Structural counterfactual analysis in macroeconomics: theory and inference*

Endong Wang[†]

December 26, 2024

Latest Version

*I am indebted to my advisors, Jean-Marie Dufour, Victoria Zinde-Walsh, and Russell Davidson, for their guidance, encouragement, and support during my Ph.D. studies at McGill University. I owe Eugène Dettaa, John Galbraith, and Silvia Goncalves for their valuable suggestions and comments. I greatly appreciate Santiago Camara for the insightful advice and data references.

[†]Department of Economics, McGill University, Montréal, Québec H3A2T7, Canada. TEL: 514 772 7078, Email: endong.wang@mail.mcgill.ca, Web: www.endongwang.com.

Abstract

We propose a robust methodology for analyzing two types of macroeconomic policy path counterfactuals: hypothetical trajectories and policy interventions. Our method is grounded in the general linear framework of a structural vector moving-average (SVMA) model and relies solely on the identification of policy shocks, thereby obviating the need for a fully specified structural model. Analytical solutions for counterfactual parameters are derived, and statistical inference is conducted using the Delta method. To enhance identification and estimation, we integrate our counterfactual framework with the local projection-instrumental variable (LP-IV) approach, offering straightforward economic interpretations and robust statistical inference without requiring corrections for serial correlation in the projection error. To account for potential nonlinearities and time-varying relationships between variables, we extend the framework by incorporating a simulation-based approach that utilizes nonlinear parametric models. The applicability of our method is demonstrated through three counterfactual studies on U.S. monetary policy: (1) a historical analysis of a hypothetical post-pandemic interest rate trajectory, (2) a forward-looking evaluation of hawkish or dovish interest rate policies, and (3) an assessment of the effects of an oil price shock, isolating the systematic responses of the interest rate.

Keywords: Policy path counterfactuals; Structural Vector Moving-Average model; Impulse responses; Linear projection-instrumental variable; Statistical inference; Monetary policy.
Journal of Economic Literature Classification: C13, C32, C36, C51, E52.

Contents

1 Introduction

A central research question UB in macroeconomic policy analysis is how the economy would behave under a hypothetical policy path. Counterfactual exercises related to hypothetical policy path can be classified into two main categories: (1) *hypothetical trajectory*, where the policy variable diverges from its baseline path, and (2) *policy intervention*, where the policy variable remains unresponsive to a shock of interest, resulting in a deviation from its impulse response function. Conventional methods, such as conditional forecasting for *hypothetical trajectory* and Bernanke et al.'s monetary policy counterfactuals for *policy intervention* typically rely on fully specified structural models like Structural Vector Autoregressive (SVAR) frameworks. These approaches often involve intricate algorithms and assumptions embedded within the model structure.[1] This paper introduces an empirical robust structural methodology utilizing a projection estimation approach. Our method offers the advantage of robustness to model misspecification, a projection-based estimation and simple statistical inference for counterfactual outcomes, and establishes a link between counterfactual analysis and the macroeconometrics literature on Local Projections.

Counterfactual analysis of treatment effects has been widely studied in microeconomics, primarily through the potential outcomes framework and the concept of randomized control trials (RCTs) (Fisher, 1925; Rubin, 1974). However, conducting RCTs in macroeconomics is impractical. As a result, a substantial body of literature focuses on studying policy effects by identifying unanticipated and exogenous interventions, commonly referred to as structural shocks (Ramey, 2016, Kilian and Lütkepohl, 2017, Stock and Watson, 2018). While much of the macroeconometrics literature has focused on the shock identification and estimation methods since the seminal work of Sims (1980),[2] there has been limited research on counterfactuals involving a sequence of variable deviations, which can be viewed as a series of treatments on the policy variable over multiple periods. This paper addresses this by providing a comprehensive methodology for such counterfactual exercises.

Our paper advances the literature by unifying the commonly used counterfactual analyses of hypothetical trajectory and policy intervention through the concept of policy path deviation. We develop a robust approach to evaluating economic performance under these counterfactual scenarios within a linear framework. The term robust signifies that our method does not require the specification of a rule equation for the policy variable (e.g., the Taylor Rule), nor the specification of a structural model such as a SVAR or Dynamic Stochastic General Equilibrium (DSGE) model. Instead, our approach focuses solely on identifying policy shocks, thereby circumventing the risk of misspecification inherent in both rule equations and structural models. In the context of monetary policy, the rule equation (e.g., the Taylor Rule) can be prone to misspecification due to issues such as omitted

[1] See, for example, Sims and Zha (1995), Bernanke et al. (1997), Waggoner and Zha (1999), Leeper and Zha (2003), Hamilton and Herrera (2004), Sims and Zha (2006), Kilian and Lewis (2011), Baumeister and Kilian (2014), Bańbura et al. (2015), Davis and Zlate (2019), Antolin-Diaz et al. (2021), Christian and Wolf (2024) and among them.

[2] See, for example, Blanchard and Quah (1989), Stock and Watson (2001), Jordà (2005), Dufour, Pelletier, and Renault (2006), Kilian and Kim (2011), Inoue and Kilian (2013), Ramey (2016), Kilian and Lütkepohl (2017), Stock and Watson (2018), Inoue and Kilian (2020), Montiel Olea and Plagborg-Møller (2021), Plagborg-Møller and Wolf (2021), Bruns and Lütkepohl (2022), Lusompa (2023), Jordà (2023), Herbst and Johannsen (2024), Li et al. (2024), Dufour and Wang (2024a) and among them.

variables, inadequate consideration of expectations, measurement errors, or nonlinearities (see Del Negro and Schorfheide (2009), Hamilton et al. (2011), Beaudry et al. (2023), among others). As a result, structural models that incorporate these rule equations are not immune to misspecification issues. The traditional approach to macroeconomic counterfactual analysis, such as conditional forecasting (see, e.g., Doan et al. (1984) and Waggoner and Zha (1999)) and policy intervention (see, e.g., Bernanke et al. (1997), Hamilton and Herrera (2004)), typically relies on SVAR models, which are susceptible to omitted variable bias. This concern has led to recent macroeconometric literature, such as Stock and Watson (2018) and Plagborg-Møller and Wolf (2022), where the underlying process is modelled as a structural vector moving average (SVMA) model. Our work, from this perspective, aligns with the rising popularity of the "sufficient statistics" approach to counterfactual analysis, see Chetty (2009), Barnichon and Mesters (2023), and McKay and Wolf (2023). However, our approach is more general, as previous studies often depend on specific rule equation or impose certain assumptions on the model, such as the separation between the policy and non-policy blocks.

This paper deliberately considers counterfactuals—hypothetical trajectory and policy intervention—in terms of policy path deviation. In contrast, recent literature, such as Beraja (2023) and McKay and Wolf (2023), examines the systematic change of policy rules by altering the rule parameter(s) in the policy equation. There are three key reasons for focusing on path deviations rather than rule changes. (1) In the context of monetary policy, central banks influence the market through short-term interest rates. It is more straightforward to work with the interest rate path, as it directly pertains to decisions regarding rate hikes, holds, or cuts. (2) A one-time shift in the Taylor rule parameter cannot arbitrarily manipulate the interest rate path. Conversely, our approach, which centers on path deviations, can directly address any scenario involving the policy variable path, thereby offering greater flexibility in practice. (3) The policy rule counterfactual proposed by McKay and Wolf (2023) requires the separation of policy and non-policy blocks and the existence of a counterfactual equilibrium. In contrast, our method does not necessitate this separation, making it more adaptable to a less restrictive data-generating process.

The counterfactual approach proposed in this paper makes five key contributions to the existing literature. First, we interpret policy path deviations between the baseline and counterfactual economies as arising solely from external interventions by policymakers through selected policy shocks:

$$\text{baseline economy} + \text{policy shocks} = \text{counterfactual economy}. \tag{1.1}$$

The idea of combining a sequence of hypothetical policy shocks with the baseline economy to mimic the counterfactual scenario is rooted in the thought experiment by Bernanke et al. (1997), who consider a hypothetical case where the interest rate remains unresponsive to an oil price shock. As they state:

"This procedure (setting the federal funds rate at zero level) is equivalent to combining the initial non-policy shock with a series of policy innovations just sufficient to offset the endogenous policy response."

Thus, the magnitude of the policy shocks is determined by a rule in which the policymaker selects shocks that best replicate the counterfactual scenario. This approach encompasses both the "period-by-period shocks" method proposed by Sims and Zha (1995) and Sims and Zha (2006), as well as the "multi-shock at initial date" method introduced by McKay and Wolf (2023). Moreover, our framework allows policymakers to select discretionary policy shocks, enabling targeted market interventions tailored to specific policy objectives.

Second, our robust counterfactual approach offers a unique analytical solution for evaluating the performance of hypothetical economies. This approach in the exercise of hypothetical trajectory presents a more robust alternative to Bayesian conditional forecasting methods (see, e.g., Waggoner and Zha (1999)), that may be sensitive to prior distributions, require intensive computational resources, and risk misspecification due to their reliance on VAR models. Additionally, our approach in the exercise of policy intervention is simpler and more robust against misspecification than the conventional recursive method founded on SVAR model (see, e.g., Bernanke et al. (1997), Hamilton and Herrera (2004), and Kilian and Lewis (2011)).

Third, we propose a novel identification approach based on linear projection with external instruments for the parameters used in our counterfactual analysis. While existing Local Projection-instrumental variable methods (see, e.g., Stock and Watson (2018)) are typically limited to impulse response estimation, our proposed linear projection model for policy intervention parameters incorporates future policy variables as controls. This allows for the identification of counterfactual causal effects that capture shock transmission while excluding the systematic response of policy variables. Consequently, the policy intervention parameter in the projection model is interpreted as a direct causal effect, capturing the impact of the shock of interest but eliminating the indirect effect from the policy variables. Conversely, the hypothetical trajectory parameters in the projection model capture the causal effects of policy variables on the outcome variable, with the inner product of the hypothetical trajectory parameters and the systematic response of the policy variables representing the indirect effect. Our approach thereby connects the two types of policy counterfactuals with the direct and indirect causal effects through a linear projection model.

Fourth, we provide statistical inference for the counterfactuals performance, while the existing robust counterfactual literature (see, e.g., Beraja (2023) and McKay and Wolf (2023)) has mostly overlooked statistical inference. In general, we derive a Delta-method based statistical inference for the parameter estimates computed from the analytical solution. Moreover, based on the proposed linear projection identification method, we also provide a simple and practical version of statistical inference for projection estimates. Considering the serial correlation of projection errors and the challenges associated with choosing kernel functions and bandwidths for heteroskedasticity and autocorrelation consistent (HAC) covariance estimates, we derive heteroskedasticity-consistent/robust (HC/HR) covariance matrix estimates, relying on techniques developed by Dufour and Wang (2024a).

Fifth, we extend our methodology to a nonlinear model founded on the potential outcome framework and provide a simulation algorithm for the counterfactual exercises. The nonlinear model setup helps us address Lucas' critique by allowing model parameters to change with the economic situation. The simulation algorithm is analogous to the Monte Carlo simulation for nonlinear impulse responses considered by Koop et al. (1996) and more generally discussed for statistical tests by Dufour (2006).

Relation to literature.—Our paper is closely related to the macroeconomic robust counterfactual literature, building on and distinguishing itself from the existing works in this area. For instance, both McKay and Wolf (2023) and Beraja (2023) perform counterfactual analyses of alternative policy rules without relying on specific parametric models. However, their approach, which necessitates the separation of policy and non-policy blocks, can result in non-unique or non-existent counterfactual equilibrium and lacks statistical inference for the resulting counterfactual outputs. Barnichon and Mesters (2023) explore a robust method for the optimality of the policy rule. Their focus is primarily on assessing the optimality of policy decisions, without considering the impact of a non-responsive policy variable under a given shock or incorporating projection methods for counterfactual analysis. Our approach also parallels the policy path exercises of Leeper and Zha (2003), yet we innovate by introducing a robust method coupled with a straightforward identification, estimation, and inference procedure based on projection-based method. Moreover, our simple counterfactual method complements the analyses of the "Lucas program" (as outlined by Christiano, Eichenbaum, and Evans, 1999) by eliminating the need to construct a micro-founded structural model.

Our hypothetical trajectory exercise is particularly relevant to scenario analysis (as discussed by Leeper and Zha (2003), Kilian and Lee (2014), and Baumeister and Kilian (2014)) and to conditional forecasts (Waggoner and Zha (1999); (Kilian and Lütkepohl, 2017, Section 4.4); Antolin-Diaz et al. (2021)). Unlike Bayesian methods, our estimation approach is grounded in frequentist statistics. The counterfactual exercise of policy intervention draws inspiration from mediation analysis in social sciences (Baron and Kenny 1986), it provides an appropriate framework to explore the role of specific variables in the propagation of structural shocks. Our method, which employs linear projection with external instruments, aligns with the instrument method to solve the problem of the mediator's endogeneity in the causal mediation analysis, see, e.g., Celli (2022).[3] Our linear projection method for policy intervention exercises incorporates future policy variables as controls. Including future controls in the projection model to capture direct effects is relevant to the structural interpretation of lagged "generalized impulse response coefficients" introduced by Dufour and Renault (1998) and extended by Dufour and Wang (2024b). This approach extends Sims' impulse response function in Local Projection, which typically includes only pre-treatment controls.

