios.

7 Conclusion

In this work I present, in Section 2, causal definitions of impulse responses for time series based on structural causal models. These are counterfactual impulse response function (CIRF), generalized counterfactual impulse response function (GCIRF), expected generalized counterfactual impulse response function (EGCIRF), causal generalized impulse response (CGI) and expected causal generalized impulse response (ECGI). In Section 5, the expected counterfactual difference sequential plan response (ECDSP) is also presented. Each one has its applications, and additional definitions may be given based on the same ideas.

In Sections 3 and 4 I apply some of these definitions on simulated data through Monte Carlo experiments, showing how the ideas work well for a large class of causal dynamical systems. Results in these sections are presented to persuade the reader to conclude about the power and simplicity of given definitions.

In Section 5 the ECDSP is presented together with an application on simulated data and detailed explanation of its usefulness and interpretation. Results showing how endogenous policy variables may be analyzed with different perspectives serve to show how flexible the techniques are. Also, Proposition 5.2 gives a formula for the ECDSP in G-identifiable systems when linear approximations for conditional expectations are valid.

In Section 6, the theoretical concepts developed are applied on real data from the BCB, the Brazilian Central Bank. Results corroborate theoretical predictions, are relatively aligned with

previous findings in the literature and show better precision and reliability. The findings are also robust (see Appendices A.5, A.6 and A.7) and open the doors to more sophisticated approaches.

Additionally, the analysis shades some light on controversial issues, namely, price puzzles and the response of real exchange rates to monetary policy. The former is recurrent in econometric analysis probably because of poor control for confounding in inflation impulse responses estimates. The latter is "solved" in favor of an appreciation of domestic currency after a contractionary monetary shock, but delayed by several months, approximately following inflation reduction.

In these practical applications there is room for extensions. Possible ones are to examine non-linear effects and consider other confounders not granted in the analysis performed. My goal was to make things as simple as possible, with a direct application of techniques developed in Section 2 to show their usefulness in practical problems.

Conceivable improvements in estimates are a better shrinking of covariates using LASSO (Tibshirani [1996]), ridge regression (Saleh et al. [2019]), principal components regression (Amemiya [1985]) or partial least squares (Haenlein and Kaplan [2004]), for example. Also, it is possible to translate the estimation of conditional distributions into prediction problems through what Angrist and Pischke [2008] call the CEF decomposition property and CEF prediction property.

Since identifiable causal effects can be asserted as expressions composed of conditional distributions by the SCM approach, a full set of modern prediction techniques in machine learning may be applied to causal inference, despite widespread beliefs that these techniques can only express correlations (see Mullainathan and Spiess [2017]). Indeed, the SCM approach is all about expressing causal effects by means of correlations inferable from observable data.

References

T. Amemiya. *Advanced econometrics*. Harvard university press, 1985.

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press, 2008.

J. D. Angrist, Ò. Jordà, and G. M. Kuersteiner. Semiparametric estimates of monetary policy effects: string theory revisited. *Journal of Business & Economic Statistics*, 36(3):371–387, 2018.

A. Banerjee, J. J. Dolado, J. W. Galbraith, D. Hendry, et al. Co-integration, error correction, and the econometric analysis of non-stationary data. *OUP Catalogue*, 1993.

H. J. Bierens. *Topics in advanced econometrics: estimation, testing, and specification of cross-section and time series models*. Cambridge university press, 1996.

G. Blondel, M. Arias, and R. Gavaldà. Identifiability and transportability in dynamic causal networks. *International journal of data science and analytics*, 3(2):131–147, 2017.

J. Bogdanski, A. A. Tombini, and S. R. d. C. Werlang. Implementing inflation targeting in brazil. *Werlang, Sergio R., Implementing Inflation Targeting in Brazil (July 2000). Banco Central do Brasil Working Paper*, (1), 2000.

M. Castro, S. Gouvea, A. Minella, R. Santos, and N. S. Souza-Sobrinho. Stochastic analytical model with a bayesian approach. banco central do brasil. Technical report, Working Papers Series, 2011.

B. Céspedes, E. Lima, and A. Maka. Monetary policy, inflation and the level of economic activity in brazil after the real plan: stylized facts from svar models. *Revista Brasileira de Economia*, 62 (2):123–160, 2008.

K. Chalak and H. White. An extended class of instrumental variables for the estimation of causal effects. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(1):1–51, 2011.

S. Chatterjee and J. S. Simonoff. *Handbook of regression analysis*, volume 5. John Wiley & Sons, 2013.

C. F. Christ. The Cowles Commission's contributions to econometrics at Chicago, 1939-1955. *Journal of Economic Literature*, 32(1):30, 1994.

A. E. d. Costa Filho. Monetary policy in brazil: Evidence from new measures of monetary shocks. *Estudos Econômicos (São Paulo)*, 47(2):295–328, 2017.

S. Cunningham. Causal inference: The mixtape. *Unpublished manuscript*, 2018.

J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. iii. *Biometrika*, 58(1):1–19, 1971.

R. F. Engle, D. F. Hendry, and J.-F. Richard. Exogeneity. *Econometrica: Journal of the Econometric Society*, pages 277–304, 1983.

R. Frisch and F. V. Waugh. Partial time regressions as compared with individual trends. *Econometrica: Journal of the Econometric Society*, pages 387–401, 1933.

R. D. Gill and J. M. Robins. Causal inference for complex longitudinal data: the continuous case. *Annals of Statistics*, pages 1785–1811, 2001.

W. H. Greene. Econometric analysis fifth edition. *New Yersey: Prentice Hall*, 2003.

M. Haenlein and A. M. Kaplan. A beginner's guide to partial least squares analysis. *Understanding statistics*, 3(4):283–297, 2004.

J. Y. Halpern. Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12: 317–337, 2000.

J. D. Hamilton. *Time Series Analysis*. Princeton University Press, 1 edition, Jan. 1994. ISBN 0691042896.

J. D. Hamilton. Why you should never use the hodrick-prescott filter. *Review of Economics and Statistics*, 100(5):831–843, 2018.

T. Hastie and R. Tibshirani. Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386, 1987.

T. J. Hastie. Generalized additive models. In *Statistical models in S*, pages 249–307. Routledge, 2017.

F. Hayashi. *Econometrics*. 2000.

J. Heckman and R. Pinto. Causal analysis after Haavelmo. *Econometric Theory*, 31(1):115–151, 2015.

P. Hünermund and E. Bareinboim. Causal inference and data-fusion in econometrics. *arXiv preprint arXiv:1912.09104*, 2019.

G. W. Imbens. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature*, 58(4):1129–79, 2020.

Ò. Jordà. Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182, 2005.

H. Kiiveri, T. P. Speed, and J. B. Carlin. Recursive causal models. *Journal of the australian Mathematical Society*, 36(1):30–52, 1984.

H. Kincaid. Making progress on causal inference in economics. *A modern guide to philosophy of economics. Cheltenham, UK: Edward Elgar*, 2021.

E. Kohlscheen. The impact of monetary policy on the exchange rate: A high frequency exchange rate puzzle in emerging economies. *Journal of International Money and Finance*, 44:69–96, 2014.

G. Koop, M. H. Pesaran, and S. M. Potter. Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147, 1996.

T. C. Koopmans. When is an equation system complete for statistical purposes? In *General-to-specific modelling*, pages 304–320. Edward Elgar, 2005.

D. Kwiatkowski, P. C. Phillips, P. Schmidt, Y. Shin, et al. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of econometrics*, 54(1-3):159–178, 1992.

S. Lauritzen. Lectures on Contingency Tables, Aalborg, Denmark, 1982.

E. E. Leamer. Vector autoregressions for causal inference? In *Carnegie-rochester conference series on Public Policy*, volume 22, pages 255–304. North-Holland, 1985.

M. C. Lovell. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, 58(304):993–1010, 1963.

R. Lucas. Econometric policy evaluation: A critique. In *Carnegie-Rochester Conference Series on Public Policy*, volume 1, pages 19–46, 1976.

M. J. C. Mendonça, L. Medrano, and A. Sachsida. Efeitos da política monetária na economia brasileira: resultados de um procedimento de identificação agnóstica. 2010.

A. Minella and N. F. Souza-Sobrinho. Monetary policy channels in brazil through the lens of a semi-structural model. *Economic Modelling*, 30:405–419, 2013.

A. Moneta and F. Russo. Causal models and evidential pluralism in econometrics. *Journal of Economic Methodology*, 21(1):54–76, 2014.

S. Mullainathan and J. Spiess. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106, 2017.

J. Neter, W. Wasserman, and M. H. Kutner. Applied linear regression models. 1989.

W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Technical report, National Bureau of Economic Research, 1986.

J. Pearl. Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings of the 7th Conference of the Cognitive Science Society, 1985*, pages 329–334, 1985.

J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009. ISBN 052189560X, 9780521895606.

J. Pearl. Reflections on heckman and pinto's causal analysis after haavelmo. Technical report, CALIFORNIA UNIV LOS ANGELES DEPT OF COMPUTER SCIENCE, 2013.

J. Pearl. Trygve Haavelmo and the emergence of causal calculus. *Econometric Theory*, 31(1): 152–179, 2015.

J. Pearl and J. Robins. Probabilistic evaluation of sequential plans from causal models with hidden variables," from: Uncertainty in artificial intelligence: Proceedings of the eleventh conference on artificial intelligence, august 18-20, 1995, mcgill university, montreal, quebec, canada, 1995.

J. Pearl, M. Glymour, and N. P. Jewell. *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

P. Perron. Trends and random walks in macroeconomic time series: Further evidence from a new approach. *Journal of economic dynamics and control*, 12(2-3):297–332, 1988.

K. Popper. *The logic of scientific discovery*. Routledge, 2005.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL `https://www.R-project.org/`.

J. M. Robins. Causal inference from complex longitudinal data. In *Latent variable modeling and applications to causality*, pages 69–117. Springer, 1997.

C. D. Romer and D. H. Romer. A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084, 2004.

S. E. Said and D. A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

A. M. E. Saleh, M. Arashi, and B. G. Kibria. *Theory of ridge regression estimation with applications*, volume 285. John Wiley & Sons, 2019.

S. Sheather. *A modern approach to regression with R*. Springer Science & Business Media, 2009.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

A. Tversky and D. Kahneman. Causal schemas in judgments under uncertainty. *Progress in social psychology*, 1:49–72, 1980.

H. Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419, 2005.

N. Wermuth and S. L. Lauritzen. *Graphical and recursive models for contingency tables*. Institut for Elektroniske Systemer, Aalborg Universitetscenter, 1982.

H. White and K. Chalak. A unified framework for defining and identifying causal effects. *UCSD Department of Economics Discussion Paper*, 2006.

S. N. Wood. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC, 2017.

J. Zettelmeyer. The impact of monetary policy on the exchange rate: evidence from three small open economies. *Journal of Monetary Economics*, 51(3):635–652, 2004.

A Appendices

A.1 Structural Causal Models

The Structural Causal Models (SCM) approach detailed in Pearl [2009] is under development since the 1980's, having been born from Bayesian networks (Lauritzen [1982], Wermuth and Lauritzen [1982], Kiiveri et al. [1984], Pearl [1985]), and draws its inspiration mainly from Haavelmo's concept of causal intervention (Pearl [2015]). This approach rests on a philosophic viewpoint close to a Laplacian quasi-deterministic concept of causality (Pearl [2009]), where variables in a system relates to each other through *deterministic* equations with stochastic inputs.