The counterfactual method we propose requires the identification of policy shocks, drawing on established macroeconometric literature such as Ramey (2016) and Kilian and Lütkepohl (2017). Projection-based estimation in macroeconometrics has a long-standing history, beginning with the works of Dufour and Renault (1998), Jordà (2005), Dufour, Pelletier, and Renault (2006), Montiel Olea and Plagborg-Møller (2021), Xu (2023), and Lusompa (2023). Our linear projection estimation method using external instruments closely relates to the Linear Projection-Instrumental Variable (LP-IV) method of Stock and Watson (2018), based on the SVMA model. Our approach utilizes external instruments, such as high-frequency financial market instruments for monetary policy shocks, as discussed

[3] A key assumption in causal mediation analysis is the "sequential ignorability assumption" for the mediator. This assumption requires that no confounder, influenced by the treatment, affects the relationship between the mediator and the outcome (see, e.g., Assumption 1 in Imai, Keele, and Yamamoto (2010)). This is a stringent condition, as it necessitates that nothing along the path from the treatment to the mediator also influences the outcome, as highlighted by Celli (2022). For example, to analyze mediation effects within a dynamic macroeconomic model, Dufour and Wang (2024b) address the ignorability assumption by incorporating post-treatment confounders within an SVAR framework.

in Gertler and Karadi (2015), Nakamura and Steinsson (2018), and Bauer and Swanson (2023). The projection inference, using an Eicker–Huber–White heteroskedasticity-robust covariance matrix estimate, parallels HAC-free local projection inference methods (see, e.g., Montiel Olea and Plagborg-Møller (2021), Xu (2023), Dufour and Wang (2024a)).

Empirical research.—We apply our theoretical approach to three empirical counterfactual exercises in the evaluation of monetary policy: (1) *historical scenario analysis*, (2) *future scenario analysis*, and (3) *policy intervention*.

In the *historical scenario analysis*, we evaluate the potential effects of a less aggressive interest rate hike implemented earlier in the post-pandemic period compared to the actual policy. Our counterfactual results suggest that such a policy could have temporarily reduced industrial production by 5%, with a subsequent rebound to the actual level by April 2024. The benefits, on the other hand, include a reduction in inflation by 1.5% in 2022, reaching a 2.5% level by January 2024.

We perform a *future scenario analysis* to examine the impact of a dovish or a hawkish monetary policy counterfactual that deviate from the Federal Reserve's Survey of Professional Forecasters (SPF). The dovish scenario, which assumes a 25 basis point lower interest rate than forecasted in the first month, could result in a 2.6% inflation rate—0.3% higher than the SPF prediction—while increasing industrial production by nearly 1.2% over one year. In contrast, the hawkish scenario, which maintains the interest rate within the 5.25-5.5% range for an additional quarter before lowering it, would lead to a 2% inflation rate—0.35% lower than the SPF projection—but would decrease industrial production by 1.3%.

In the *policy intervention* exercise, we conduct a thought experiment in which the Federal Reserve remains unresponsive to an oil price shock. Our analysis indicates that the Fed did not respond to rising oil prices until the inflation rate had already increased for several periods. The quantitative findings suggest that, in systematically addressing an oil price shock, the Fed effectively trades off two units of industrial production for each unit of inflation in its effort to stabilize the inflation rate potentially elevated by the oil price shock.

Outline.—The structure of this paper is as follows. Section 2 formalizes the concepts of policy intervention and hypothetical trajectory within a linear model framework. In Section 3, we provide a structural interpretation of macroeconomic counterfactuals, deriving both an analytical solution and the corresponding optimal policy trajectory. Section 4 introduces the IV-based linear projection identification methodology, while Section 5 develops the simple projection-based estimation and inference approach. Section 6 presents statistical inference for the closed-form estimates of the counterfactual parameters, utilizing the Delta method. Section 7 extends our procedure to accommodate counterfactual exercises within a nonlinear framework. Three empirical exercises of the U.S. monetary policy counterfactuals are detailed in Section 8. Finally, Section 9 concludes the paper.

2 Framework

This section establishes a linear framework and introduces two primary macroeconomic counterfactuals related to deviations from the policy path. The assumption of linearity is fundamental to our robust methodology, offering substantial advantages in both interpretation and computation. In Section 7, we will extend this analysis to a more general parametric model.

2.1 Linear structural model

Consider a three-dimensional variable $W_t = (X_t, Y_t, R_t)'$.[4] Suppose W_t is generated by a function of current and past shocks, $\varepsilon_{1:t}$, and a given initial condition at period zero, $\boldsymbol{w}_0$,

$$W_t = \boldsymbol{g}_t(\varepsilon_{1:t}; \boldsymbol{w}_0), \tag{2.1}$$

for $1 \leq t \leq T$, where $\boldsymbol{g}_t(\cdot)$ is a 3×1 vector function, $\varepsilon_{1:t}$ is a $nt \times 1$ stacked vector of shocks,

$$\varepsilon_{1:t} := (\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_t')',$$

and ε_t is a n-dimensional vector of uncorrelated structural shocks, both intratemporal and intertemporal, with zero mean and a full rank diagonal covariance matrix for dimension $n \geq 3$. Throughout this paper, let $W_t, X_t, Y_t, R_t, \varepsilon_t$ denote random variables, and $\boldsymbol{w}_t, x_t, y_t, r_t, \boldsymbol{\epsilon}_t$ denote their realized value. To accommodate the three-dimensional vector $W_t = (X_t, Y_t, R_t)'$, we partition the 3×1 vector function $\boldsymbol{g}_t(\cdot)$ as

$$\boldsymbol{g}_t(\cdot) := \left[g_{x,t}(\cdot), g_{y,t}(\cdot), g_{r,t}(\cdot)\right]'. \tag{2.2}$$

For notation simplicity, $\boldsymbol{g}_t(\varepsilon_{1:t}; \boldsymbol{w}_0)$ is used interchangeably with $\boldsymbol{g}_t(\varepsilon_{1:t})$.

The structural shock ε_t can be partitioned as $\varepsilon_t := (\varepsilon_{x,t}, \varepsilon_{r,t}', \varepsilon_{I,t})'$, where $\varepsilon_{x,t}$ represents the structural shock corresponding to variable X_t, $\varepsilon_{r,t}$ is the vector of policy shocks (which may reduce to a scalar) to the policy variable R_t, and $\varepsilon_{I,t}$ includes the remaining shocks. It shall be highlighted that allowing $\varepsilon_{r,t}$ to be a vector of shocks accommodates not only the counterfactual method founded on models with one shock per variable (Sims and Zha (1995) and (2006)), typical of standard SVAR models, as well as those incorporating multi-shock per variable[5]. The rationale for models with multiple shocks per variable is to facilitate the empirical Lucas-robust counterfactuals proposed by McKay and Wolf (2023). These counterfactuals are designed for incorporating ex-ante private-sector expectations, wherein the private sector is informed of the policy intervention planned at an initial date but realized in the future.

In particular, we employ a linear moving-average model, namely, Structural Vector Moving Average (SVMA) model[6]:

$$\boldsymbol{g}_t(\varepsilon_{1:t}; \boldsymbol{w}_0) = \Theta_0 \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \cdots + \Theta_{t-1}\varepsilon_1 + \boldsymbol{w}_0, \tag{2.3}$$

where Θ_h is the structural matrix at horizon h with dimensions $3 \times n$, such that $\Theta_h = \partial W_{t+h}/\partial \varepsilon_t'$. When the dimensions of the structural shocks and observables are identical,

[4] Readers who wish to have an empirical interpretation in mind may think of X_t as oil price, Y_t as output variable, and R_t as interest rate. This setting can be easily extended to include other macro-aggregates.

[5] For literature on identifying and estimating multiple monetary policy shocks, see Inoue and Rossi (2021), Jarociński (2024), and references therein.

[6] Alternately, an SVMA model without initial conditions can be written as $W_t = \Theta(L)\varepsilon_t$, where $\Theta(L) = \sum_{h=0}^{\infty} \Theta_h L^h$, L is the lag operator.

$n = 3$, and the roots of the moving average polynomial lie outside the unit circle, the SVMA model can be transformed into an SVAR model. The structural impulse response can then be computed recursively using VAR slope coefficient estimates, then the SVAR and Local Projection have the same estimand of impulse responses (Plagborg-Møller and Wolf (2021)). In contrast, if there are more shocks than observables ($n > 3$), the SVMA model falls into the category of non-recoverable SVMA, as described by Stock and Watson (2018). This classification implies that recovering all structural shocks as a rotation of the reduced-form VAR error is unlikely and and it could lead to the discrepancy between the impulse response estimands of the SVAR and Local Projection method. This distinction has recently garnered significant attention in macroeconometrics, as the assumption that the reduced-form VAR error spans the same space as the structural shocks is overly restrictive. It limits the model's ability to account for phenomena such as measurement error, omitted variable bias, and time-aggregated data.

Our linear framework can be extended to incorporate exogenous variables, which are commonly used to determine the 'state' of the economy. This extension is feasible because the linear model outlined is compatible with the SVAR model, which in turn can be further expanded to include the Vector Autoregression with exogenous variables (VARX) model. The inclusion of exogenous variables in a dynamic framework is equivalent to imposing specific restrictions on the equations governing these variables. As discussed by Kilian and Lütkepohl, 2017, Section 2, a VARX model is essentially a specialized form of a VAR model, where a VARX model effectively imposes Granger non-causality from endogenous to exogenous variables.

2.2 Hypothetical trajectory

Counterfactual analysis in macroeconomics involves exploring hypothetical trajectories of a policy variable that deviate from the baseline scenario. These analyses are essential for structural scenario evaluation. For instance, one might examine alternative economic outcomes had the central bank implemented an earlier interest rate hike during the post-pandemic period or assess the impact on output forecasts under a dovish or hawkish future interest rate policy.

Consider a baseline economy characterized by a sequence of observable variables generated by baseline shocks. Now, imagine a counterfactual scenario in which the policy variable (e.g., interest rate) deviates from its baseline trajectory starting at time t and extending to $t+H$, spanning $(H+1)$ periods. Denote $\boldsymbol{r}_{t+H} := (r_t, r_{t+1}, \cdots, r_{t+H})'$ as the stacked vector of the policy variable in the baseline economy, and $\tilde{\boldsymbol{r}}_{t+H} := (\tilde{r}_t, \tilde{r}_{t+1}, \cdots, \tilde{r}_{t+H})'$ as the corresponding vector in the counterfactual scenario,

$$\begin{array}{ll} \text{baseline economy:} & \boldsymbol{r}_{t:t+H} \\ \text{counterfactual economy:} & \tilde{\boldsymbol{r}}_{t:t+H}, \end{array} \tag{2.4}$$

and denote the police path deviation as,

$$\boldsymbol{d}_H^{(ht)} := \boldsymbol{r}_{t:t+H} - \tilde{\boldsymbol{r}}_{t:t+H}. \tag{2.5}$$

In practice, researchers typically assign values to $\tilde{\boldsymbol{r}}_{t:t+H}$ based on their specific research designs and objectives.

The objective of hypothetical trajectory analysis is to evaluate the impact of deviations in the policy path on other macroeconomic variables. Let Y_{t+h} represent the output variable of interest. Given that the baseline economy is observed, the goal is to derive the counterfactual output by examining the difference between the output variables in the baseline and counterfactual economies, denoted as

$$\psi_h := y_{t+h} - \tilde{y}_{t+h}, \tag{2.6}$$

where y_{t+h} and $\tilde{y}_{t+h}$ represent the realized output variables in the baseline and counterfactual economies, respectively. The output deviation, ψ_h, is thus impacted by the policy path deviation, $\boldsymbol{d}_H^{(ht)}$.

Given the data generating process (2.1), the variables in baseline and hypothetical economy are driven by a sequence of structural shocks, $\boldsymbol{\epsilon}_{1:T}$ and $\tilde{\boldsymbol{\epsilon}}_{1:T}$, respectively. They are the stacked vectors of realized shocks in the corresponding economy: $\boldsymbol{\epsilon}_{1:T} := (\boldsymbol{\epsilon}_1', \boldsymbol{\epsilon}_2', \cdots, \boldsymbol{\epsilon}_T')'$ and $\tilde{\boldsymbol{\epsilon}}_{1:T} := (\tilde{\boldsymbol{\epsilon}}_1', \tilde{\boldsymbol{\epsilon}}_2', \cdots, \tilde{\boldsymbol{\epsilon}}_T')'$. The central task in hypothetical trajectory analysis is to identify the sequence of counterfactual policy shocks based on the baseline observables and the hypothetical performance of policy variables. If the complete sequence of hypothetical shocks is known, determining the trajectory of the hypothetical economy becomes straightforward. However, challenges arise because typically, the number of policy variable deviations often exceeds the number of freely varying structural shocks, implying that the mapping from the path deviation to hypothetical shocks may not be one-to-one without imposing additional constraints.

It is essential to consider a specific structural interpretation of policy path deviation. In the context of conditional forecasting exercises (see, e.g., Waggoner and Zha (1999)), where all future shocks are allowed to deviate from their unconditional distribution to align with the conditional information, Jarociński (2010) noted that multiple solutions may exist, potentially leading to infinitely many solutions for the conditional shocks. Antolin-Diaz et al. (2021) provide unique approximated solutions for the deviation of the mean and variance of the future shocks from their unconditional counterparts by using the Moore-Penrose inverse. Additionally, Kilian and Lütkepohl, 2017, Section 4 present a scenario analysis approach in which only policy shocks deviate from their unconditional distribution, while all other shocks remain within their unconditional distribution. Consequently, when determining hypothetical shocks is crucial to counterfactual analysis, different structural interpretations of policy path deviation can result in distinct hypothetical shocks, potentially leading to substantial differences in the economic outcomes within the hypothetical scenario.

Historical Scenario Analysis

The historical scenario analysis is frequently employed in policy evaluation to assess how the economy might have performed under an alternative policy path. In this scenario, the baseline economy is represented by historical data, where observables are driven by a sequence of historical shocks. In contrast, the hypothetical economy, characterized by observables $\tilde{\boldsymbol{w}}_t$, is driven by counterfactual shocks $\tilde{\boldsymbol{\epsilon}}_{1:t}$. When researchers apply a structural parametric model, such as SVAR, they can recover historical shocks through consistent model estimation. The primary challenge in this exercise lies in determining the counterfactual shocks, $\tilde{\boldsymbol{\epsilon}}_{1:t}$. Once these counterfactual shocks are identified under specific restrictions, the performance of other output variables in the counterfactual economy can be derived, enabling a comprehensive analysis of hypothetical scenarios and their broader economic implications.

For example, Kilian and Lee (2014) employ historical decomposition to perform counterfactual analysis. They use an SVAR model to decompose each observable into functions of historical structural shocks and then conduct counterfactuals by removing the cumulative contribution of a specific structural shock to the variables at a given point in time. In their counterfactual economy, the shock of interest is set to zero, $\tilde{\epsilon}_{x,t} = 0$, for all t, while the remaining shocks are kept equal to their baseline counterparts, $\tilde{\boldsymbol{\epsilon}}_{-x,t} = \boldsymbol{\epsilon}_{-x,t}$. This approach is valuable for identifying how changes in variables over time can be attributed to the removal of a particular structural shock.

Future Scenario Analysis

In future scenario analysis, the baseline economy consists of unconditional forecasts for macroeconomic aggregates, while the counterfactual economy represents a conditional forecast in which the policy variable is constrained to follow a predetermined path distinct from its baseline counterpart. The baseline economy is driven by historical shocks, $\boldsymbol{\epsilon}_{1:t-1}$, with future shocks fixed at their mean value, $\boldsymbol{\epsilon}_{t:T} = 0$. The hypothetical scenario, generated by a sequence of hypothetical shocks, reflects an alternative economy where the policy variable deviates. Typically, this deviation arises from future shocks diverging from their mean value, while past shocks in the counterfactual economy are identical to those in the baseline, $\tilde{\boldsymbol{\epsilon}}_{1:t-1} = \boldsymbol{\epsilon}_{1:t-1}$. The counterfactual policy trajectory imposes an implicit restriction on future shocks, such that $g_{r,t+h}(\boldsymbol{\epsilon}_{1:t-1}, \tilde{\boldsymbol{\epsilon}}_{t:t+H}) = \tilde{r}_{t+h}$ for all integer $h \in [0, H]$. Once these future shocks are recovered from the imposed restrictions, the counterfactual economy can be constructed.

Conventional conditional forecasting is a widely used method for future scenario analysis, often employing Bayesian techniques. This approach involves obtaining the posterior distribution, which represents updated beliefs about the parameters (see Doan et al. (1984) and Waggoner and Zha (1999)). Within our framework, the deviation of the policy variable path can be interpreted as a "hard conditional forecast." In this context, $\boldsymbol{r}_{t:t+H}$, as defined in (2.4), represents the mean of the unconditional forecast, while $\tilde{\boldsymbol{r}}_{t:t+H}$, also defined in (2.4), serves as the hard condition imposed on future policy variables. The conditional forecasting method then updates the distribution of both the parameters and the shocks by incorporating this conditional information.

Alternative forecast scenarios, as implemented by Baumeister and Kilian (2014), utilize scenarios derived from the structural moving average representation of the VAR model. Unlike Bayesian conditional forecasting, which relies on posterior distributions to update parameter beliefs, their method adopts a frequentist approach. This involves directly manipulating the sequence of future structural shocks while keeping the model parameters fixed. For the baseline variable, Baumeister and Kilian set it as the unconditional forecast by assuming all future baseline shocks are zero. They then intentionally adjust the sequence of structural shocks within the range of historical experience. This approach enables them to explore the consequences of certain historical events recurring in economic forecasts. They argue that the critique by Lucas Jr (1976) does not apply to their counterfactuals, as their forecast scenario analysis is grounded in shock sequences directly based on historical precedent.

2.3 Policy intervention

The evaluation of policy intervention in macroeconomics requires an investigation into the extent to which a policy variable contributes to a dynamic causal effect by constructing a counterfactual scenario in which the policy variable is held unresponsive. A prominent example is the seminal study by Bernanke et al. (1997) (hereafter, BGW), which investigated whether the systematic response of interest rates was the primary cause of the recession following the oil crisis in the 1970s. BGW demonstrated that systematic monetary policy responses played a significant role in contributing to the 1970s recession in the United States. Their claim is based on constructing a hypothetical scenario where the federal funds rate remains constant throughout the transmission of the oil price shock. This approach, often referred to as zeroing-out policy intervention, underscores the benefits of assessing dynamic causal effects within a counterfactual framework where the endogenous response of specific variables is nullified.[7]

The 'zeroing-out' policy intervention is often referred to as the direct effect in social science mediation analysis, as it isolates the causal impact by eliminating the indirect influence of certain endogenous variables.[8] This approach recognizes that within a dynamic system, the impulse response function reflects the total effect, fully capturing the impact when these endogenous variables are allowed to respond to the shock of interest. Consequently, the total effect can be decomposed into direct and indirect effects (Baron and Kenny, 1986; Imai et al., 2010, 2013). By suppressing the endogenous response of these variables to the shock, the indirect effect is removed, allowing for an assessment of the direct contribution of these variables to the overall causal effect.

The policy intervention fundamentally relies on impulse response functions with restrictions imposed on certain endogenous variables. To clarify, we first define the impulse response function as follows:

$$\boldsymbol{\theta}_{x,h} := W_{t+h}(1) - W_{t+h}(0), \tag{2.7}$$

where $\boldsymbol{\theta}_{x,h}$ is a 3×1 impulse response vector, partitioned as $\boldsymbol{\theta}_{x,h} = [\theta_{xx,h}, \theta_{yx,h}, \theta_{rx,h}]'$, and $W_{t+h}(d)$ is a three-dimensional treated/untreated random variable defined as

$$W_{t+h}(d) := g_{t+h}\left([\varepsilon'_{1:t-1}, \epsilon_{x,t} = d, \varepsilon'_{-x,t}, \varepsilon'_{t+1:t+h}]'\right), \tag{2.8}$$

for $d = 0, 1$, where the notation $\varepsilon_{-x,t}$ represents the vector ε_t with the element $\varepsilon_{x,t}$ excluded. For the purpose of subsequent discussion, we partition $W_{t+h}(d)$ as

$$W_{t+h}(d) = [X_{t+h}(d), Y_{t+h}(d), R_{t+h}(d)]'. \tag{2.9}$$

[7] Kilian and Lewis (2011) conduct an alternative thought experiment to evaluate whether the endogenous response to oil price shocks is the primary cause of subsequent economic recessions. They depart from BGW's approach by not imposing the "zeroing-out" constraints on the policy interest rate. Their argument challenges the interpretation that completely holding the interest rate unresponsive to oil price shocks represents an implausible counterfactual scenario with near-zero probability. Instead, they propose an alternative method to interpret this "non-responsiveness." In their SVAR framework, they impose period-by-period constraints on the interest rate equation, assuming that all coefficients of contemporaneous or lagged oil price variables are set to zero. Ultimately, they interpret the interest rate path deviation from its equilibrium response as the effect of period-by-period monetary policy shocks and compute the performance of the economy in this counterfactual scenario.

[8] Leeper and Zha (2003) refer to conditional projections as the direct effects of policy when evaluating alternative policy actions.

The definition provided in (2.7) offers a causal interpretation of impulse responses, as discussed in Bojinov and Shephard (2019) and Rambachan and Shephard (2021), as it is grounded in the potential outcome framework. This setting effectively bridges the conceptual gap between macroeconomic impulse responses and microeconomic treatment effects.[9] Unlike the classical definition of impulse response, typically based on the difference between two forecasts (see, e.g., Hamilton (1994) and Jordà (2005)), the definition in (2.7) adheres to the principle of ceteris paribus, capturing the difference in the output variable between two economies subject to an identical sequence of random structural shocks, where one economy experiences a specific shock of interest valued at 1, while the other experiences it at 0.

It is crucial to note that the term $Y_{t+h}(d)$ remains random because only the shock of interest, $\epsilon_{x,t}$, is assigned a realization value of d, while all other shocks continue to be random variables. In the context of a linear moving average model, as defined in (2.3), they cancel each other out due to the linearity-constant coefficients, and therefore the impulse response $\theta_{yx,h}$ is a constant term. Conversely, in the case of a nonlinear function, $\theta_{yx,h}$ may be stochastic and dependent on the sequence of random shocks. This distinction highlights the differing behaviours of linear and nonlinear models in response to shocks.

Our objective in analyzing 'zeroing-out' policy intervention is to evaluate the response of the output variable to a one-time impulse in the shock of interest, under the assumption that the policy variable remains unresponsive from period t to period $t+H$. To formalize this discussion, we introduce the following notation:

$$Y_{t+h}\{1, R_{t:t+H}(0)\}, \tag{2.10}$$

where

$$R_{t:t+H}(0) := \left[R_t(0), R_{t+1}(0), \dots, R_{t+H}(0)\right]',$$

and $R_{t+h}(0)$ is defined in (2.9). The expression $Y_{t+h}\{1, R_{t:t+H}(0)\}$ represents the value of the output variable when the shock of interest, $\varepsilon_{x,t}$, occurs as value of $\epsilon_{x,t} = 1$, while the policy variables from period t to $t+H$ remain unresponsive, as if the shock of interest had not occurred, i.e., $\epsilon_{x,t} = 0$. This definition is closely related to the concept of direct effects in mediation analysis (e.g., see Imai et al. (2013)). Our policy intervention effect aims to recover the causal effect of the shock of interest, excluding the indirect effects mediated through the policy variables.

In situations where the policy variables exhibit the same response pattern as the output variable, we denote this equivalence as

$$Y_{t+h}\{d, R_{t:t+H}(d)\} \equiv Y_{t+h}(d), \tag{2.11}$$

for $d = 0, 1$, where $Y_{t+h}(d)$ is defined in (2.9). This notation reflects the structural interpretation that the policy variable responds to the shock of interest in a manner consistent with its natural dynamics.

As illustrated, the zeroing-out policy intervention captures the causal effect of the shock of interest on the output variable while keeping the policy variable unresponsive. Therefore,

[9] See, e.g., the potential-outcome framework for time series causal research in the context of monetary policy (Angrist and Kuersteiner (2011), Angrist et al. (2018)).

the policy intervention parameter is defined as

$$\phi_h := Y_{t+h}\{1, R_{t:t+H}(0)\} - Y_{t+h}\{0, R_{t:t+H}(0)\}. \tag{2.12}$$

This definition can be interpreted through a two-step decomposition,

$$\begin{aligned} \phi_h =& \big(Y_{t+h}\{1, R_{t:t+H}(1)\} - Y_{t+h}\{0, R_{t:t+H}(0)\}\big) \\ &- \big(Y_{t+h}\{1, R_{t:t+H}(1)\} - Y_{t+h}\{1, R_{t:t+H}(0)\}\big). \end{aligned} \tag{2.13}$$

In the first step, the economy undergoes an one-time shock; in the second step, the policy-maker intervenes to suppress the policy variable, transitioning it from a responsive state to an unresponsive one.

The first step, within the linear framework of (2.3), involves measuring the impulse response function as defined in (2.11).

$$\text{Step 1 (one-time shock)} \;:\quad Y_{t+h}\{1, R_{t:t+H}(1)\} - Y_{t+h}\{0, R_{t:t+H}(0)\} = \theta_{yx,h}, \tag{2.14}$$

where $\theta_{yx,h}$ is defined in (2.7). The identification of ϕ_h, however, critically depends on the structural interpretation of the second step. In this step, the policy variable is suppressed, up to period H, transitioning from a responsive to an unresponsive state in reaction to a unit-sized shock. This process entails comparing two scenarios: the baseline economy, where the shock of interest propagates freely through the system, and a hypothetical economy, where the sequence of policy variables is artificially constrained to behave as though unaffected by the initial shock.

$$\text{Step 2 (policy variable muted)} \;:\quad \begin{array}{ll} \text{baseline economy:} & R_{t:t+H}(1) \\ \text{counterfactual economy:} & R_{t:t+H}(0) \end{array} \tag{2.15}$$

The disparity in the policy path between the baseline and hypothetical economies arises from the transition in status from responsiveness, $R_{t:t+H}(1)$, to unresponsiveness, $R_{t:t+H}(0)$. This vector of the deviation is denoted as:

$$\boldsymbol{d}_H^{(po)} := R_{t:t+H}(1) - R_{t:t+H}(0), \tag{2.16}$$

The deviation, $R_{t:t+H}(1) - R_{t:t+H}(0)$, is a vector of stacked impulse response functions under the linear model framework defined in (2.3). Given the impulse response functions defined in (2.7), the deviation can be rewritten as

$$\boldsymbol{d}_H^{(po)} = (\theta_{rx,0}, \theta_{rx,1}, \cdots, \theta_{rx,H})' \tag{2.17}$$

where $\theta_{rx,h}$ is the scalar impulse response of the policy variable to the shock of interest, defined in (2.7).