SCMs are the workhorse of this paper and are designed to address causal questions, especially those related with manipulation of variables. Here I briefly describe them and also make some conceptual considerations relative to notions in statistics and econometrics which are somewhat controversial, in order to facilitate the understanding for those unfamiliar with SCMs. This appendix is based in Pearl [2009].

Formally, SCMs are defined as:

Definition A.1 (Pearl [2009])**.** A structural causal model is a triple $M = \langle U, V, F \rangle$ where:

1. U is a set of **background**, **predetermined** or **exogenous** variables, determined by factors outside the model;

2. V is a set $\{V_1, \ldots, V_n\}$ of **endogenous** variables, that is, they are determined by variables in $U \cup V$;

3. F is a set of functions $\{f_1, \ldots, f_n\}$ such that each f_i is a mapping from its domain in $U_i \cup PA_i$ to V_i, where $U_i \subseteq U$ and $PA_i \subseteq V \setminus V_i$ and the entire set F forms a mapping from U to V.

Since F is a mapping from U to V, for each instantiation u of U the system F has a unique solution $V(u)$. Every causal model M has an associated graph $G(M)$ in which each node corresponds to a variable and the directed edges point from members of PA_i and U_i toward V_i. If $G(M)$ is acyclic, then M is acyclic and uniqueness of $V(u)$ is guaranteed. Halpern [2000] also extends axiomatization of causal models to "arbitrary theories", where equations may have multiple solutions or no solutions at all. For our purposes, the acyclic case is sufficient.

We are interested in statistical analysis of causal phenomena, so a probabilistic extension of Definition A.1 is necessary. A probabilistic causal model is a pair $\langle M, P(u) \rangle$ where M is a causal model and $P(u)$ is a probability function defined over U. The operation of intervention on a variable X_i is denoted $do(X_i = x)$, and is done removing the equations $x_i = f_i(pa_i, u_i)$ and substituting all instances of X_i in the model by x. This leads to the concept of causal effect:

Definition A.2 (Pearl [2009])**.** Given two disjoint sets X and Y, the causal effect of X on Y, denoted as $P(y|do(X = x))$, $P(y|do(x))$ or, more compactly, $P(y|\hat{x})$, is a function from X to the space of probability distributions on Y. For each realization x of X, $P(y|\hat{x})$ gives the probability (or density) of $Y = y$ induced by deleting from the model $x_i = f_i(pa_i, u_i)$ all equations corresponding to variables in X and substituting $X = x$ in the remaining equations.[4]

The potential response of Y in model $M = \langle U, V, F \rangle$ to action $do(X = x)$ in situation $U = u$, denoted $Y^x(u)$, is the solution for Y of the system of equations $F_x = \{f_i : V_i \notin X\} \cup \{X = x\}$ where $f_i \in F$ and $U = u$ (Pearl [2009]). This is just a formal statement of the equation substitution mechanism with fixed background $U = u$, but helps to define the important concept of counterfactual:

Definition A.3 (Pearl [2009])**.** Let X and Y be subsets of V in causal model $M = \langle U, V, F \rangle$. The **counterfactual sentence** "Y would be y in situation $U = u$, had X been x" is interpreted as $Y^x(u) = y$, the potential response of Y to $X = x$ in situation u.

Letting u vary according to $P(u)$, counterfactuals may be seen as random variables Y^x, and a simple direct relation among counterfactuals and the "do" operator (Pearl et al. [2016]) is

$$P(Y^x = y) = P(Y = y|do(X = x)). \tag{A.1}$$

Besides their close relation, counterfactuals and do-expressions are not the same. Counterfactuals express relations among variables that evolves in *different histories*, and do-expressions represent

[4]We will denote the intervention $do(X = x)$ also as $\hat{x}$ or $do(x)$ without distinction, only for notational convenience.

interventions in probability distributions of the system. This issue and the greater flexibility of counterfactuals are discussed in Pearl et al. [2016].

Another important property of causal models is the directionality of causality. As observed in Pearl [2009], if the equation

$$y = \beta x + \epsilon \tag{A.2}$$

is regarded as structural, so the change in $E(Y)$ resulting from a unit change in X is β. If we rewrite (A.2) as

$$x = \frac{y - \epsilon}{\beta} \tag{A.3}$$

what is the interpretation of $1/\beta$? According to Pearl [2009], some authors simply deny any causal interpretation to β, concentrating on its statistical interpretation. In contrast, the econometrics standard response is that the error term in (A.3) is correlated with y, since it is not correlated with x by hypothesis in equation (A.2). But, as is well known, any pair of bivariate normal variables can be written as

$$\begin{aligned} y &= \beta x + \epsilon_y \\ x &= \alpha y + \epsilon_x \end{aligned} \tag{A.4}$$

with $Cov(X, \epsilon_x) = Cov(Y, \epsilon_y) = 0$. So, the condition $Cov(X, \epsilon_x) = 0$ does not guarantee that a variation of Y by one unit *cause* a variation of α in $E(X)$, since Equations (A.4) are not necessarily structural (causal). Notice that, observationally, these equations are true, but what if Y is manipulated? To think clearly about this, a simple example is helpful. Consider the following system:

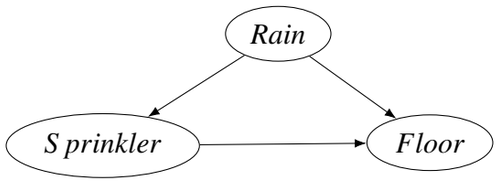

Figure 24: DAG representing the sprinkler system described on text.

This model has only 3 variables: one tell if it is raining, another if the sprinkler is on or off and the last one represents the state of the floor (wet or dry). Suppose that the sprinkler is automatic, and turns on if weather is dry enough. The arrows represent causal relations. In the $sprinkler \rightarrow floor$ link, turn the sprinkler on wets the floor, but not the other way around. So, there is an intrinsic *asymmetry* in causal relations which is somewhat masked by the equal sign in algebraic equations like (A.2), but is made clear by the arrow direction in causal graphs. One way to tackle these inconsistencies is through an *operational definition*[5] of structural equations (Pearl [2009]):

Definition A.4 (Pearl [2009])**.** An equation like (A.2) is said to be **structural** if in an ideal experiment where we control for $X = x$ and any other set of variables $Z = z$, Z not containing X and Y, then, the value of Y is given by $y = \beta x + \epsilon$, where ϵ is not a function of the settings x and z.

Consequently, an operational definition for structural parameters like β in equation (A.2) is simply (Pearl [2009]):

$$\beta = \frac{\partial E[y|do(x)]}{\partial x}. \tag{A.5}$$

These considerations also provide an operational definition for the structural error term (Pearl [2009]):

$$\epsilon = y - E[y|do(x)]. \tag{A.6}$$

This gives a definition to error terms that prescribes how they can be measured, avoiding worrying about how they originate. This contrasts with the widespread econometrics point of view.

One point about Definition A.4 is that it prescribes how measures must be made to observe structural parameters of interest, but it does not mean that these measurements are possible. Control for $X = x$ may be impossible in practice. But, the point of causal analysis of observational data is that, even if this hypothetical experiment is impossible, there are tools to transform equation (A.5), in some circumstances, into expressions free of do-operator, that is, expressions representing statistical models that may be fitted.

[5] An operational definition is a procedural way to define something, that is, a procedure to compute what is being defined.

All in all, a causal model is a set of non-parametric structural equations representing autonomous mechanisms of variables determination that can also be represented (in the acyclic case) by a DAG (directed acyclic graph). An intervention in X is done as substitutions of some functions in the model, or as removing all parents of X in the corresponding DAG (Pearl [2009]). Indeed, the DAG is more than a simple representation. Frequently operations are easily done in DAGs than equations. Some of these operations may be used to identify causal effects using only observational data. DAGs also make the asymmetric nature of causality clear.

Counterfactuals are the response of variables of the model to interventions in a given situation. The situation, represented by $U = u$, is interpreted as an individual or experimental unit in Neyman-Rubin approach[6], which is a particular case of Definition A.3.

Another related concept, which pervades econometrics and traces its origins to the Cowles Commission for Research in Economics (Christ [1994]), is that of exogeneity (Koopmans [2005], Engle et al. [1983], Leamer [1985]). Pearl [2009] contextualizes it and the related concepts of weak, strong and superexogeneity into the framework of SCMs together with instrumental variables and counterfactuals, showing the power of SCMs as a language for the analysis of causality in general.

Pearl [2009] gives some compelling reasons for the use of causal graphs and structural equations, at least as a common language for researchers on causality. I wish to highlight three main points in favor of his view: first, SCMs are non-parametric, and so more general, a desirable property for a broad language. Second, the Neyman-Rubin potential outcome framework and SEMs, which have their own merits, are embedded in SCM framework (Pearl [2009])[7]. Third, causal thinking is more natural to human mind than conditional independence assumptions, the basis for Neyman-Rubin framework.

Corroborating this third point, Tversky and Kahneman [1980] analyze various experiments investigating judgments about conditional probabilities. In their words:

> It is a psychological commonplace that people strive to achieve a coherent interpretation of the events that surround them, and that the organization of events by schemas of cause-effect relations serves to achieve this goal.

In their analysis of judgments about conditional probabilities $P(X|D)$ of a target X given evidence D, Tversky and Kahneman [1980] distinguish different types of relations between X and D. If D is perceived as a cause of X, then D is a *causal datum*. Otherwise, if X is a possible cause of D, the latter is called a *diagnostic datum*. Finally, D is *indicational* if X and D have a common cause and if they are not related, D is called *incidental*. The authors present evidence in support of causal inferences having greater efficacy than diagnostic inferences. The implication is that model construction based on causal relations is more natural and less error prone than model construction based on conditional independence assumptions.

Notice that the distinctions made in Tversky and Kahneman [1980] about the role of D are non-statistical, and the conditional probability $P(X|D)$ is the same for collected data if D is causal, diagnostic, indicational or incidental. Pearl [2009] argues that much confusion in econometrics and statistics are related to an implicit assumption of $P(X|D)$ as representing a causal relationship from D to X (see, e.g., Simpson's paradox in Pearl et al. [2016] or Pearl [2009]). This is false in general, but may be true in some circumstances, something that have an important role in identification of causal effects using observational data.

[6] In our notation we put the "experimental condition" $X = x$ in superscript instead of subscript, as is common in Neyman-Rubin framework, because we leave the subscript for time indexing.

[7] Neyman-Rubin is embedded through what Pearl et al. [2016] called the *fundamental law of counterfactuals*. SEMs are a particular case of SCMs.

A.2 Sequential plan response proof

Proposition A.5. *The expected causal effect in* (5.4) *may be approximated by the equation*

$$\begin{aligned}E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \alpha^{y_h} &+ \sum_{m=2}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\alpha(m-1)^{m,n} \\ &+ \sum_{j=1}^{l_1}\left[\beta_{1,j}^{y_h} + \sum_{m=2}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\beta(m-1)_{1,j}^{m,n}\right]E(w_{1,j}) \\ &+ \sum_{i=1}^{h-1}\left[\gamma_i^{y_h} + \sum_{m=i+1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\gamma(m-1)_i^{m,n}\right]x_i + \gamma_h^{y_h}x_h,\end{aligned}$$

where α^{y_h}, $\beta_{m,n}^{y_h}$ *e* $\gamma_m^{y_h}$ *are given by*

$$E(y_h|w_{1,1},\ldots,w_{h,l_h},x_1,\ldots,x_h) = \alpha^{y_h} + \sum_{m=1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}w_{m,n} + \sum_{m=1}^{h}\gamma_m^{y_h}x_m$$

and $\alpha(m-1)^{m,n}$, $\beta(m-1)_{1,j}^{m,n}$ *and* $\gamma(m-1)_i^{m,n}$ *are recursively calculated by the following equations:*

$$\begin{aligned}\alpha(q+1)^{m,n} &= \alpha(q)^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\alpha(1)^{m-q,p} \\ \beta(q+1)_{i,j}^{m,n} &= \beta(q)_{i,j}^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\beta(1)_{i,j}^{m-q,p} \\ \gamma(q+1)_i^{m,n} &= \gamma(q)_i^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\gamma(1)_i^{m-q,p},\end{aligned}$$

where $1 \leq q \leq m-1$ *and* $\alpha(1)^{m,n}$, $\beta(1)_{i,j}^{m,n}$ *and* $\gamma(1)_i^{m,n}$ *are given by*

$$E(w_{m,n}|w_{1,1},\ldots,w_{m-1,l_{m-1}},x_1,\ldots,x_{m-1}) = \alpha(1)^{m,n} + \sum_{i=1}^{m-1}\sum_{j=1}^{l_i}\beta(1)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(1)_i^{m,n}x_i.$$

Proof. Consider the expected causal effect given by (5.4). We may approximate the conditional expectation by:

$$E(y_h|w_{1,1},\ldots,w_{h,l_h},x_1,\ldots,x_h) = \alpha^{y_h} + \sum_{m=1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}w_{m,n} + \sum_{m=1}^{h}\gamma_m^{y_h}x_m. \tag{A.7}$$

Substituting (A.7) into (5.4), we have:

$$\begin{aligned}&E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \alpha^{y_h} + \sum_{m=1}^{h}\gamma_m^{y_h}x_m + \\ &+ \sum_{m=1}^{h}\sum_{n=1}^{l_m}\beta_{m,n}^{y_h}\left[\sum_{w_{1,1},\ldots,w_{h,l_h}} w_{m,n}\prod_{k=1}^{h}P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1})\right].\end{aligned} \tag{A.8}$$

Now, it is necessary to evaluate the term inside square brackets in (A.8):

$$W_{m,n}^{h} = \sum_{w_{1,1},\ldots,w_{h,l_h}} w_{m,n}\prod_{k=1}^{h}P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}). \tag{A.9}$$

To do so, we adopt the same strategy and substitute conditional expectations with linear functions. First, note that (A.9) may be rewritten, for $m > 1$, in the following way:

$$\begin{aligned}W_{m,n}^{h} = &\sum_{w_{1,1},\ldots,w_{h,l_h}} w_{m,n}\prod_{k=1}^{h}P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) = \\ &\sum_{w_{1,1},\ldots,w_{m,l_m}} w_{m,n}\prod_{k=1}^{m}P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) = \\ &\sum_{w_{1,1},\ldots,w_{m-1,l_{m-1}}}\prod_{k=1}^{m-1}P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) \\ &\times E(w_{m,n}|w_{1,1},\ldots,w_{m-1,l_{m-1}},x_1,\ldots,x_{m-1}).\end{aligned} \tag{A.10}$$

Substituting conditional expectations gives

$$W_{m,n}^{h} = \sum_{w_{1,1},\ldots,w_{m-1,l_{m-1}}} \prod_{k=1}^{m-1} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) \times \left[\alpha(1)^{m,n} + \sum_{i=1}^{m-1}\sum_{j=1}^{l_i}\beta(1)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(1)_{i}^{m,n}x_i\right], \tag{A.11}$$

where we adopt the following notation for $E(w_{m,n}|w_{1,1},\ldots,w_{m-1,l_{m-1}},x_1,\ldots,x_{m-1})$ linear approximations:

$$\alpha(1)^{m,n} + \sum_{i=1}^{m-1}\sum_{j=1}^{l_i}\beta(1)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(1)_{i}^{m,n}x_i. \tag{A.12}$$

Here, $\alpha(1)^{m,n}$ is the constant term, $\beta(1)_{i,j}^{m,n}$ is the coefficient of $w_{i,j}$ and $\gamma(1)_{i}^{m,n}$ is the coefficient of x_i. The number 1 inside brackets is an index used for the induction argument.

Suppose that

$$W_{m,n}^{h} = \sum_{w_{1,1},\ldots,w_{m-q,l_{m-q}}} \prod_{k=1}^{m-q} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) \times \left[\alpha(q)^{m,n} + \sum_{i=1}^{m-q}\sum_{j=1}^{l_i}\beta(q)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(q)_{i}^{m,n}x_i\right]. \tag{A.13}$$

for any q, $1 \le q \le m-1$. Summing through $w_{m-q,1},\ldots,w_{m-q,l_{m-q}}$, we have:

$$\begin{aligned} W_{m,n}^{h} = & \sum_{w_{1,1},\ldots,w_{m-(q+1),l_{m-(q+1)}}} \prod_{k=1}^{m-(q+1)} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) \\ & \times \left[\alpha(q)^{m,n} + \sum_{i=1}^{m-(q+1)}\sum_{j=1}^{l_i}\beta(q)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(q)_{i}^{m,n}x_i \right. \\ & \left. + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\left(\alpha(1)^{m-q,p} + \sum_{i=1}^{m-(q+1)}\sum_{j=1}^{l_i}\beta(1)_{i,j}^{m-q,p}w_{i,j} + \sum_{i=1}^{m-1}\gamma(1)_{i}^{m-q,p}x_i\right)\right], \end{aligned} \tag{A.14}$$

where the term inside parentheses is the linear approximation of

$$E(w_{m-q,p}|w_{1,1},\ldots,w_{m-(q+1),l_{m-(q+1)}},x_1,\ldots,x_{m-(q+1)}), \tag{A.15}$$

given by (A.12). Rearranging and grouping (A.14) gives

$$W_{m,n}^{h} = \sum_{w_{1,1},\ldots,w_{m-(q+1),l_{m-(q+1)}}} \prod_{k=1}^{m-(q+1)} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) \times \left[\alpha(q+1)^{m,n} + \sum_{i=1}^{m-(q+1)}\sum_{j=1}^{l_i}\beta(q+1)_{i,j}^{m,n}w_{i,j} + \sum_{i=1}^{m-1}\gamma(q+1)_{i}^{m,n}x_i\right], \tag{A.16}$$

where we define:

$$\begin{aligned} \alpha(q+1)^{m,n} &= \alpha(q)^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\alpha(1)^{m-q,p} \\ \beta(q+1)_{i,j}^{m,n} &= \beta(q)_{i,j}^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\beta(1)_{i,j}^{m-q,p} \\ \gamma(q+1)_{i}^{m,n} &= \gamma(q)_{i}^{m,n} + \sum_{p=1}^{l_{m-q}}\beta(q)_{m-q,p}^{m,n}\gamma(1)_{i}^{m-q,p}. \end{aligned} \tag{A.17}$$

So, by induction, we conclude that (A.13) is valid for $1 \le q \le m-1$. Thus, substituting q by its maximum, we have:

$$\begin{aligned} W_{m,n}^{h} = & \sum_{w_{1,1},\ldots,w_{h,l_h}} w_{m,n}\prod_{k=1}^{h} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) = \\ & \sum_{w_{1,1},\ldots,w_{1,l_1}} P(w_{1,1},\ldots,w_{1,l_1}) \times \left[\alpha(m-1)^{m,n} + \sum_{j=1}^{l_1}\beta(m-1)_{1,j}^{m,n}w_{1,j} + \sum_{i=1}^{m-1}\gamma(m-1)_{i}^{m,n}x_i\right], \end{aligned} \tag{A.18}$$

which reduces to

$$W^h_{m,n} = \sum_{w_{1,1},\ldots,w_{h,l_h}} w_{m,n} \prod_{k=1}^{h} P(w_{k,1},\ldots,w_{k,l_k}|w_{1,1},\ldots,w_{k-1,l_{k-1}},x_1,\ldots,x_{k-1}) = \\ \alpha(m-1)^{m,n} + \sum_{j=1}^{l_1} \beta(m-1)^{m,n}_{1,j} E(w_{1,j}) + \sum_{i=1}^{m-1} \gamma(m-1)^{m,n}_{i} x_i. \tag{A.19}$$

Since $q \geq 1$, equation (A.19) is valid for $m \geq 2$. If $m = 1$, equation (A.9) reduces to $E(w_{1,n})$. Considering this and substituting (A.19) into (A.8), we have

$$E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \alpha^{y_h} + \sum_{j=1}^{l_1} \beta^{y_h}_{1,j} E(w_{1,j}) + \sum_{i=1}^{h} \gamma^{y_h}_{i} x_i \\ + \sum_{m=2}^{h} \sum_{n=1}^{l_m} \beta^{y_h}_{m,n} \left[\alpha(m-1)^{m,n} + \sum_{j=1}^{l_1} \beta(m-1)^{m,n}_{1,j} E(w_{1,j}) + \sum_{i=1}^{m-1} \gamma(m-1)^{m,n}_{i} x_i \right], \tag{A.20}$$

which may be rearranged[8] in the following way:

$$E(y_h|\hat{x}_1,\ldots,\hat{x}_h) = \alpha^{y_h} + \sum_{m=2}^{h} \sum_{n=1}^{l_m} \beta^{y_h}_{m,n} \alpha(m-1)^{m,n} \\ + \sum_{j=1}^{l_1} \left[\beta^{y_h}_{1,j} + \sum_{m=2}^{h} \sum_{n=1}^{l_m} \beta^{y_h}_{m,n} \beta(m-1)^{m,n}_{1,j} \right] E(w_{1,j}) \\ + \sum_{i=1}^{h-1} \left[\gamma^{y_h}_{i} + \sum_{m=i+1}^{h} \sum_{n=1}^{l_m} \beta^{y_h}_{m,n} \gamma(m-1)^{m,n}_{i} \right] x_i + \gamma^{y_h}_{h} x_h, \tag{A.21}$$

where α^{y_h}, $\beta^{y_h}_{m,n}$ and $\gamma^{y_h}_{m}$ are given by (A.7) and $\alpha(m-1)^{m,n}$, $\beta(m-1)^{m,n}_{1,j}$ and $\gamma(m-1)^{m,n}_{i}$ are recursively given by (A.17). □

A.3 Data

Data source and transformations are shown in Tables 2 and 3. Since meeting days are irregular and do not coincide exactly with available data points, I made daily linear interpolations in all time series and matched these interpolations with meeting days. A more complicated scheme were drawn for variables predicting future. Focus survey data is not an issue because it is daily available, but the BCB projections in inflation reports are not, and so I proceeded with a different interpolation scheme.