The effect of policy intervention underscores the dynamic causal relationship that emerges when a policy variable deviates from its equilibrium path to a zero response. This "zero-out" exercise highlights the pivotal role of the policy variable in transmitting the shock of interest. Much like hypothetical trajectory analyses, this thought experiment emphasizes the importance of the structural interpretation of such deviations. The following section provides a structural interpretation of the policy path deviation and presents an analytical solution for the policy intervention effect.

3 Structural interpretation and analytical solution

In this section, we offer a structural interpretation of the disparity between the baseline economy and the hypothetical economy and derive an analytical solution for the parameters of interest in two macroeconomic counterfactuals.

3.1 Structural interpretation

As discussed in Section 2.2, the hypothetical trajectory exercise illustrates how the policy variable deviates over certain periods from its baseline path, $\boldsymbol{r}_{t:t+H}$, to a hypothetical one, $\tilde{\boldsymbol{r}}_{t:t+H}$, as defined in (2.4). In contrast, the policy intervention exercise imposes a restriction on the policy variable, requiring it to revert from a responsive state, $R_{t:t+H}(1)$, to an unresponsive state, $R_{t:t+H}(0)$, as defined in (2.15). The disparity between the two economies in

	baseline	counterfactual	policy path deviation
hypothetical trajectory	$\boldsymbol{r}_{t:t+H}$	$\tilde{\boldsymbol{r}}_{t:t+H}$	$\boldsymbol{d}_H^{(ht)}$
policy intervention	$R_{t:t+H}(1)$	$R_{t:t+H}(0)$	$\boldsymbol{d}_H^{(po)}$

our structural interpretation is attributed to exogenous interventions by policymakers. This interpretation has a longstanding history in economic and social science research, dating back to Box and Tiao (1975). In our study, we examine scenarios where a sequence of policy variables deviates from their baseline path, interpreting these deviations as a result of such exogenous interventions. We offer an analytical solution to address and explain these deviations.

Note that the policy path deviation, $\boldsymbol{d}_H^{(ht)}$, for a hypothetical trajectory is a constant vector, arbitrarily determined by researchers. In contrast, the deviation $\boldsymbol{d}_H^{(po)}$ for policy intervention represents the endogenous responses of the policy variable to the initial shock of interest. As a result, it includes a sequence of impulse response functions and must be estimated in practical applications.

We consider the scenario where policymakers intervene in the market through specific policy shocks, constrained by the limits of their institutional powers. This limitation arises from the common belief that the central bank can only influence the market through the monetary policy rate, i.e., the prime interest rate. In our dynamic model, the policy variable is denoted by R_t, and the corresponding policy shocks are represented by $\varepsilon_{r,t}$. The disparity in the variable from period t to period $(t+H)$ is attributed to the manipulation by policymakers through the selected policy shocks,

$$\{\varepsilon_{t,H,(S)}\} \subseteq \{\varepsilon_{r,t}, \varepsilon_{r,t+1}, \cdots, \varepsilon_{r,t+H}\}. \tag{3.1}$$

where the dimension of the selected policy shocks, $\dim(\varepsilon_{t,H,(S)}) = n_e$. In the existing literature, there are typically two approaches to selecting policy shocks. One approach, considered by Sims and Zha (1995), involves period-by-period shocks $\varepsilon_{t,H,(S)} = (\varepsilon_{r,t}, \varepsilon_{r,t+1}, \cdots, \varepsilon_{r,t+H})'$, where $\varepsilon_{r,t}$ is reduced to a scalar structural shock. The other approach, discussed in McKay and Wolf (2023), involves applying multiple shocks at the initial date, where $\varepsilon_{t,H,(S)} = \varepsilon_{r,t}$, with $\varepsilon_{r,t}$ being a vector of policy shocks. A detailed discussion of these policy shock choices is provided in Discussion 1 below.

As specified in our framework of (2.1), the variables are driven by a sequence of shocks. The linear model framework enables us to derive analytical solutions for the parameters of interest in two counterfactual exercises without the need to specify a particular structural model. Given the structural interpretation that the variable disparity between the two economies is solely attributed to the selected policy shocks, we can represent the two economies as follows:

$$\begin{array}{lll} \text{Baseline economy } (W_{t+h}): & \{\varepsilon_{t,H,(S)} & , \quad U_{t+H}\} \\ \text{Counterfactual economy } (\tilde{W}_{t+h}): & \{\varepsilon_{t,H,(S)} - \boldsymbol{\delta}_H & , \quad U_{t+H}\} \end{array} \tag{3.2}$$

where $\boldsymbol{\delta}_H$ denotes the $n_e \times 1$ vector of the difference of the selected policy shocks between the baseline and counterfactual economy, and U_{t+H} represents the remaining shocks up to period $t+H$, $\{U_{t+H}\} = \{\varepsilon_{1:t+H}\} \backslash \{\varepsilon_{t,H,(S)}\}$.

Due to the linearity assumption imposed on the data generating process in (2.3), it is straightforward to derive the disparity between the variables from two economies:

$$W_{t+h} - \tilde{W}_{t+h} = \Theta_{e,h} \boldsymbol{\delta}_H, \tag{3.3}$$

where $\Theta_{e,h} := \partial W_{t+h} / \partial \varepsilon'_{t,H,(S)}$ and we partition $\Theta_{e,H}$ comfortably with three-dimensional vector $W_t = (X_t, Y_t, R_t)'$,

$$\Theta_{e,h} = \left[\boldsymbol{\theta}'_{xe,h}, \boldsymbol{\theta}'_{ye,h}, \boldsymbol{\theta}'_{re,h} \right]'. \tag{3.4}$$

In the next subsection, we provide a method to uniquely determine the size of the intervention with an analytical solution under our structural interpretation.

Discussion 1: (Policy shocks selection and the Lucas critique) The selection of policy shocks is potentially vulnerable to Lucas' critique (1976). The classical approach, formalized by Sims and Zha (1995) and Sims and Zha (2006), treats $\varepsilon_{r,t}$ as a scalar variable, consistent with the majority of empirical macroeconomic literature on SVAR models. In this context, the selected policy shocks, denoted as $\varepsilon_{t,H,(S)} = (\varepsilon_{r,t}, \varepsilon_{r,t+1}, \cdots, \varepsilon_{r,t+H})'$, represent a sequence of interventions designed to bridge the gap between the baseline and hypothetical economies on a period-by-period basis. However, this methodology has been criticized by macroeconomic theorists for its lack of ex ante consideration of private sector expectations. If the private sector anticipates the planned interventions, it is unlikely to delay its response until future periods; rather, it will adjust its behavior immediately at the initial date of the intervention.

This recognition has prompted the exploration of multi-shock interventions at the initial date, where $\varepsilon_{t,H,(S)} = \varepsilon_{r,t}$, with $\varepsilon_{r,t}$ considered as a vector of policy shocks, as discussed in McKay and Wolf (2023). This approach partially addresses Lucas' critique by allowing the private sector to respond immediately to the intervention. While this multi-shock approach deviates from the traditional SVAR literature, which typically considers a single shock per variable, recent research on monetary policy has made significant strides in empirically identifying multi-shocks, particularly concerning interest rates. Notable examples include the work by Inoue and Rossi (2021), which introduces new methods for identifying such multi-shocks. This paper does not advocate for one policy shock selection method over another; rather, we present a general approach for conducting counterfactual exercises.

Discussion 2: (Parameter Stability and the Lucas Critique) We consider a general linear model, specifically an SVMA, for our counterfactual exercises. A central tenet of Lucas' critique is that it is naïve to predict the effects of a change in economic policy based solely on historical observations. In essence, private sector agents will adjust their behavior in response to policy changes, leading to shifts in the underlying economic model. In the empirical macroeconomics literature, nonlinear models such as threshold VAR, smooth-transition VAR, and time-varying VAR are potential candidates to address this issue from an econometric perspective (see Koop, Pesaran, and Potter (1996), Pesaran and Shin (1998), Gonçalves, Herrera, Kilian, and Pesavento (2021)). In Section 7, we move beyond the linear model and incorporate the concept of a nonlinear model to provide a simulation-based procedure for the counterfactual exercises.

3.2 Analytical solution

In this subsection, we derive a unique solution for the magnitude of shock deviation between the two economies. We assume that the difference in performance between the baseline and counterfactual economies only contributes to the deviation of the selected policy shocks. Based on equation (3.3), the policy path period t to $t+H$ is expected to deviate as follows:

$$R_{t:t+H} - \tilde{R}_{t:t+H} = \Theta_{re,H}\boldsymbol{\delta}_H, \tag{3.5}$$

where $\Theta_{re,H}$ is an $(H+1)\times n_e$ impulse response matrix of $R_{t:t+H}$ to $\varepsilon_{t,H,(S)}$,

$$\Theta_{re,H} = [\boldsymbol{\theta}_{re,0}, \boldsymbol{\theta}_{re,1}, \cdots, \boldsymbol{\theta}_{re,H}]', \tag{3.6}$$

and $\boldsymbol{\theta}_{re,h}$ is defined in (3.4).

As demonstrated, the hypothetical trajectory and policy intervention are characterized by deviations in the policy variables, as defined by $\boldsymbol{d}_H^{(ht)}$ and $\boldsymbol{d}_H^{(po)}$, respectively. The magnitude of these shock deviations can be determined using equation (3.5), where the left-hand side of the equation is substituted with the vector that represents the specific policy path deviation in each scenario. Assuming that the matrix $\Theta_{re,H}$ is full-rank and square, this indicates that the number of manipulable policy shocks matches the number of deviated policy variables. Consequently, the magnitude of the shock deviation can be readily derived as

$$\boldsymbol{\delta}_H^{(\cdot)} = \Theta_{re,H}^{-1}\boldsymbol{d}_H^{(\cdot)}. \tag{3.7}$$

where ${}^{(\cdot)}$ is ${}^{(po)}$ for policy intervention exercise and ${}^{(ht)}$ for hypothetical trajectory exercise.

It is important to note that a full-rank square matrix $\Theta_{re,H}$ represents a special case. In a more general context, the matrix $\Theta_{re,H}$ may be singular, for instance, when the number of policy shocks manipulable by the policymaker is fewer than the dimensions of the policy path deviation vector. To resolve this, we utilize the Moore-Penrose inverse to derive a unique solution for the shock deviation vector,

$$\boldsymbol{\delta}_H^{(\cdot)} = \Theta_{re,H}^{-}\boldsymbol{d}_H^{(\cdot)}. \tag{3.8}$$

where the notation A^{-} represents the Moore–Penrose inverse of any matrix A. By utilizing the Moore–Penrose inverse, the vector of the shock deviation can be uniquely determined even when $\Theta^{(\cdot)}_{re,H}$ is non-square. This approach is equivalent to optimizing a general target function for the vector of the shock deviation,

$$\boldsymbol{\delta}_H^{(\cdot)} = \underset{\boldsymbol{\delta}_H}{\operatorname{argmin}} \left\{ \|\boldsymbol{d}_H^{(\cdot)} - \Theta_{re,H}\boldsymbol{\delta}_H\| \text{ subject to } \|\boldsymbol{\delta}_H\| \right\}. \tag{3.9}$$

This implies that the optimal shock deviation is the one that minimizes the Euclidean norm among all possible best-fitted policy path deviation. The economic interpretation for three specific cases is outlined below.

(i) Too few shocks ($\text{rank}(\Theta_{re,H}) < (H+1)$): The policymaker manipulates the available shocks as effectively as possible to replicate the counterfactual path. However, due to the limited number of shocks they can control, they are unable to perfectly replicate the path and must instead select the closest possible approximation.

(ii) Exact shocks ($\text{rank}(\Theta_{re,H}) = (H+1) = n_e$: The policymaker has just enough policy tools to replicate the counterfactual path exactly.

(iii) Too many shocks ($\text{rank}(\Theta_{re,H}) = (H+1) < n_e$): The policymaker has more than sufficient tools. They replicate the counterfactual path precisely in the least "surprising" way to the private sector. This "least surprising" approach is characterized by minimizing the Euclidean norm of the shock deviation.