Table 4 represents data on internal forecasts made in days t, $t+1$, $t+2$, etc. These are represented in each line of the table. Also, each entry of the table, labeled in row-column convention as $(t+k, t+l)$, represents the forecast made in date $t+k$ for variable at $t+l$. Forecasts represented by "O" are available data, and the ones represented by "X" are unavailable and I proceed to interpolate then. So, from the table it is implicit that forecasts are made every k days and are also made for every k days in the future. Notice that the entries for forecast data are empty for $(t+k, t+l)$ where $l < k$ because these represent predictions of past data.

For each line where there is forecast data we linearly interpolate then daily. So, we arrive at Table 5, where H represents a linear horizontal interpolation.

To interpolate periods where there are not projections, things got trickier. I may interpolate then vertically, but columns where there is not data will became empty at bottom points, or must be extrapolated. A better option is to interpolate diagonally, but this is going to put information of future forecasts in interpolations, something plausible, since forecasters probably have more information close to forecast announcement, but debatable. For example, the interpolated forecast at date $t+k-1$ for $t+k$ will be very close to forecast available at $t+k$ to $t+k$. A possible critique

[8] In deducing (A.21), we applied the following summation inversion formula:

$$\sum_{m=2}^{h} \sum_{i=1}^{m-1} f(m,i) = \sum_{i=1}^{h-1} \sum_{m=i+1}^{h} f(m,i).$$

The proof is easy using Iverson brackets.

	t	$t+1$	$t+2$	$\dots$	$t+k$	$t+k+1$	$t+k+2$	$t+k+3$	$\dots$	$t+2k$
t	O	X	X	$\dots$	O	X	X	X	$\dots$	O
$t+1$		X	X	$\dots$	X	X	X	X	$\dots$	X
$t+2$			X	$\dots$	X	X	X	X	$\dots$	X
$\vdots$				$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$t+k$					O	X	X	X	$\dots$	O
$t+k+1$						X	X	X	$\dots$	X
$t+k+2$							X	X	$\dots$	X
$t+k+3$								X	$\dots$	X
$\vdots$									$\ddots$	$\vdots$
$t+2k$										O

Table 4: Schematic representation of BCB internal forecast data.

	t	$t+1$	$t+2$	$\dots$	$t+k$	$t+k+1$	$t+k+2$	$t+k+3$	$\dots$	$t+2k$
t	O	H	H	$\dots$	O	H	H	H	$\dots$	O
$t+1$		X	X	$\dots$	X	X	X	X	$\dots$	X
$t+2$			X	$\dots$	X	X	X	X	$\dots$	X
$\vdots$				$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$t+k$					O	H	H	H	$\dots$	O
$t+k+1$						X	X	X	$\dots$	X
$t+k+2$							X	X	$\dots$	X
$t+k+3$								X	$\dots$	X
$\vdots$									$\ddots$	$\vdots$
$t+2k$										O

Table 5: Schematic representation of BCB internal forecast data. "H" represents horizontally linearly interpolated data.

on this procedure is that the forecast at $t+1$ for $t+k$ is probably closer to the forecast made at t to $t+k$, and this is not being considered. So, I combine both ideas.

Consider the diagonal linear interpolation made after the horizontal one. This is represented in Table 6. Consider now an auxiliary variable α which varies linearly between t and $t+k$, $t+k$ and $t+2k$, etc., such that $\alpha(t+nk)=0$ and $\alpha(t+(n+1)k)=1$, $n=0,1,\dots$. Let $A_{t+i,t+j}$ be the element $(t+i,t+j)$ of Table 6 and $I_{t+i,t+j}$ be the element $(t+i,t+j)$ of the final interpolation. So, I proceed in the following way: $I_{t+i,t+j}=\alpha(t+i)A_{t+i,t+j}+[1-\alpha(t+i)]A_{h,t+j}$ such that h is the maximum row value with $h<t+i$ and containing original data.

	t	$t+1$	$t+2$	$\dots$	$t+k$	$t+k+1$	$t+k+2$	$t+k+3$	$\dots$	$t+2k$
t	O	H	H	$\dots$	O	H	H	H	$\dots$	O
$t+1$		D	D	$\dots$	D	D	D	D	$\dots$	D
$t+2$			D	$\dots$	D	D	D	D	$\dots$	D
$\vdots$				$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$t+k$					O	H	H	H	$\dots$	O
$t+k+1$						D	D	D	$\dots$	D
$t+k+2$							D	D	$\dots$	D
$t+k+3$								D	$\dots$	D
$\vdots$									$\ddots$	$\vdots$
$t+2k$										O

Table 6: Schematic representation of BCB internal forecast data. "H" represents horizontally linearly interpolated data and "D" represents diagonally linearly interpolated data.

In this way I maintain interpolations of forecasts closer to real ones when dates are close. This procedure is just a coherence exercise, because in practice, only longer horizon forecasts are relevant for the estimates, and the interpolation procedure has negligible impact on results.

A.4 GAM model

In Section 4 I analyze non-linear systems. In the estimation procedure, generalized additive models to deal with non-linearities were adopted. References for techniques used are Wood [2017], Hastie [2017] and Hastie and Tibshirani [1987].

Generally speaking, a GAM is a model of the form

$$g(\mu_i) = A_i\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_{3,4}(x_{3i}, x_{4i}) + \ldots \tag{A.22}$$

where g is a link function, $\mu_i = E(Y_i)$ with $Y_i \sim EF(\mu_i, \phi)$ as a response variable distributed according to an exponential family distribution, $A_i\theta$ represents strictly parametric model components and functions f_k are smooth functions of the covariates x_k.

In all specifications of Section 4 we adopted an identity link, Gaussian response and one multivariate smooth function embracing all covariates. For estimation purposes we adopted the dimension of the basis of the smoother as 200. Details on GAM estimation are available in Wood [2017]. So, the models presented have the form

$$E(Y_{t+h}) = f(X_t, Y_{t-1}, Z_t), \tag{A.23}$$

where variables are those in equations (4.2). I run (A.23) to estimate conditional expectations in equation (4.4). The procedure adopted for EGCIRF and ECGI computations is explained in Section 4.

A.5 Robustness check - changing regressors set

In Tables 2 and 3 I present the covariates used in Section 6.2 regressions. A wider set of variables, presented in Tables 7, 8 and 9, were chosen from Inflation Reports of BCB, Focus survey and macroeconomic aggregates. This wider set does not present all possible variables because lots of them simply do not have any affect on impulse responses.

We can compare the EGCIRFs computed in Section 6.2 with the ones computed using covariates of Tables 7, 8 and 9 together with 4 lags of policy variable and lags 2 to 4 of real GDP gap, as explained in Section 6.2. This comparison is shown in Figures 25, 26, 27 and 28.

Results are remarkably similar, and are not statistically different. For the sake of robustness, results presented in Section 6.2 use less 20 variables, and these variables may be added or removed without significant change in EGCIRFs. Actually, even various variables in Tables 2 and 3 can be removed or redefined without drastically changing conclusions.

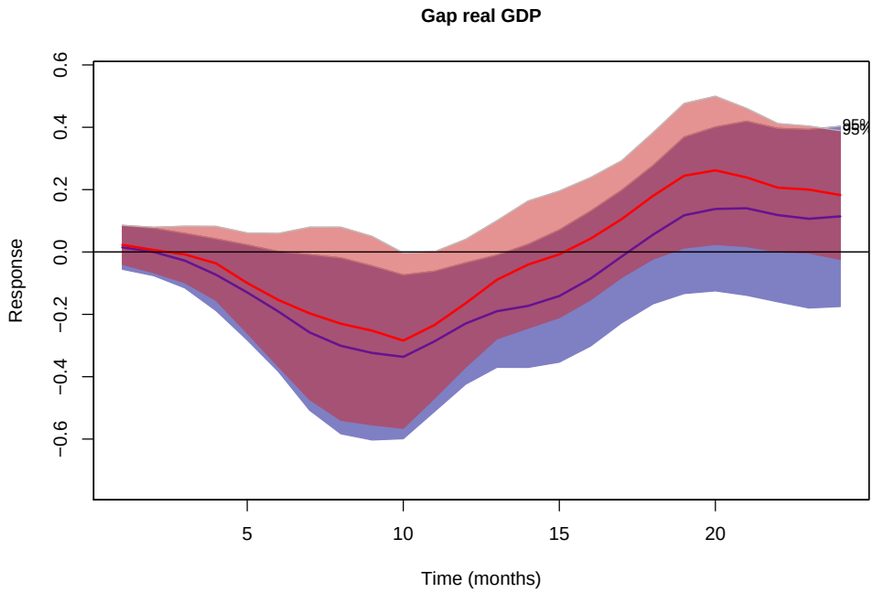

Figure 25: EGCIRF of real GDP gap for 25-basis-point rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using Tables 7, 8 and 9 covariates.

Regressor	Description	Source
S	Target Selic rate change. This is the policy variable, determined on each COPOM meeting. Its coefficient is the causal effect of interest.	BCB - Copom Meetings
R_1	3-month difference on target Selic rate	BCB - SGS
R_2	12-month difference on target Selic rate	BCB - SGS
R_3	Last 3-month change of inflation index	BCB - SGS
R_4	Last 6-month change of inflation index	BCB - SGS
R_5	Last 12-month change of inflation index	BCB - SGS
R_6	Median of Focus IPCA 1-month-ahead projections - seasonally adjusted	BCB - Market Expectations System
R_7	Median of Focus IPCA 6-month-ahead projections - seasonally adjusted	BCB - Market Expectations System
R_8	Median of Focus IPCA 12-month-ahead projections - seasonally adjusted	BCB - Market Expectations System
R_9	Difference between the reference scenario IPCA forecast by the BCB for meeting day and 12-month change in the IPCA	BCB - Inflation Reports
R_{10}	Difference between the reference scenario IPCA forecast by the BCB for 3 months ahead and the BCB forecast for meeting day, both adjusted for inflation target	BCB - Inflation Reports
R_{11}	Difference between the reference scenario IPCA forecast by the BCB for 6 months ahead and the BCB forecast for 3 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_{12}	Difference between the reference scenario IPCA forecast by the BCB for 9 months ahead and the BCB forecast for 6 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_{13}	Difference between the reference scenario IPCA forecast by the BCB for 12 months ahead and BCB forecast for 9 months ahead, both adjusted for inflation target	BCB - Inflation Reports
R_{14}	Monthly variation of median meeting day Focus Selic rate forecasts	BCB - Market Expectations System

Table 7: Regressors used in preliminary estimates, part I.

Regressor	Description	Source
R_{15}	Monthly variation of median 6-month-ahead Focus Selic rate forecasts	BCB - Market Expectations System
R_{16}	Monthly variation of median 12-month-ahead Focus Selic rate forecasts	BCB - Market Expectations System
R_{17}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for meeting day by Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{18}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for 1 quarter ahead by Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{19}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for 2 quarters ahead by Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{20}	Difference between the median of 1 quarter accumulated annualized GDP forecasts for 3 quarters ahead Focus and accumulated 12 months GDP	BCB - Market Expectations System
R_{21}	Real GDP 3-month log variation	BCB - SGS
R_{22}	Real GDP 6-month log variation	BCB - SGS
R_{23}	Real GDP 9-month log variation minus real GDP 6-month log variation	BCB - SGS
R_{24}	Real GDP 12-month log variation minus real GDP 9-month log variation	BCB - SGS
R_{25}	Real exchange rate 1-month log variation - US dollar	BCB - SGS
R_{26}	Real exchange rate 6-month log variation minus real exchange rate 1-month log variation - US dollar	BCB - SGS
R_{27}	Real exchange rate 12-month log variation minus real exchange rate 6-month log variation - US dollar	BCB - SGS
R_{28}	Nominal exchange rate 1-month log variation - US dollar	BCB - SGS
R_{29}	Nominal exchange rate 6-month log variation - US dollar	BCB - SGS
R_{30}	Nominal exchange rate 12-month log variation - US dollar	BCB - SGS

Table 8: Regressors used in preliminary estimates, part II.

Regressor	Description	Source
R_{31}	Focus end-of-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{32}	Focus 6-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{33}	Focus 12-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{34}	Monthly variation of Focus end-of-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{35}	Monthly variation of Focus 6-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{36}	Monthly variation of Focus 12-month expected exchange rate variation - US dollar	BCB - Market Expectations System
R_{37}	Focus expected industrial production 1 month ahead minus 1-year industrial production variation	BCB - Market Expectations System
R_{38}	Focus expected industrial production 6 months ahead minus 1-year industrial production variation	BCB - Market Expectations System
R_{39}	Focus expected industrial production 9 months ahead minus 1-year industrial production variation	BCB - Market Expectations System
R_{40}	Three-month difference of the reference scenario IPCA forecast by the BCB for meeting day	BCB - Inflation Reports
R_{41}	Three-month difference of the reference scenario IPCA forecast by the BCB for 3 months ahead	BCB - Inflation Reports
R_{42}	Three-month difference of the reference scenario IPCA forecast by the BCB for 6 months ahead	BCB - Inflation Reports
R_{43}	Three-month difference of the reference scenario IPCA forecast by the BCB for 9 months ahead	BCB - Inflation Reports
R_{44}	Three-month difference of the reference scenario IPCA forecast by the BCB for 12 months ahead	BCB - Inflation Reports
R_{45}	Difference between market scenario IPCA forecast by the BCB for 12 months ahead and the same BCB forecast for 6 months ahead, both adjusted for inflation target	BCB - Inflation Reports

Table 9: Regressors used in preliminary estimates, part III.

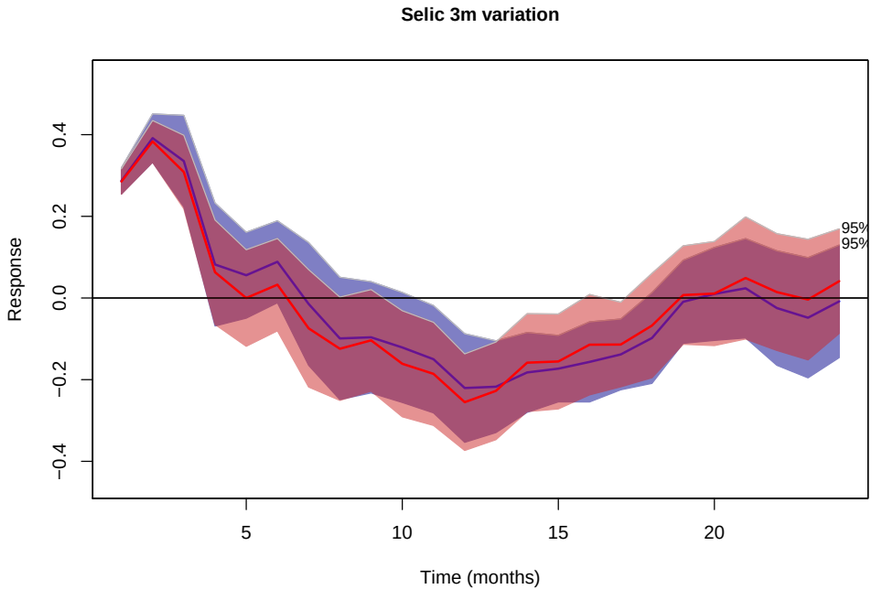

Figure 26: EGCIRF of 3-month Selic rate change for 25-basis-point rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using Tables 7, 8 and 9 covariates.

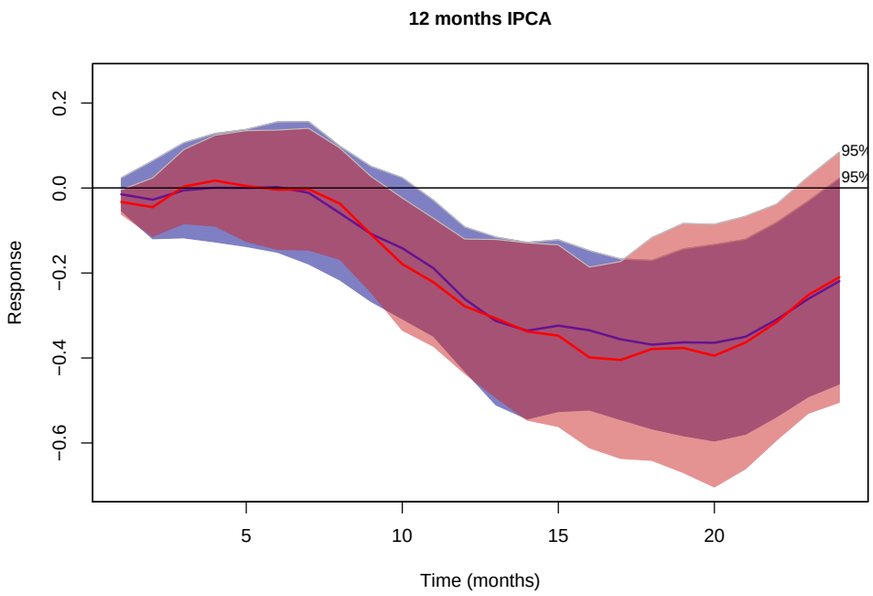

Figure 27: EGCIRF of 12-month IPCA for 25-basis-point rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using Tables 7, 8 and 9 covariates.

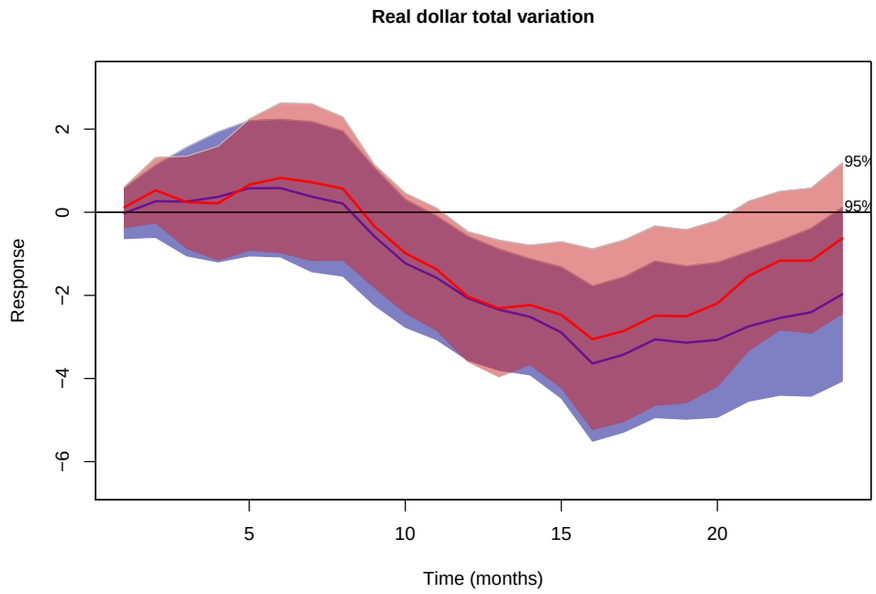

Figure 28: EGCIRF of real US dollar total variation for 25-basis-point rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using Tables 7, 8 and 9 covariates.

A.6 Robustness check - changing sample range

Another robustness exercise is to change the sample size, computing EGCIRFs for shorter time samples. This additionally gives clues about possible changes in policy makers or economic agents behavior. Final results presented in Figures 17, 18, 19 and 20 ranges from 2003-02-19 to 2019-02-06 in their first horizon[9]. Cutting all data beyond 2014 gives a sample from 2003-02-19 to 2014-12-03, reducing the estimation period by more than 4 years from the end. Figures 29, 30, 31 and 32 show comparisons among final results and reduced sample estimates. Notice that results are reasonably similar and, indeed, not statistically different.

Things get worse deleting data from sample beginning, starting from 2007-07-18 instead of 2003-02-19, in a more than 4 years cut from original starting point. Confidence intervals got much bigger and results behave more wildly. This happens probably because the period 2002-2003 was turbulent for monetary policy in Brazil. Markets feared the election of Luiz Inácio Lula da Silva, and strong depreciations of domestic currency were experienced, accompanied by higher inflation and interest rates. So, the results are particularly sensitive to this period. But, looking for comparisons in Figures 33, 34, 35 and 36, they are not so different, specially considering the confidence intervals.

It can be argued that, despite economic turbulence, the BCB is yet looking for inflation control, trying to put it on target, and did not fundamentally changed its behavior, only rising Selic rate strongly because of stronger inflationary forces. One point corroborating this view is the presence of much bigger confidence intervals for the short samples estimations cutting data before 2007 in comparison with the former sample range cut exercise. The causal interpretation of the regressions carried out in our analysis depends on the existence of unpredictable behavior from policy makers. Indeed, more unpredictability increases estimation precision, other things equal. This is so because causal estimations based on SCM may be interpreted as IPTW (Inverse Probability of Treatment Weighting) estimators, as shown in Pearl [2009], and so with more variation on treatment choice (Selic rate determination), results may be more precise, and possibly this is what happens in the analysis. Of course that this assertion must be investigated more deeply, although I do not do this here. After all, considering the confidence intervals, results in Figures 33, 34, 35 and 36 are yet reasonably close to our main results.

A.7 Regressions diagnostics

Since each EGCIRF is comprised of h regressions, where h is the horizon of the response, we may analyze the properties of these regressions for diagnostics. First of all, let's test the premise of stationarity of regressors. Figure 37 shows a Venn diagram summarizing all tests.

[9]Different horizons may comprise different sample sizes because the regression for each horizon uses all available data. Since for longer horizons less data is available, data availability changes for longer horizons. So, I use as reference the sample for horizon 1.

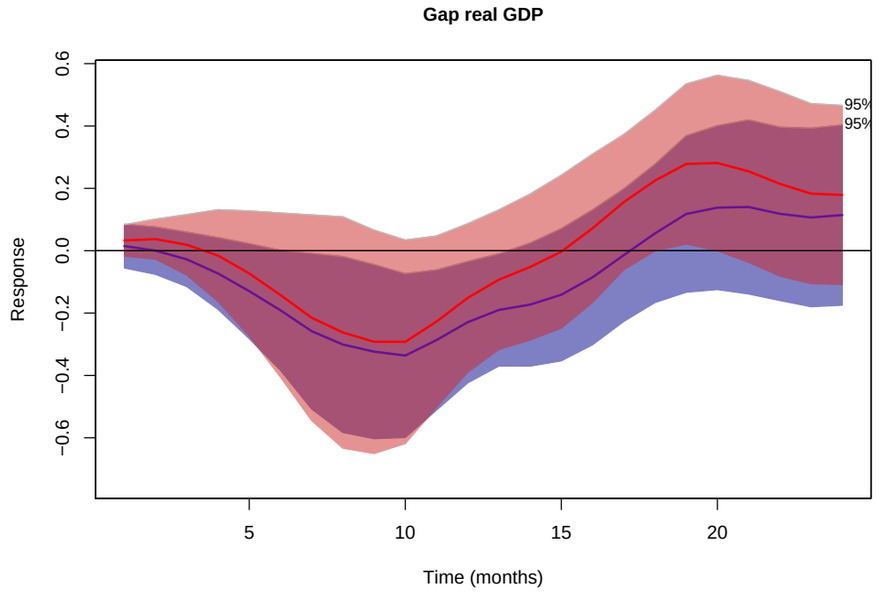

Figure 29: EGCIRF of real GDP gap for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only up to 2014.