3.2.1 Hypothetical trajectory

In the analysis of hypothetical trajectories, the discrepancy vector for the policy variable is denoted by $\boldsymbol{d}_H^{(ht)}$, as defined in (2.5). According to (3.7), the magnitude of the hypothetical policy shocks in this context is explicitly determined by:

$$\boldsymbol{\delta}_H^{(ht)} = \Theta^{-}_{re,H}\boldsymbol{d}_H^{(ht)}. \tag{3.10}$$

The objective of interest, ψ_h, in this counterfactual can be calculated using (3.3), specifically as $\psi_h = \boldsymbol{\theta}'_{ye,h}\boldsymbol{\delta}_H^{(ht)}$. Substituting $\boldsymbol{\delta}_H^{(ht)}$ with $\Theta^{-}_{re,H}\boldsymbol{d}_H^{(ht)}$, we obtain:

$$\psi_h = \boldsymbol{\beta}'_{h,H}\boldsymbol{d}_H^{(ht)}, \tag{3.11}$$

where $\boldsymbol{\beta}_{h,H}$ is a $(H+1)\times 1$ vector of parameters that represents the mapping of the policy variable deviation to the corresponding output variable at horizon h:

$$\boldsymbol{\beta}_{h,H} := \Theta^{-'}_{re,H}\boldsymbol{\theta}_{ye,h}. \tag{3.12}$$

We refer to $\boldsymbol{\beta}_{h,H}$ as the *hypothetical trajectory parameter*.

This expression underscores a key point: in counterfactual analyses of hypothetical trajectories, it is sufficient to know the counterfactual parameter $\boldsymbol{\beta}_{h,H}$ and the policy variable discrepancy $\boldsymbol{d}_H^{(ht)}$. Since the hypothetical trajectory parameter $\boldsymbol{\beta}_{h,H}$ is a nonlinear transformation of the impulse responses to policy shocks, identifying the policy shocks and the policy variable discrepancy is sufficient for conducting these counterfactual analyses, without the need to specify the entire structural model, such as an SVAR model.

3.2.2 Policy intervention

The policy intervention effect, as described in (2.12), involves two key steps. The first step captures the impulse response function, while the second step measures the impact of altering the responsive status of the policy variables.

Altering the responsive status results in a deviation in the policy path, representing the shift from a responsive to an unresponsive state, defined as $\boldsymbol{d}_H^{(po)}$ in (2.16). It has an explicit form displayed in (2.17) under a linear framework. Consequently, (3.7) specifies that the magnitude of the hypothetical policy shocks in this policy intervention scenario is given by:

$$\boldsymbol{\delta}_H^{(po)} = \Theta_{re,H}^{-} \boldsymbol{d}_H^{(po)}. \tag{3.13}$$

Thus, the two steps involved in the policy intervention can be summarized as follows:

$$\begin{aligned} \text{Step 1}: \quad & Y_{t+h}\{1, R_{t:t+H}(1)\} - Y_{t+h}\{0, R_{t:t+H}(0)\} = \theta_{yx,h}, \\ \text{Step 2}: \quad & Y_{t+h}\{1, R_{t:t+H}(1)\} - Y_{t+h}\{1, R_{t:t+H}(0)\} = \boldsymbol{\theta}'_{ye,h} \boldsymbol{\delta}_H^{(po)}, \end{aligned} \tag{3.14}$$

where Step 2 follows from (3.3) and (3.13). The economic interpretation of Step 2 involves eliminating the endogenous response of the policy variable by shifting the policy variable path from $R_{t:t+H}(1)$ to $R_{t:t+H}(0)$. This shift is induced by a policy shock deviation of size $\boldsymbol{\delta}_H^{(po)}$, and its impact on the output variable of interest is determined by the mapping encapsulated in the impulse response function $\boldsymbol{\theta}_{ye,h}$.

By combining the causal effects of both steps, the policy intervention effect can be explicitly expressed as:

$$\phi_h = \theta_{yx,h} - \boldsymbol{\theta}'_{ye,h} \boldsymbol{\delta}_H^{(po)}. \tag{3.15}$$

Substituting $\boldsymbol{\delta}_H^{(po)}$ with its explicit form from (3.13), we obtain:

$$\phi_h = \theta_{yx,h} - \boldsymbol{\beta}'_{h,H} \boldsymbol{d}_H^{(po)}, \tag{3.16}$$

where $\boldsymbol{\beta}_{h,H}$ is defined in (3.11). The economic interpretation of (3.16) is that the policy intervention effect is derived by subtracting the impact of the endogenous response of the policy variables, $\boldsymbol{\beta}'_{h,H} \boldsymbol{d}_H^{(po)}$, from the impulse response function, $\theta_{yx,h}$. Here, $\boldsymbol{d}_H^{(po)}$ represents the endogenous response of the policy variable to the shock of interest, and its impact on the output variable is captured by the hypothetical trajectory parameter $\boldsymbol{\beta}_{h,H}$.

This expression decomposes the impulse response ($\theta_{yx,h}$, total effect) into two components: the causal effect excluding the endogenous responses of the policy variable (ϕ_h, direct effect) and the causal effect attributed solely to the endogenous responses of the policy variable ($\boldsymbol{\beta}'_{h,H} \boldsymbol{d}_H^{(po)}$, indirect effect),

$$\underset{(\theta_{yx,h})}{\text{total effect}} = \underset{(\phi_h)}{\text{direct effect}} + \underset{(\boldsymbol{\beta}'_{h,H} \boldsymbol{d}_H^{(po)})}{\text{indirect effect}}. \tag{3.17}$$

The analytical solution provided here offers a robust approach to conducting counterfactual exercises, effectively isolating the indirect effect of certain endogenous policy variables. This method requires only the identification of policy shocks for the parameter

$\beta_{h,H}$, as well as the identification of the initial shock of interest for the impulse response function $\theta_{yx,h}$ and the endogenous response of policy variables $\boldsymbol{d}_H^{(po)}$. Traditional methods typically involve constructing a structural model, such as an SVAR, and recursively suppressing the endogenous response of the policy variable through policy shocks period by period. Our approach bypasses these complexities, providing a direct expression to measure both direct and indirect effects without the need to specify the entire structural model. This simplification facilitates direct estimation and statistical inference, which will be discussed in Sections 5 and 7.

3.3 Desired policy path

In this subsection, we provide an analytical solution to determine the optimal policy path, building on the expression derived from the hypothetical trajectory exercise. We consider two scenarios for the desired policy path: (1) when the policymaker has a specific target trajectory for an output variable, and (2) when the policymaker seeks to maximize the utility function associated with the output variable.

Ideal output trajectory

Consider $\boldsymbol{y}_{t:t+H}$ as the vector of output variables from period t to $t+H$ in the baseline economy, with $\tilde{\boldsymbol{y}}^*_{t:t+H}$ representing the ideal trajectory the policymaker aims to achieve in a counterfactual scenario. As previously illustrated, the disparity between observable variables in the baseline and counterfactual economies is attributed to deviations in selected policy shocks. Our objective now is to identify the optimal policy path that corresponds to the desired deviation in the output variables.

Given the linear data-generating process and (3.3), the ideal trajectory for the output variables in the counterfactual economy requires a policy shock intervention, denoted by $\boldsymbol{\delta}^*_H$, which can be computed as follows::

$$\boldsymbol{\delta}^*_H = \Theta^-_{ye,H}(\boldsymbol{y}_{t:t+H} - \tilde{\boldsymbol{y}}^*_{t:t+H}), \tag{3.18}$$

where $\Theta_{ye,H} := \partial Y_{t:t+H}/\partial \varepsilon'_{t,H,(S)} = [\boldsymbol{\theta}_{ye,0}, \boldsymbol{\theta}_{ye,1}, \cdots, \boldsymbol{\theta}_{ye,H}]'$ and $\boldsymbol{\theta}_{ye,h}$ is defined in (3.4). This approach is analogous to determining the magnitude of shock deviation corresponding to a hypothetical policy path, as outlined in (3.8). Previously, the counterfactual exercise centered on identifying the specific policy shock deviation required to align with the counterfactual policy path. The current task builds on this by determining the unique policy shock deviation necessary to achieve the counterfactual output trajectory.

As a result, the desired policy path $\tilde{\boldsymbol{r}}^*_{t:t+H}$ can be computed as follows:

$$\tilde{\boldsymbol{r}}^*_{t:t+H} = \boldsymbol{r}_{t:t+H} - \Theta_{re,H}\boldsymbol{\delta}^*_H, \tag{3.19}$$

where $\Theta_{re,H}$ is defined in (3.5).

Maximizing utility function

If the policymaker seeks to maximize the performance of output variables based on a given utility function, the necessary policy shock intervention can be derived, enabling the computation of the desired policy path. To illustrate this, consider a quadratic utility function:

$$U(\boldsymbol{\delta}_H; \boldsymbol{y}_{t:t+H}) := \tilde{\boldsymbol{y}}'_{t:t+H} A \tilde{\boldsymbol{y}}_{t:t+H}, \tag{3.20}$$

where $\tilde{\boldsymbol{y}}_{t:t+H}$ represents the counterfactual output path, $\tilde{\boldsymbol{y}}_{t:t+H} = \boldsymbol{y}_{t:t+H} - \Theta_{ye,H}\boldsymbol{\delta}_H$ and $\Theta_{ye,H}$ is defined in (3.18). The matrix $\boldsymbol{A}$ is a weighting matrix defined as $\boldsymbol{A} = \text{diag}[1, \alpha, \alpha^2, \ldots, \alpha^H]$, where α is the discount factor. To ensure the existence of a global maximum for the utility function, assume the Hessian matrix of the utility function is negative definite:

$$\mathbf{H}_U := \frac{\partial^2 U(\boldsymbol{\delta}_H; \boldsymbol{y}_{t:t+H})}{\partial \boldsymbol{\delta}_H \partial \boldsymbol{\delta}_H'} = \Theta_{ye,H}' \boldsymbol{A} \Theta_{ye,H} \prec 0 \tag{3.21}$$

Thus, the desired policy shocks is derived through:

$$\nabla U(\boldsymbol{\delta}_H^*; \boldsymbol{y}_{t:t+H}) = \mathbf{0} \tag{3.22}$$

where $\nabla U := \frac{\partial U(\boldsymbol{\delta}_H; \boldsymbol{y}_{t:t+H})}{\partial \boldsymbol{\delta}_H}$. This yields the desired policy shocks as:

$$\boldsymbol{\delta}_H^* = -(\Theta_{ye,H}' \boldsymbol{A} \Theta_{ye,H})^{-1} \Theta_{ye,H}' \boldsymbol{A} \boldsymbol{y}_{t:t+H} \tag{3.23}$$

Lastly, apply (3.19), we obtain the desired policy path.

Remarks:

(i) Negative definiteness is useful for ensuring the existence of a global maximum, which depends on the properties of the matrix $\Theta_{ye,H}$. If the condition of negative definiteness does not hold, it may still be possible to identify a local maximum by constraining the domain of $\boldsymbol{\delta}_H$.

(ii) The weighting matrix $\boldsymbol{A}$ can be generalized to include nonzero off-diagonal terms, enabling the utility function to account for interaction effects between outputs across different periods.

(iii) The quadratic function can be interpreted as a Constant Elasticity of Substitution (CES) utility function with an elasticity of substitution equal to 2. One advantage of the quadratic function is that its Hessian matrix is constant, which simplifies the process of verifying the existence of a global maximum and computing it.

(iv) In practice, policymakers are often concerned with multiple output variables. To account for this, it is feasible to consider a more general utility function, which can be expressed as a weighted summation of several utility functions.

4 Linear projection identification

In this section, we investigate a linear projection-based methodology for identifying these parameters of interest for the counterfactual exercises through external instruments.

4.1 Hypothetical trajectory parameters

We propose a projection-based method to identify the hypothetical trajectory parameter, $\boldsymbol{\beta}_{h,H} := \Theta_{re,H}^{-\prime} \boldsymbol{\theta}_{ye,h}$, as defined in Section 3.

To illustrate this, consider a structural equation that includes the hypothetical trajectory parameter $\boldsymbol{\beta}_{h,H}$,

$$Y_{t+h} = \boldsymbol{\beta}'_{h,H} R_{t:t+H} + u_{t,h} \tag{4.1}$$

where $u_{t,h} = Y_{t+h} - \boldsymbol{\beta}'_{h,H} R_{t:t+H}$. As our objective is to develop a projection-based method for identifying the hypothetical trajectory parameter, it is crucial to examine the properties of the residual $u_{t,h}$. Considering the moving average model specified in (2.3), the independent variable can be rewritten as:

$$R_{t:t+H} = R_{t:t+H} + \Theta_{re,H}\varepsilon_{t,H,(S)} - \Theta_{re,H}\varepsilon_{t,H,(S)} \tag{4.2}$$

where $\Theta_{re,H}$ is the impulse response matrix, defined in (3.6). With substitution of $\boldsymbol{\beta}_{h,H}$ by its definition $\boldsymbol{\beta}_{h,H} := \Theta^{-'}_{re,H}\boldsymbol{\theta}_{ye,h}$, we rewrite $u_{t,h}$ as

$$u_{t,h} = Y_{t+h} - \boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H}(R_{t:t+H} + \Theta_{re,H}\varepsilon_{t,H,(S)} - \Theta_{re,H}\varepsilon_{t,H,(S)}) \tag{4.3}$$

The expression above includes the term $\Theta^{-}_{re,H}\Theta_{re,H}$. By definition, the Moore-Penrose inverse satisfies $A^{-}A = I$ when the matrix A has full column rank. It is desirable to obtain $\Theta^{-}_{re,H}\Theta_{re,H}$ as an identity matrix, ensuring that $u_{t,h}$ is orthogonal to the selected policy shocks, $\varepsilon_{t,H,(S)}$. This reasoning is due to our structural interpretation, where the selected policy shock is considered the sole contributor to the deviation in the policy path and the resulting disparity between the baseline and counterfactual outputs.