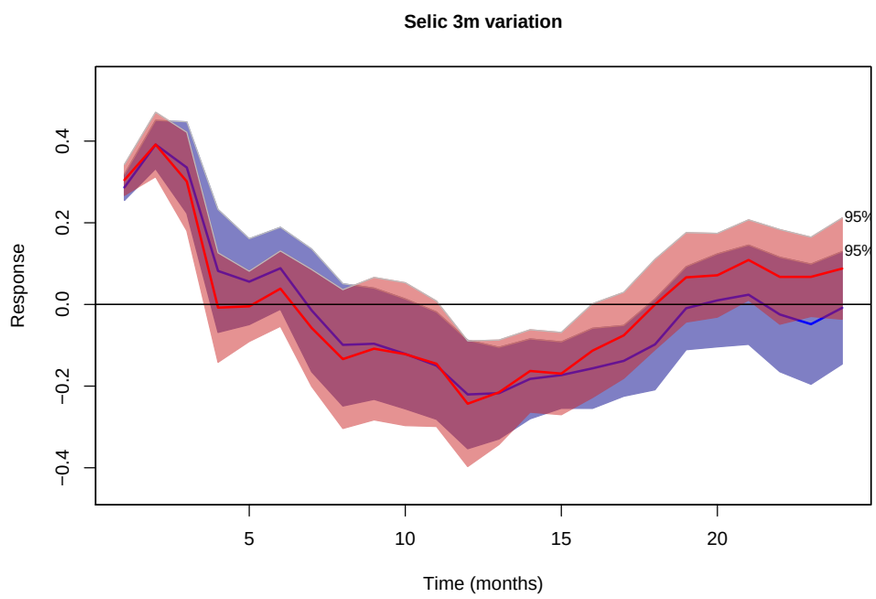

Figure 30: EGCIRF of 3 months Selic rate variation for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only up to 2014.

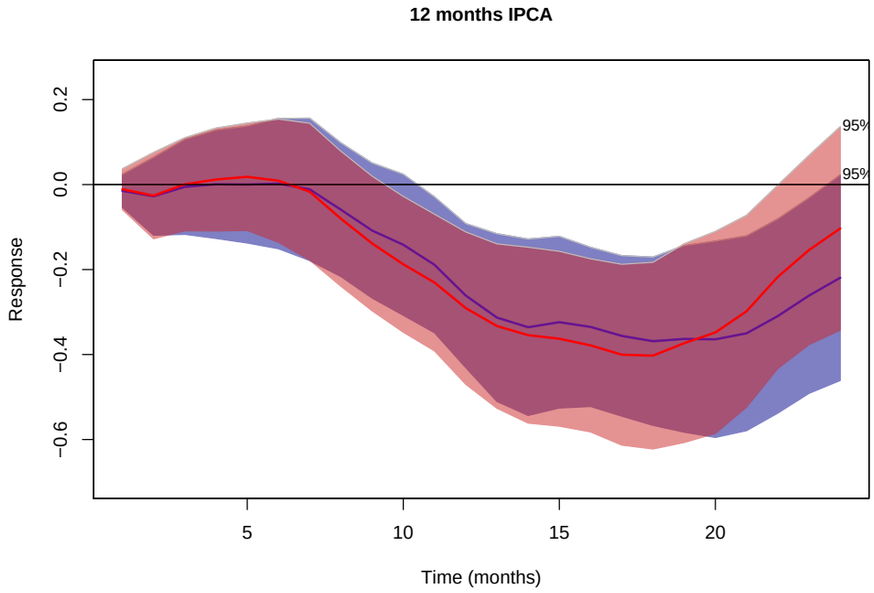

Figure 31: EGCIRF of 12 months IPCA for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only up to 2014.

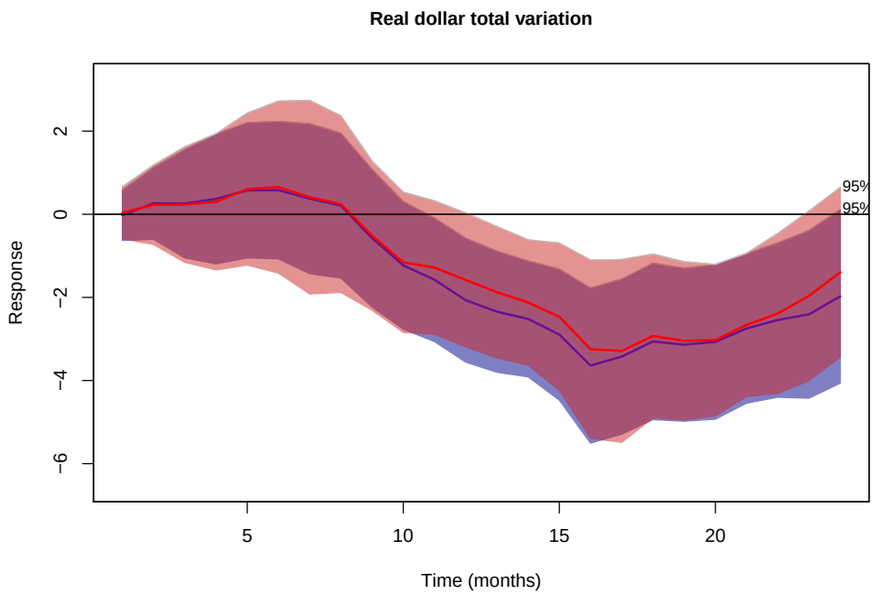

Figure 32: EGCIRF of real dollar total variation for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only up to 2014.

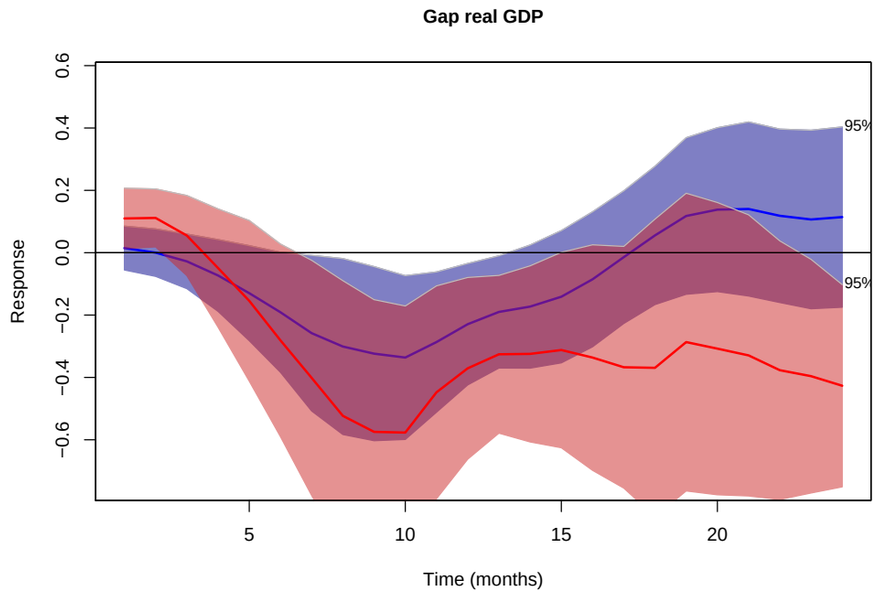

Figure 33: EGCIRF of real GDP gap for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only from mid 2007 on.

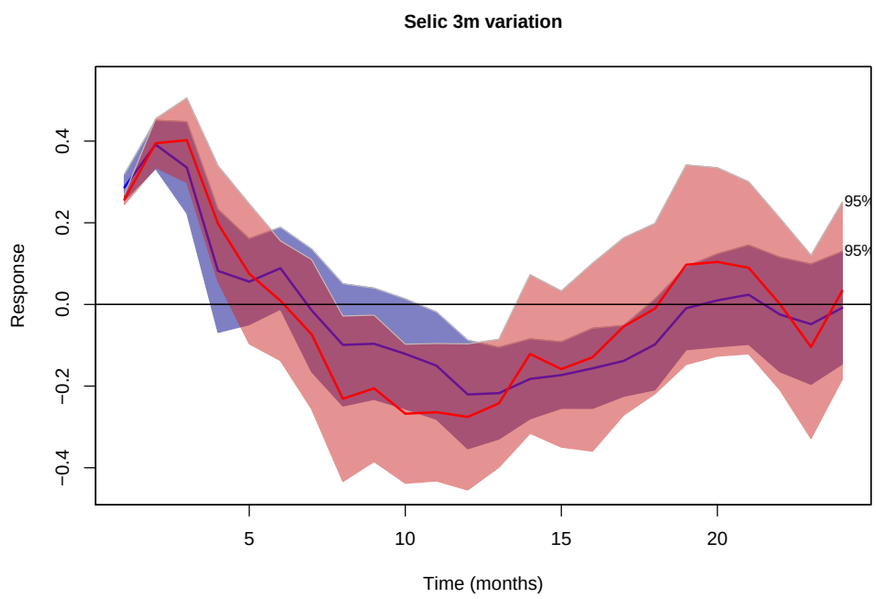

Figure 34: EGCIRF of 3 months Selic rate variation for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only from mid 2007 on.

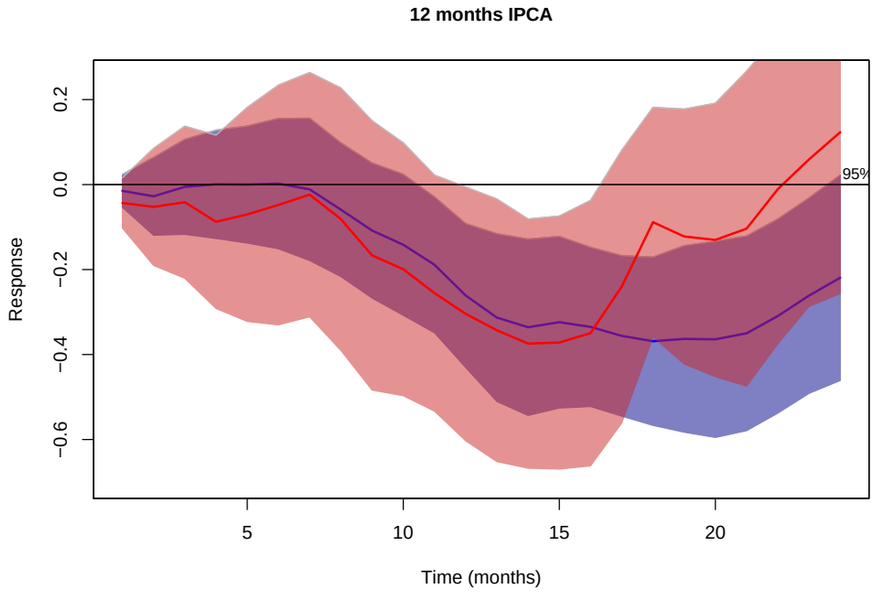

Figure 35: EGCIRF of 12 months IPCA for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only from mid 2007 on.

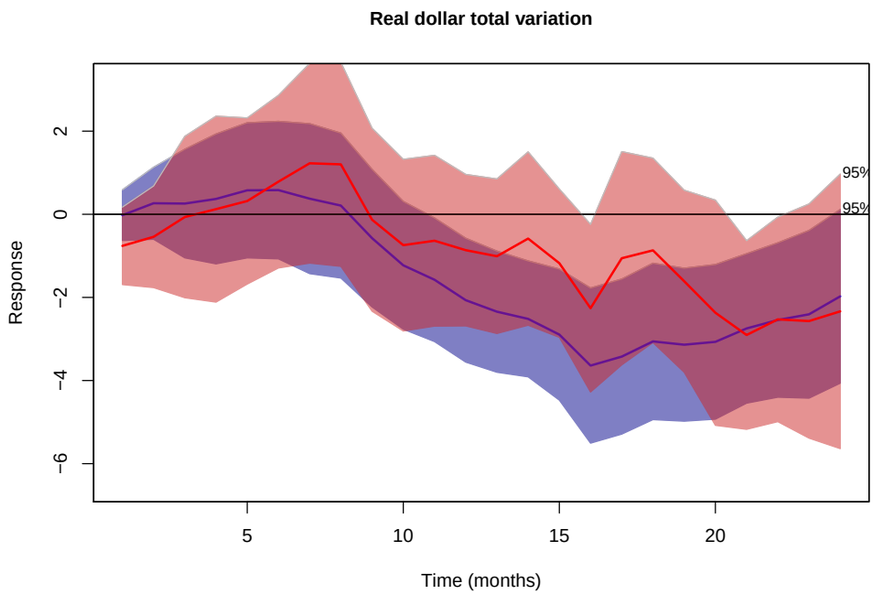

Figure 36: EGCIRF of real dollar total variation for 25 basis points rise in Selic rate against no variation. Blue results are the ones in Section 6.2 and red results correspond to estimates using data only from mid 2007 on.

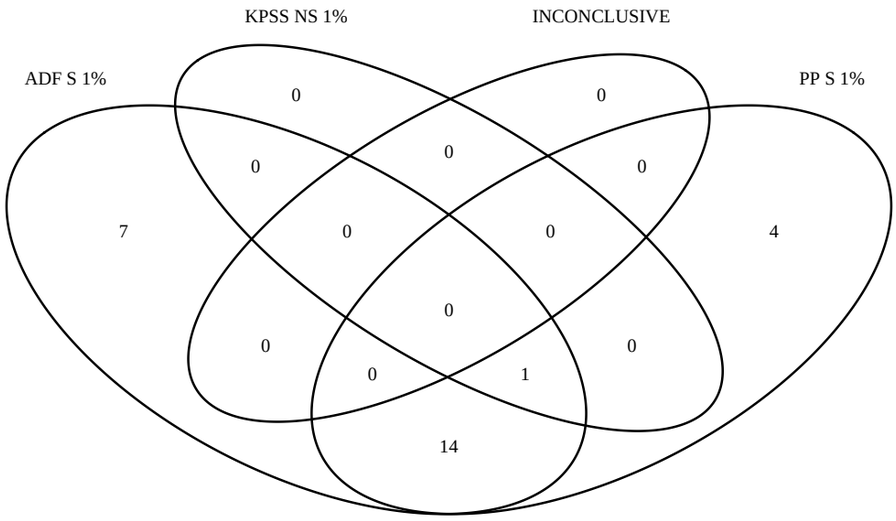

Figure 37: Venn diagram representing ADF, KPSS and PP stationarity tests results at 1% level.

ADF is the Augmented Dickey–Fuller test (Said and Dickey [1984], Banerjee et al. [1993], Greene [2003]) whose null hypothesis is the presence of unit root, KPSS is the Kwiatkowski-Phillips-Schmidt-Shin test (Kwiatkowski et al. [1992]) whose null hypothesis is the stationarity of the series and PP is the Phillips–Perron unit root test (Perron [1988], Banerjee et al. [1993]) whose null hypothesis is the presence of unit root. The diagram reads as the number of series decided by each test, that is, the sets ADF and PP contain series whose presence of unit root was rejected at 1% level, the KPSS contains series whose stationarity was rejected at 1% level and the INCONCLUSIVE set contains series for which no conclusion is reached. As can be seen, only one series is decided non-stationary by KPSS test, but ADF and PP together decided for its stationarity, so I suppose it is stationary.

For the response variables, results are:

1. Real GDP gap with unit root rejected at 5% by ADF and stationarity not rejected by any standard[10] level by KPSS;

2. 3 months Selic variation with unit root rejected at 1% by ADF and PP and stationarity not rejected by any standard level by KPSS;

3. 12 months IPCA with unit root rejected at 1% by ADF and stationarity not rejected at 1% by KPSS;

4. Real dollar total variation with unit root rejected at 1% by ADF and PP and stationarity not rejected at 1% by KPSS.

So, all variables are relatively well behaved regarding their stationarity.

Now, consider residual tests for each horizon of each EGCIRF. Figures 38, 39, 40 and 41 show p-values for Shapiro-Wilk normality test, KPSS[11], ADF and PP unit root tests, Durbin-Watson autocorrelation test and the Variance Inflation Factor (VIF) for the policy variable coefficient (Sheather [2009]). The red dashed line represents 5% significance level. In general, results point to stationary residuals, with some more severe deviations from normality by real GDP gap and real exchange rate responses. This deviation from normality is not so severe if we consider the wider set of covariates of Section A.5, and since results are close, this does not seem too much important.

Another important feature are the VIF values, between 5 and 10 (close to 7.5 for all estimates), showing some concern about correlations among policy variable and covariates. According to Neter et al. [1989], VIF > 10 is reason for concern, so we are fine. But, Sheather [2009] points to a cutoff of VIF > 5. One strong source of these correlations is the presence of lags of the policy variable, which we added by reasons presented in Section 6.2. Without these lags, all VIFs became closer to 5, in more acceptable levels. Since, with the exception of 1 month EGCIRF for 3 months Selic rate

[10] Standard levels are 1%, 5% and 10%.

[11] P-values above 10% for KPSS are not computed and so are represented as 10%.

variation, all results are pretty close with or without the lags added, we conclude that VIF is not a concern for the estimations.

Finally, consider the Durbin-Watson p-values for all estimations (Durbin and Watson [1971], Chatterjee and Simonoff [2013]). It points to the presence of autocorrelation in residuals, and this is unavoidable as explained in Section 6.2. For this reason we compute all confidence intervals using Newey-West HAC estimators.

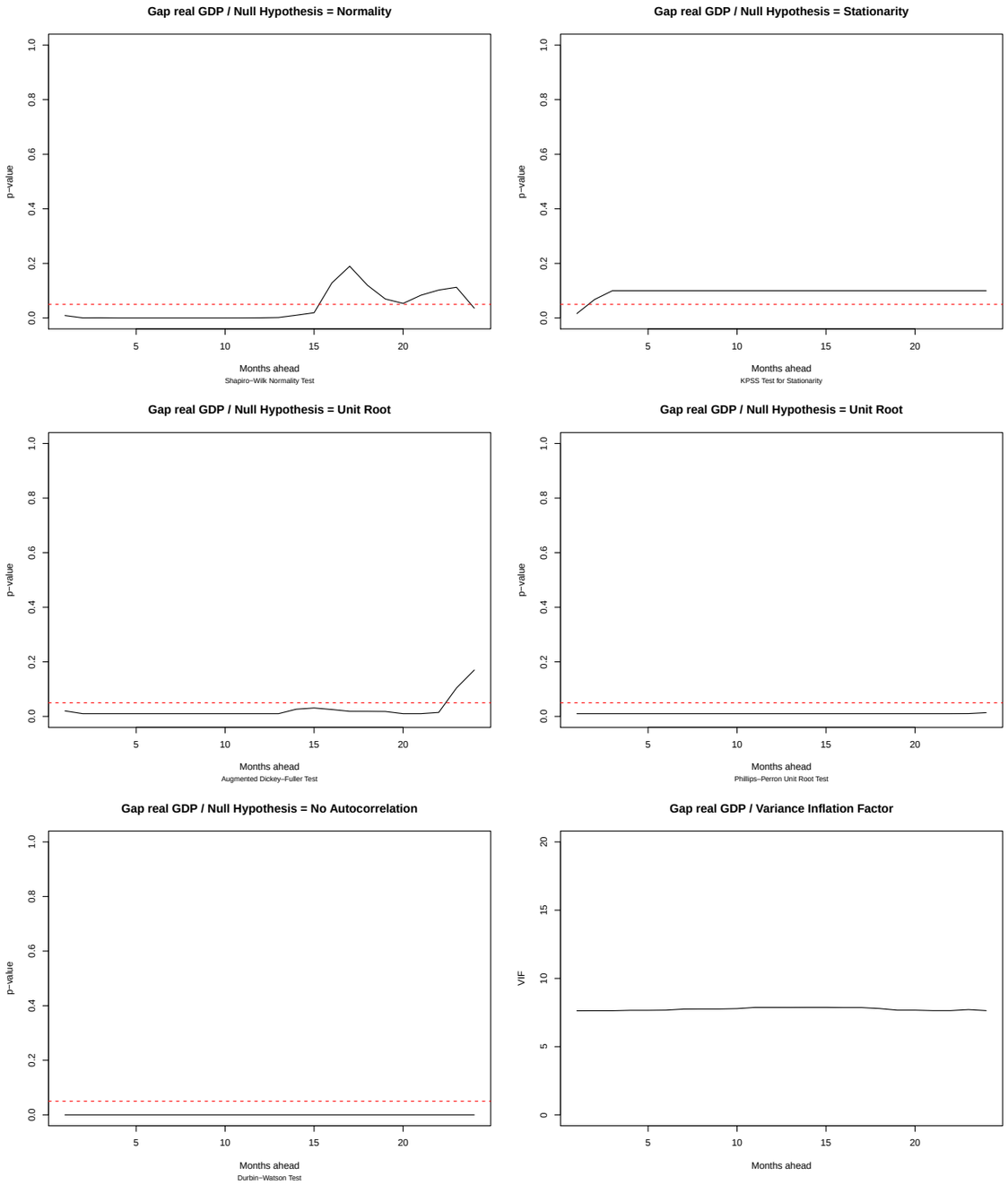

Figure 38: Residuals diagnostics and VIF for real GDP gap.