Thus, we introduce the 'full column rank assumption':

Assumption 4.1 (Full column rank assumption). *The matrix $\Theta_{re,H}$ has full column rank.*

This assumption implies that the number of manipulable policy shocks is less than or equal to the number of variable deviations, and that these shocks are distinct from one another. In practice, this assumption is typically satisfied, as seen in methods such as the period-by-period shock approach and the initial multi-shock method.

Under Assumption 4.1, the residual $u_{t,h}$ can be rewritten as

$$u_{t,h} = Y_{t+h} - \boldsymbol{\theta}'_{ye,h}\varepsilon_{t,H,(S)} - \boldsymbol{\beta}'_{h,H}(R_{t:t+H} - \Theta_{re,H}\varepsilon_{t,H,(S)}) \tag{4.4}$$

Given that $\boldsymbol{\theta}_{ye,h}$ and $\Theta_{re,H}$, as respectively defined in (3.4) and (3.6), are the impulse response functions, the linearity imposed in (2.3) implies that $\mathrm{P}(Y_{t+h} \mid \varepsilon_{t,H,(S)}) = \boldsymbol{\theta}'_{ye,h}\varepsilon_{t,H,(S)}$ and $\mathrm{P}(R_{t:t+H} \mid \varepsilon_{t,H,(S)}) = \Theta_{re,H}\varepsilon_{t,H,(S)}$. Consequently, the residual $u_{t,h}$ is orthogonal to the selected policy shocks,

$$\mathrm{P}(u_{t,h} \mid \varepsilon_{t,H,(S)}) = 0. \tag{4.5}$$

The zero projection result provides a moment equation to identify the hypothetical trajectory parameter. However, since the selected policy shocks are latent in practice, we consider variable, $z_{t,H}$, as an instrumental variable to address this issue.

Assumption 4.2 (IV assumption). *Suppose $\boldsymbol{z}_{t,H}$ follows $\boldsymbol{z}_{t,H} = \Pi\varepsilon_{t,H,(S)} + \boldsymbol{\eta}_{t,H}$ and satisfies*

(i) (Relevance and exact identification) Π *is a $n_e \times n_e$ full rank matrix,*

(ii) (Exogeneity) $\mathbb{E}[\boldsymbol{\eta}_{t,H}\boldsymbol{\varepsilon}'_{t,H,(S)}] = \mathbf{0}$, *for all* $t \geq 1$.

The instrumental variable (IV) assumption is grounded in the LP-IV (Local Projection - Instrumental Variable) framework, as detailed by Stock and Watson (2018).[10] The instrument $z_{t,H}$ is assumed to be exclusively correlated with the selected policy shocks and uncorrelated with all other shocks.

It is important to note that the instrument $\boldsymbol{z}_{t,H}$ does not necessarily have to be external. The trivial case is encompassed where $\boldsymbol{z}_{t,H} = \boldsymbol{\varepsilon}_{t,H,(S)}$, with Π as an identity matrix and $\boldsymbol{\eta}_{t,H}$ equal to zero. As a result, the statistical methods presented in this paper can readily accommodate most of the SVAR literature, where structural shocks of interest are identified by imposing conditions that allow the structural shock to serve as a valid instrument for itself.

According to Assumption 4.2, we obtain a moment equation,

$$\mathbb{E}[\boldsymbol{z}_{t,H}(Y_{t+h} - \boldsymbol{\beta}'_{h,H}R_{t:t+H})] = \mathbf{0}. \tag{4.6}$$

Since the full column rank assumption implies that the dimension of the instruments $z_{t,H}$, which matches the dimension of the selected policy shocks by Assumption 4.2(i), can be less than the dimension of the deviated policy variables, it is considered to identify the hypothetical trajectory parameter through second moments using the Moore-Penrose inverse. Typically, this scenario could lead to under-identification, making the parameter of interest difficult to accurately identify. However, due to the careful construction of the parameter $\boldsymbol{\beta}_{h,H} := \Theta^{-'}_{re,H}\boldsymbol{\theta}_{ye,h}$, it can be well identified. The results are presented in the following proposition, with the proof provided in Appendix C.2.

Proposition 4.1 (Projection-based identification for the hypothetical trajectory parameter)**.** *Let W_t follows equation* (2.3) *and Assumption 4.1 and 4.2 are satisfied. Then*

$$\boldsymbol{\beta}_{h,H} = \mathbb{E}[\boldsymbol{z}_{t,H}R'_{t:t+H}]^{-}\mathbb{E}[\boldsymbol{z}_{t,H}Y_{t+h}]. \tag{4.7}$$

Remarks:

(i) In contrast to the causal interpretation of the impulse response function in response to a one-time shock, the hypothetical trajectory parameter, $\boldsymbol{\beta}_{h,H}$, captures the causal effect of a sequence of interventions on policy variables over multiple periods. The external instrument, correlated solely with the selected policy shocks, ensures that the intervention on policy variables occurs exclusively through these shocks. Capturing this sequence of interventions is challenging because each policy variable is endogenous to the previous policy shock. Conventional methods often account for the endogenous response of the policy variable at each period to all prior shocks, requiring the computation of new shocks with precise magnitudes to replicate the path

[10] The literature on external instruments in macroeconomics often focuses on high-frequency instrumental variables for monetary policy, such as those studied by Gertler and Karadi (2015) and Bauer and Swanson (2023). These works identify monetary policy shocks by capturing changes in financial market instruments in response to Federal Open Market Committee (FOMC) announcements. Typically, the instrumental variable is processed by partialling out lagged variables to satisfy the exogeneity assumption (Assumption 4.2 (ii)). This choice of instruments is based on the premise that financial market responses are plausibly uncorrelated with shocks other than the monetary policy shock within a narrow time window.

deviation. Our projection method, however, provides a more straightforward and interpretable identification approach, thereby enhancing the practicality of counterfactual analysis.

(ii) The projection-based identification method provides a straightforward approach for estimating hypothetical trajectory parameters, utilizing the well-established two-stage least squares (2SLS) technique for easy implementation. To determine the output variable gap between the baseline and hypothetical economy, practitioners simply compute the inner product of the parameter estimates and the vector of policy path deviation. This approach offers a practical and accessible approach of assessing the impact of counterfactual path deviations on the output variable across multiple periods. Moreover, if the matrix $\mathbb{E}[\boldsymbol{z}_{t,H} R'_{t:t+H}]$ is square, statistical inferences, as presented in Section 5, can be readily obtained. This allows for the straightforward construction of confidence intervals for the output variable disparity.

(iii) Another rationale for imposing the assumption of full column rank on the matrix $\Theta_{re,H}$ lies in its significance for the identification process when using an external instrument, $\boldsymbol{z}_{t,H}$. If $\Theta_{re,H}$ has full row rank but lacks full column rank, applying the Moore-Penrose inverse to the covariance matrix $\mathbb{E}[\boldsymbol{z}_{t,H} R'_{t:t+H}]$ is equivalent to using an optimal weighting matrix that accounts for the strength of each instrument. In such cases, researchers should forego the linear projection method and instead rely on the analytical expression of $\boldsymbol{\beta}_{h,H}$ with the impulse response functions, $\Theta_{re,H}$ and $\boldsymbol{\theta}_{ye,h}$.

4.2 Policy intervention parameter

In this subsection, we introduce a projection-based method to identify the policy intervention parameter $\phi_h = \theta_{yx,h} - \boldsymbol{\beta}'_{h,H} \boldsymbol{d}_H^{(po)}$, as derived in (3.16).

To illustrate this, consider a structural equation that includes the policy intervention parameter ϕ_h,

$$Y_{t+h} = \phi_h \varepsilon_{x,t} + \boldsymbol{\beta}'_{h,H} R_{t:t+H} + v_{t,h}, \tag{4.8}$$

where $v_{t,h} = u_{t,h} - \phi_h \varepsilon_{x,t}$ and $u_{t,h}$ is defined in (4.1). Based on the following system of linear equations, (4.8) offers a causal mediation interpretation for the parameter ϕ_h,

$$\begin{aligned} Y_{t+h} &= \theta_{yx,h} \varepsilon_{x,t} + (\text{residuals}), \\ R_{t:t+H} &= \boldsymbol{d}_H^{(po)} \varepsilon_{x,t} + (\text{residuals}), \\ Y_{t+h} &= \phi_h \varepsilon_{x,t} + \boldsymbol{\beta}'_{h,H} R_{t:t+H} + v_{t,h}. \end{aligned} \tag{4.9}$$

where $\theta_{yx,h}$, defined in (2.7), represents the impulse response and $\boldsymbol{d}_H^{(po)}$, specified in (2.17), denotes the vector of endogenous responses of the policy variable to the shock of interest. The term, (residuals), in each equation captures the linear combination of all remaining structural shocks corresponding to the respective dependent variable.

The three-equation system is typically seen in causal mediation literature, see, e.g., Imai et al. (2010). The first equation captures the total effect of the shock of interest on the output variable. The second equation reflects the causal effect of the policy variables, acting as mediators, on the output variable. The third equation presents the direct effect, ϕ_h, with the policy variables held constant. This interpretation aligns with the analytical solution for ϕ_h in (3.16), $\phi_h = \theta_{yx,h} - \boldsymbol{\beta}'_{h,H}\boldsymbol{d}_H^{(po)}$.

Since structural shocks are latent, the feasibility of (4.8) depends on the ability to recover the shock $\varepsilon_{x,t}$ from the observables.

Assumption 4.3 (Partial identification)**.** *Suppose the structural shock $\varepsilon_{x,t}$ is recoverable, that is, $\varepsilon_{x,t} = X_t - \mathrm{P}_L(X_t \mid \overline{W}_{t-1})$, where $\overline{W}_{t-1} = (W'_{t-1}, W'_{t-2}, \cdots, W'_{t-p})'$.*

The concept of partial invertibility is analogous to Definition 1 in Miranda-Agrippino and Ricco (2023) and represents a weaker version of the invertibility condition presented in (24) of Stock and Watson (2018). This condition implies that the shock $\varepsilon_{x,t}$ can be identified using observable variables. Specifically, it assumes that $\varepsilon_{x,t}$ can be expressed as a linear combination of X_t and $\overline{W}_{t-1}$. This includes the commonly applied recursive ordering assumption in SVAR models, where $\varepsilon_{x,t}$ is the structural shock ordered first, as outlined in equation (2) of Gonçalves, Herrera, Kilian, and Pesavento (2021).

Assumption 4.3 can be extended to a more general scenario where $\varepsilon_{x,t}$ belongs to the linear space spanned by contemporaneous and lagged variables, such that $\varepsilon_{x,t} = \mathrm{P}_L(\varepsilon_{x,t} \mid W_t, W_{t-1}, \ldots)$. This implies that a linear combination of X_t with some contemporaneous and past controls can recover the structural shocks.[11]

To identify the parameter ϕ_h, it is essential to examine the properties of $v_{t,h}$. Given that $v_{t,h} = u_{t,h} - \phi_h \varepsilon_{x,t}$, and using the expression for $u_{t,h}$ provided in (4.1), we obtain

$$\mathrm{P}(u_{t,h} \mid \varepsilon_{x,t}) = (\theta_{yx,h} - \boldsymbol{\beta}'_{h,H}\boldsymbol{d}_H^{(po)})\varepsilon_{x,t} = \phi_h \varepsilon_{x,t}. \tag{4.10}$$

In addition to the result that $u_{t,h}$ is orthogonal to the selected policy shocks, we also obtain

$$\mathrm{P}(v_{t,h} \mid \varepsilon_{x,t}, \boldsymbol{\varepsilon}_{t,H,(S)}) = 0. \tag{4.11}$$

The zero projection result provides a moment equation for identifying the policy intervention parameter, as well as the hypothetical trajectory parameter.[12]

Similar to the projection-based identification of the hypothetical trajectory parameter, we consider the instrumental variable $\boldsymbol{z}_{t,H}$ and define the combined instruments as

$$\boldsymbol{z}_{x,t,H} = (\varepsilon_{x,t}, \boldsymbol{z}'_{t,H})'. \tag{4.12}$$

where $\boldsymbol{z}_{x,t,H}$ is a $(n_e + 1) \times 1$ vector of variables. Consequently, the instrument $\boldsymbol{z}_{x,t,H}$ is correlated only with the shocks $(\varepsilon_{x,t}, \boldsymbol{\varepsilon}'_{t,H,(S)})'$ and remains uncorrelated with all other shocks. Incorporating (4.11) leads to the following moment condition:

$$\mathbb{E}\Big[\boldsymbol{z}_{x,t,H}(Y_{t+h} - \phi_h \varepsilon_{x,t} - \boldsymbol{\beta}'_{h,H}\boldsymbol{R}_{t:t+H})\Big] = \boldsymbol{0}. \tag{4.13}$$

[11] A practical example is the use of recursive ordering with certain contemporaneous variables ordered before X_t. For instance, Christiano et al. (2005) identified monetary policy shocks by imposing a temporal ordering where output, consumption, investment, wages, productivity, and the price deflator do not respond to the policy rate (Federal Funds Rate).

[12] If researchers are solely interested in conducting a hypothetical trajectory analysis, (4.1) is sufficient for identifying the hypothetical trajectory parameter. However, (4.8) is constructed specifically for the identification of the policy intervention parameter.

If the dimension of the instrument $z_{x,t,H}$ matches the number of explanatory variables, that is, $n_e = H+1$, the parameter of interest ϕ_h can be identified as follows:

$$\phi_h = v(1)'\mathbb{E}[z_{x,t,H}R'_{x,t,H}]^{-1}\mathbb{E}[z_{x,t,H}Y_{t+h}], \tag{4.14}$$

where $v(1)$ is the selection vector $v(1) = (1, 0, \ldots, 0)'$, and $R_{x,t,H}$ represents the $(H+2)\times 1$ vector of the stacked explanatory variables in (4.8),

$$R_{x,t,H} := (\varepsilon_{x,t}, R'_{t:t+H})'. \tag{4.15}$$

If the dimension of the instrument $z_{x,t,H}$ is less than the number of explanatory variables, that is, $n_e < H+1$, the parameter of interest ϕ_h can be identified up to a scale parameter μ_θ for all $h \geq 0$.

$$\phi_h = v(\mu_\theta)'\mathbb{E}[z_{x,t,H}R'_{x,t,H}]^{-}\mathbb{E}[z_{x,t,H}Y_{t+h}], \tag{4.16}$$

where $v(\mu_\theta) = (\mu_\theta, 0, \ldots, 0)'$ and

$$\mu_\theta = (1 + \boldsymbol{d}_H^{(po)'} M_\Theta \boldsymbol{d}_H^{(po)}), \tag{4.17}$$

with M_Θ defined as the projection matrix $M_\Theta = I - P_\Theta$, and $P_\Theta = \Theta_{re,H}(\Theta'_{re,H}\Theta_{re,H})^{-1}\Theta_{re,H}$. Note that (4.16) encompasses the special case where the matrix $\Theta_{re,H}$ is square, in which the scale parameter μ_θ equals 1. Mathematically, this scale parameter emerges from the use of the Moore-Penrose inverse in the moment equation. Economically, this occurs because the Moore-Penrose inverse prevents the initial shock of interest from receiving a unit impact when $\Theta_{re,H}$ is not a square matrix. A more detailed explanation is provided in Remarks (ii) below.

Thus, we summarize the identification results in the following proposition:

Proposition 4.2 (Projection-based identification for policy intervention effect). *Let W_t follows equation* (2.3) *and Assumption 4.1, 4.2, and 4.3 are satisfied. Then,*

$$\phi_h = v(\mu_\theta)'\mathbb{E}[z_{x,t,H}R'_{x,t,H}]^{-}\mathbb{E}[z_{x,t,H}Y_{t+h}]. \tag{4.18}$$

where $R_{x,t,H}$ and $v(\mu_\theta)$ are defined in (4.15) *and* (4.16)*, respectively.*

The proof of Proposition 4.2 is given in Appendix C.3.

Remarks:

(i) The policy intervention parameter ϕ_h represents the response of Y_{t+h} to a one-unit change in the independent variable $\varepsilon_{x,t}$ while holding $R_{t:t+H}$ constant. Unlike conventional Local Projection models, which use pre-treatment variables as controls, our projection model includes contemporaneous and future policy variables to suppress the endogenous responses of these variables. This specific control variable setup allows for counterfactual analysis, creating scenarios where certain policy variables are restricted from responding to the shock of interest.

(ii) The identification of ϕ_h is achieved through the instrument $z_{t,H}$, which is correlated exclusively with $\varepsilon_{t,H,(S)}$. This ensures that endogenous responses are suppressed solely through the selected policy shocks $\varepsilon_{t,H,(S)}$. The parameter ϕ_h, identified using the external instrument, has a causal interpretation of the direct effect, capturing the dynamic causal impact of the shock $\varepsilon_{x,t}$ on the outcome variable, with no indirect effects through a sequence of policy variables.

(iii) The term μ_θ^{-1} arises from the Moore-Penrose inverse of the matrix $\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]$. Economically, this implies that if $\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]$ has full column rank but not full row rank, indicating an insufficient number of policy shocks to perfectly mimic the counterfactual scenario, the Moore-Penrose inverse will generate the combination of the selected policy shocks and the initial shock of interest $\varepsilon_{x,t}$ with certain magnitude that mimics the counterfactual scenario with the minimum Euclidean norm of the deviation. Consequently, the initial shock $\varepsilon_{x,t}$ may not have a unit size impact, which is inconsistent with the economic definition of ϕ_h. However, when $\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]$ is a square matrix, $\mu_\theta = 1$. It implies that the policy maker can perfectly mimic the counterfactuals with the policy shocks and thereby the parameter ϕ_h can exactly be identified.

An alternative identification method using an external instrument for the shock of interest $\varepsilon_{x,t}$ is presented in Appendix A. Instead of assuming the shock of interest is recoverable from the observables, a valid external instrument is employed. Additionally, the linear projection identification method requires an assumption that the policy shock has no contemporaneous causal effect on the variable of interest, X_t. This is because it is undesirable for the suppression of the endogenous response of the policy variable to have a contemporaneous impact on the variable of interest. Such an impact could prevent the variable of interest from receiving a unit shock, leading to the misidentification of the parameter.

5 Period-by-period shocks: estimation and inference

In this section, we discuss the estimation and statistical inference of projection-based counterfactual parameter estimates for period-by-period shock counterfactuals.

5.1 Stronger assumption

The period-by-period shock method assumes that the policy variable is subjected to only one shock per period, enabling the policymaker to manipulate the shock each period to accurately replicate the counterfactual scenario. This approach imposes specific conditions on the selected policy shocks and the instrumental variables.

Assumption 4.2*. *Suppose*

(i) the selected policy shocks $\varepsilon_{t,H,(S)} = (\varepsilon_{r,t}, \varepsilon_{r,t+1}, \cdots, \varepsilon_{r,t+H})'$ *and* $\varepsilon_{r,t}$ *is a scalar variable.*

(ii) the instrumental variable $\boldsymbol{z}_{t,H} = (z_t, z_{t+1}, \cdots, z_{t+H})'$, *such that* $z_t = \pi\varepsilon_{r,t} + \eta_t$, *and* $\pi \neq 0$ *and* $\mathbb{E}[\varepsilon_s \eta_t] = 0$ *for all* $s, t \geq 1$.

Compared to Assumption 4.2, this assumption is stronger due to the explicit specification of the selected shocks and instrumental variables. The period-by-period shock approach mirrors the conventional method, which involves recursively computing the impulse responses of variables at each horizon, then suppressing the endogenous responses of the policy variable (e.g., the Federal Funds Rate), and finally accounting for the impact of nullifying these endogenous policy responses. In contrast, this section provides a one-step estimation and inference method for this counterfactual exercise.

Under Assumption 4.2*, it becomes feasible to derive a more explicit form of the counterfactual parameters. The hypothetical trajectory parameter, as defined in Section 3, is given by $\boldsymbol{\beta}_{h,H} := \Theta_{re,H}^{-\prime}\boldsymbol{\theta}_{ye,h}$. With Assumption 4.2*, the matrix $\Theta_{re,H}$ is a lower triangular matrix, where the (ij)-th element is impulse response of the policy variable to its own structural shock, $\partial R_{t+h}/\partial \varepsilon_{r,t}$, and zero otherwise; and $\boldsymbol{\theta}_{ye,h} = (\theta_{yr,h}, \ldots, \theta_{yr,1}, \theta_{yr,0}, \mathbf{0}_{1\times(H-h)})'$. Consequently, the last $H-h$ elements in $\boldsymbol{\beta}_{h,H}$ are zero.

$$\boldsymbol{\beta}_{h,H} = [\boldsymbol{\beta}_h', \mathbf{0}_{1\times(H-h)}]', \tag{5.1}$$

by defining $\boldsymbol{\beta}_h$ as the sub-vector of $\boldsymbol{\beta}_{h,H}$ containing its first $(h+1)$ elements.

Since the last $H-h$ elements in $\boldsymbol{\beta}_{h,H}$ are zero in this case, the policy path deviation beyond period h does not influence the output disparity. Therefore, the hypothetical trajectory outcome ψ_h can be expressed as:

$$\psi_h = \boldsymbol{\beta}_h' \boldsymbol{d}_h^{(ht)}, \tag{5.2}$$

where $\boldsymbol{d}_h^{(ht)}$ is the sub-vector of $\boldsymbol{d}_H^{(ht)}$ containing its first $h+1$ elements,

$$\boldsymbol{d}_h^{(ht)} = S_h \boldsymbol{d}_H^{(ht)}, \tag{5.3}$$

with $S_h := [I_{h+1}, \mathbf{0}_{(h+1)\times(H-h)}]$.

The economic interpretation of period-by-period shock counterfactuals is that the output difference at horizon h depends solely on the policy path deviation up to that point, which is achieved through the manipulation of policy shocks. Specifically, in each period, the private sector perceives the endogenous responses (impulse responses) of all previous policy interventions and interprets the difference arising from the policy deviation $(r_{t+h} - \tilde{r}_{t+h})$ as the new policy intervention for that period. This implies that the private sector lacks prior information about future policy path deviation, relying instead on the observed equilibrium path of the policy shock response and the current intervention.

Under Assumption 4.2*, the policy intervention parameter in (5.1), $\phi_h = \theta_{yx,h} - \boldsymbol{\beta}_{h,H}' \boldsymbol{d}_H^{(po)}$, is influenced only by the path deviation up to the first $h+1$ periods. Therefore, the parameter ϕ_h can be rewritten as:

$$\phi_h = \theta_{yx,h} - \boldsymbol{\beta}_h' \boldsymbol{d}_h^{(po)}, \tag{5.4}$$

where

$$\boldsymbol{d}_h^{(po)} = S_h \boldsymbol{d}_H^{(po)}. \tag{5.5}$$

The economic interpretation of this specific case suggests that if the suppression of the endogenous response of the policy variable is achieved through period-by-period shocks, then any suppression beyond period h has no impact on the output variable at horizon h.

5.2 Infeasible estimation and statistical inference

Following the linear projection identification method, we consider an "infeasible" version of the hypothetical trajectory parameter and policy intervention effect. It is important to note that identifying the counterfactual parameters does not necessarily require the inclusion of past variables as control variables, as long as the instrument satisfies the lead-lag exogeneity condition. However, adding past variables as control variables can effectively reduce the variation in the projection residual, thereby increasing estimation precision. We denote the subscript $\perp$ to indicate the projection residual for any variable after partialling out the control variable $\overline{W}_{t-1}$, i.e., $Y^{\perp}_{t+h} = Y_{t+h} - \mathrm{P}_L(Y_{t+h} \mid \overline{W}_{t-1})$.

Consider two infeasible regression models:

$$\begin{aligned} Y^{\perp}_{t+h} &= \beta_h' R^{\perp}_{t:t+h} + u^{\perp}_{t,h}, \\ Y^{\perp}_{t+h} &= \phi_h \varepsilon_{x,t} + \beta_h' R^{\perp}_{t:t+h} + v^{\perp}_{t,h}. \end{aligned} \tag{5.6}$$

Here, we adopt Assumption 4.3 (ii) and consider $\varepsilon_{x,t} = X_t - \mathrm{P}_L(X_t \mid \overline{W}_{t-1})$. Consequently, it follows that $\varepsilon_{x,t} = \varepsilon^{\perp}_{x,t}$. This leads to two infeasible least squares (LS) estimates:

$$\tilde{\beta}_h = \left(\sum_{t=1}^{T-h} z_{t,h} R^{\perp'}_{t:t+h} \right)^{-1} \left(\sum_{t=1}^{T-h} z_{t,h} Y^{\perp}_{t+h} \right), \tag{5.7}$$

$$\tilde{\phi}_h = \nu(1)' \left(\sum_{t=1}^{T-h} z_{x,t,h} R^{\perp'}_{x,t,h} \right)^{-1} \left(\sum_{t=1}^{T-h} z_{x,t,h} Y^{\perp}_{t+h} \right), \tag{5.8}$$

where $z_{t,h} = S_h' z_{t,H}$ and $z_{x,t,h} = S_h' z_{x,t,H}$. The term "infeasible" refers to these estimates being derived from unobservable inputs, such as the latent shock variable $\varepsilon_{x,t}$ in $z_{x,t,H}$, and the projected variables $R^{\perp}_{t:t+h}$ and $Y^{\perp}_{t+h}$. In practice, these terms are typically replaced by their sample counterparts. In the subsequent section, we propose "feasible" estimates and demonstrate that the two types of estimates are asymptotically equivalent.