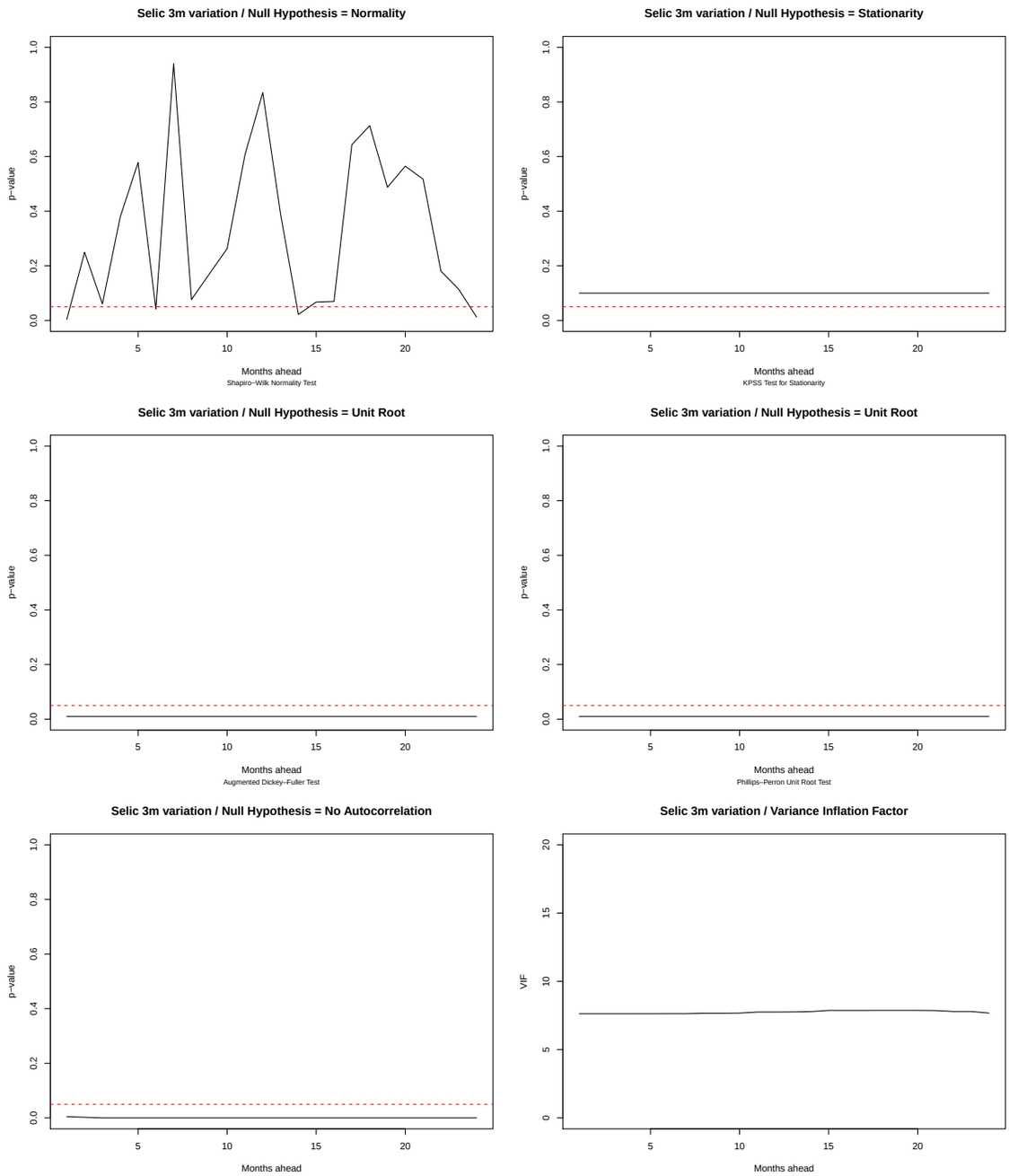

Figure 39: Residuals diagnostics and VIF for 3 months Selic variation.

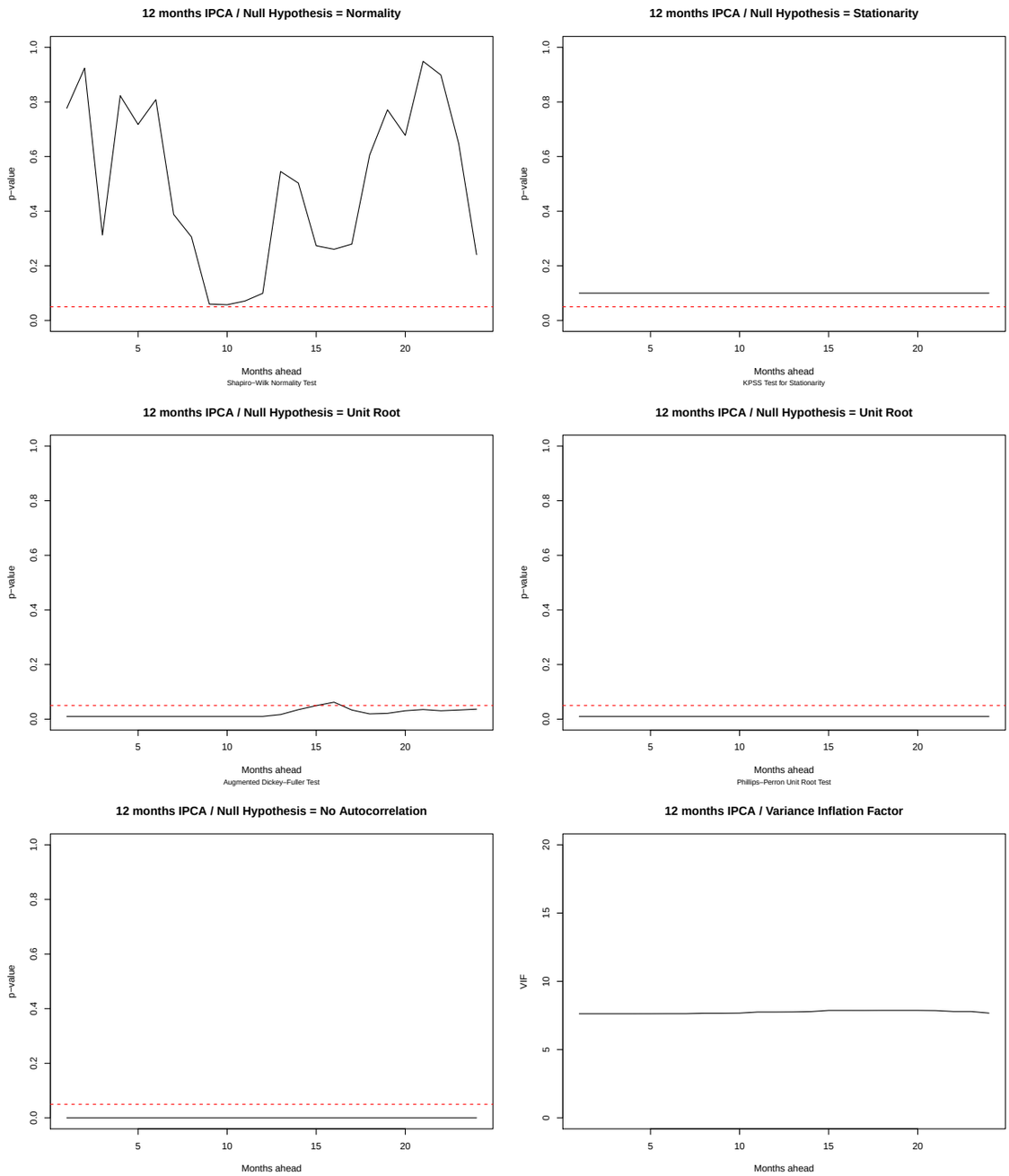

Figure 40: Residuals diagnostics and VIF for 12 months IPCA.

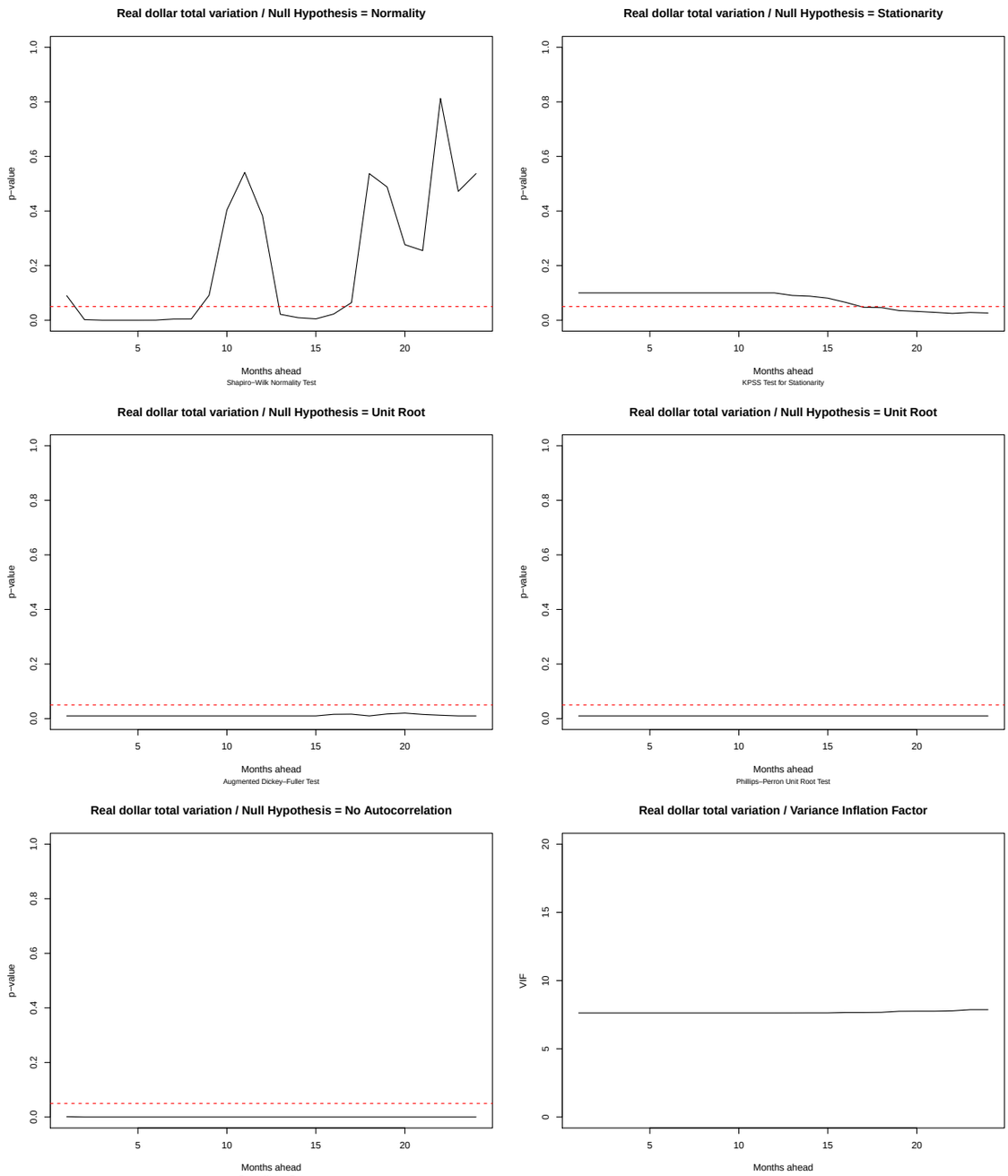

Figure 41: Residuals diagnostics and VIF for real dollar total variation.