To obtain the statistical inference, it is helpful to define the long run variance (LRV) of two regression score functions:

$$\Omega_{i,h} := \sum_{k=-\infty}^{\infty} \mathbb{E}[s_{i,t,h} s'_{i,t+k,h}], \tag{5.9}$$

for $i \in \{r, x\}$, where

$$\begin{aligned} s_{r,t,h} &:= z_{t,h} u^{\perp}_{t,h}, \\ s_{x,t,h} &:= z_{x,t,h} v^{\perp}_{t,h}. \end{aligned} \tag{5.10}$$

We impose the following regularity conditions to establish the asymptotic theory for the infeasible estimates.

Assumption 5.1 (Regularity conditions). *For all $t \geq 1$ and $h \geq 0$, let c_0, c_1 are positive constant, suppose*

(i) $\eta_t, W^{\perp}_{t+h}$ are covariance stationary.

(ii) $\varepsilon_t, \eta_t, W^{\perp}_{t+h}$ are strong mixing (α-mixing) processes with mixing size $-r/(r-2)$, for $r > 2$ and all finite integer $h \geq 0$.

(iii) $\mathbb{E}\|\varepsilon_t\|^{4r+\delta}$, $\mathbb{E}\|\eta_t\|^{4r+\delta}$, *and* $\mathbb{E}\|W_{t+h}^{\perp}\|^{4r+\delta} < c_0 < \infty$, *for* r *defined in (ii), and* $\delta > 0$.

(iv) $\lambda_{\min}\left(\Omega_{i,h}\right) > c_1 > 0$, *for* $i \in \{r, x\}$.

Assumption 5.1 imposes standard regularity conditions on the structural shocks and the partialled out random variable. Specifically, Assumption 5.1(i), which assumes that $W_{t+h}^{\perp}$ is stationary, permits the original variable W_{t+h} to be non-stationary. The strong mixing condition, along with the boundedness of moments, supports the asymptotic theory. Assumption 5.1(iv) ensures that the matrix $\Omega_{i,h}$ is non-singular. With these assumptions, we can proceed with the statistical inference for the infeasible estimates.

Proposition 5.1 (Asymptotic normality of infeasible estimates)**.** *Let* W_t *follows equation* (2.3) *and Assumption 4.1, 4.2*, and 5.1 are satisfied. Then*

$$\sqrt{T}\sigma_{\psi,h}^{-1}\left(\tilde{\psi}_h - \psi_h\right) \xrightarrow{d} N(0,1), \tag{5.11}$$

where $\sigma_{\psi,h}^2 = (\boldsymbol{d}_h^{(ht)\prime}\Sigma_{zr,h}^{-1}\Omega_{r,h}\Sigma_{zr,h}^{\prime -1}\boldsymbol{d}_h^{(ht)})^{1/2}$ *and* $\Sigma_{zr,h} := \mathbb{E}[\boldsymbol{z}_{t,h}R_{t:t+h}^{\perp\prime}]$. *Moreover, if Assumption 4.3 holds,*

$$\sqrt{T}\sigma_{\phi,h}^{-1}\left(\tilde{\phi}_h - \phi_h\right) \xrightarrow{d} N(0,1), \tag{5.12}$$

where $\sigma_{\phi,h} = (\nu(1)'\Sigma_{zx,h}^{-1}\Omega_{x,h}\Sigma_{zx,h}^{\prime -1}\nu(1))^{1/2}$ *and* $\Sigma_{zx,h} := \mathbb{E}[\boldsymbol{z}_{x,t,h}R_{x,t,h}^{\perp\prime}]$.

The proof of Proposition 5.1 is given in Appendix C.4. In the next subsection, we propose feasible estimates and demonstrate that the estimation bias between feasible and infeasible estimates is asymptotically negligible.

5.3 Feasible estimation and HAC inference

The infeasible estimates are practically unavailable because $\varepsilon_{x,t}$ and $W_{t+h}^{\perp}$ are latent variables. Additionally, obtaining the projection residual $W_{t+h}^{\perp}$ requires the control variable $\overline{W}_{t-1}$ to include all past variables, which is not feasible in practice. Consequently, it is standard to consider sample controls $\overline{W}_{t-1}$, which include p-lagged variables, with the order p allowed to grow with the sample size, albeit at a slower rate. This approach is a common procedure for handling infinite order time series models Lewis and Reinsel (1985) and is often employed in the first step of least squares-based VARMA estimation, as seen in Dufour and Pelletier (2022) and Wilms, Basu, Bien, and Matteson (2021). Consequently, we derive the sample counterparts for $\varepsilon_{x,t}$ and $W_{t+h}^{\perp}$.

$$\hat{\varepsilon}_{x,t} = X_t - \hat{\gamma}'\overline{W}_{t-1}, \qquad \hat{W}_{t+h}^{\perp} = W_{t+h} - \hat{\Gamma}_h\overline{W}_{t-1}, \tag{5.13}$$

where $\hat{\gamma}$ and $\hat{\Gamma}_h$ are two Least Square estimates, $\hat{\gamma} = (\sum_{t=1}^{T}\overline{W}_{t-1}\overline{W}_{t-1}')^{-1}\left(\sum_{t=1}^{T}\overline{W}_{t-1}X_t\right)$ and $\hat{\Gamma}_h = (\sum_{t=1}^{T-h}\overline{W}_{t-1}\overline{W}_{t-1}')^{-1}(\sum_{t=1}^{T-h}\overline{W}_{t-1}W_{t+h}')$.

Then, we obtain the feasible estimates for two counterfactual parameters.

$$\hat{\beta}_h = \left(\sum_{t=1}^{T-h} z_{t,h}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,h}\hat{Y}^{\perp}_{t+h}\right), \tag{5.14}$$

$$\hat{\phi}_h = \nu(1)'\left(\sum_{t=1}^{T-h} \hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h} \hat{z}_{x,t,h}\hat{Y}^{\perp}_{t+h}\right). \tag{5.15}$$

where $\hat{z}_{x,t,h} = (\hat{\varepsilon}_{x,t}, z'_{t,h})'$, $\hat{R}^{\perp'}_{x,t,h} = (\hat{\varepsilon}_{x,t}, \hat{R}^{\perp'}_{t:t+h})'$, $\hat{\varepsilon}_{x,t} = \hat{X}^{\perp}_t$, $\hat{R}^{\perp}_{t:t+h} = (\hat{R}^{\perp}_t, \hat{R}^{\perp}_{t+1}, \cdots, \hat{R}^{\perp}_{t+h})'$, and $\hat{X}^{\perp}_t, \hat{Y}^{\perp}_t, \hat{R}^{\perp}_t$ are elements of $\hat{W}^{\perp}_t = (\hat{X}^{\perp}_t, \hat{Y}^{\perp}_t \hat{R}^{\perp}_t)$.

To derive the statistical inference for the feasible estimates, we impose the following regularity conditions.

Assumption 5.2 (Asymptotic equivalence)**.** *For all $t \geq 1$, and finite integer $h \geq 0$, let*

(i) $\lim_{T\to\infty}\sum_{t=1}^{T-h}\left\|(\hat{W}^{\perp}_{t+h} - W^{\perp}_{t+h})(\eta_t, \varepsilon_{x,t}, u^{\perp}_{t,h})'\right\|^2 = O_p(1)$.

(ii) $\lim_{T\to\infty}\sum_{t=1}^{T-h}\left\|\hat{W}^{\perp}_{t+h} - W^{\perp}_{t+h}\right\|^4 = O_p(1)$.

Assumption 5.3 (Consistency of HAC estimates)**.** $\lim_{T\to\infty}\hat{\Omega}^{(hac)}_{i,h} \xrightarrow{p} \Omega_{i,h}$*, for $i \in \{r, x\}$.*

Assumption 5.2 is a high-level assumption on the estimation precision. It is commonly required in time series literature as some intermediate result for deriving the covariance matrix estimation. Based on the explicit form for $\hat{W}^{\perp}_{t+h} - W^{\perp}_{t+h} = (\Gamma_h - \hat{\Gamma}_h)\overline{W}_{t-1}$, the term in Assumption 5.2 (i) can be rewritten as $\sum_{t=1}^{T-h}||(\hat{\Gamma}_h - \Gamma_h)\overline{W}_{t-1}(\eta_t, \varepsilon_{x,t}, u^{\perp}_{t,h})||^2$. If the process W_t is covariance stationary, it can be readily obtained that $||\hat{\Gamma}_h - \Gamma_h||^2$ is an $O_p(T^{-1})$ term since the LS estimates are usually $\sqrt{T}$-asymptotic Gaussian distribution with certain regularity conditions. The term $\sum_{t=1}^{T-h}||\overline{W}_{t-1}(\eta_t, \varepsilon_{x,t}, u^{\perp}_{t,h})||^2$ commonly can be shown to be an $O_p(T)$ term. We provide an heuristic verification here. Without loss of generosity, focuses on the term $\sum_{t=1}^{T-h}||W_{t-1}\eta_t||^2$,

$$\sum_{t=1}^{T-h}||W_{t-1}\eta_t||^2 = \sum_{t=1}^{T-h}||\sum_{i=0}^{\infty}\Theta_i\varepsilon_{t-i}\eta_t||^2 \leq \sum_{i=0}^{\infty}||\Theta_i||^2\sum_{t=1}^{T-h}||\varepsilon_{t-i}\eta_t||^2 = O_p(T). \tag{5.16}$$

Since W_t is covariance stationary, $\sum_{i=0}^{\infty}||\Theta_i||^2$ represents the scale of the variance and is thus bounded by a constant. The term $\sum_{t=1}^{T-h}||\varepsilon_{t-i}\eta_t||^2$ is of order $O_p(T)$, as both ε_t and η_t have bounded fourth moments. Therefore, by combining this with the boundedness of $||\hat{\Gamma}_h - \Gamma_h||^2 = O_p(T^{-1})$, Assumption 5.2(i) is satisfied.

In the case where W_t contains a unit root, the boundedness still holds; however, the proof requires the use of a Dickey-Fuller type matrix to eliminate the perfect multicollinearity in W_t or $\overline{W}_{t-1}$ for cointegration (see Sims et al. (1990), Assumption 3 in Montiel Olea and Plagborg-Møller (2021), and Assumption 5.1 in Dufour and Wang (2024a)). Intuitively, although $\sum_{i=0}^{T}||\Theta_i||^2$ grows with the sample size T, some rotated estimates in $\hat{\Gamma}_h$ will exhibit super-consistency due to the presence of unit roots, ensuring that the product remains an

$O_p(1)$ term. A rigorous proof of this asymptotic equivalence can be constructed similarly to the proof of Proposition 5.2 (Estimates Equivalence) in Dufour and Wang (2024a). This similar approach can be applied to verify Assumption 5.2(ii).

Assumption 5.3 establishes a consistency requirement for heteroskedasticity and autocorrelation consistent (HAC) estimators of the long-run variance of the regression score function.

Thus, we establish the asymptotic normality of the feasible estimates, scaled by the HAC standard error.

Proposition 5.2 (Asymptotic normality of feasible estimates)**.** *Let the conditions of Proposition 5.1 hold. If Assumption 5.2 and 5.3 are satisfied, then*

$$\sqrt{T}\left(\hat{\sigma}^{(hac)}_{\psi,h}\right)^{-1}\left(\hat{\psi}_h-\psi_h\right)\xrightarrow{d}N(0,1), \tag{5.17}$$

$$\sqrt{T}\left(\hat{\sigma}^{(hac)}_{\phi,h}\right)^{-1}\left(\hat{\phi}_h-\phi_h\right)\xrightarrow{d}N(0,1), \tag{5.18}$$

where $\hat{\sigma}^{(hac)}_{\psi,h}=(\boldsymbol{d}^{(ht)\prime}_h\hat{\Sigma}^{-1}_{zr,h}\hat{\Omega}^{(hac)}_{r,h}\hat{\Sigma}^{-1\prime}_{zr,h}\boldsymbol{d}^{(ht)}_h)^{1/2}$ *and* $\hat{\sigma}^{(hac)}_{\phi,h}=\left(\nu(1)'\hat{\Sigma}^{-1}_{zx,h}\hat{\Omega}^{(hac)}_{x,h}\hat{\Sigma}^{-1\prime}_{zx,h}\nu(1)\right)^{1/2}$*, such that* $\hat{\Sigma}_{zx,h}=T^{-1}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp\prime}_{x,t,h}$ *and* $\hat{\Sigma}_{zr,h}=T^{-1}\sum_{t=1}^{T-h}z_{t,H}\hat{R}^{\perp\prime}_{t:t+h}$.

The proof of Proposition 5.2 is presented in Appendix C.5. This proposition provides the theoretical foundation for deriving confidence intervals for the estimates of the counterfactual outcomes. However, HAC inference tends to perform relatively poorly in terms of empirical test levels, particularly in small samples. In line with the literature on Local Projection, it is generally preferable to derive robust inference methods that avoid the need to correct for serial correlation in the projection residuals (e.g., see Montiel Olea and Plagborg-Møller (2021), Dufour and Wang (2024a), and Dettaa and Wang (2024)). The results for robust inference are provided in the following section.

5.4 Heteroskedastic robust inference in recoverable SVMA

In this section, we propose an alternative method for estimating the long-run variance (LRV) $\Omega_{x,h}$, based on a slightly stronger regularity condition.

Assumption 5.4 (Recoverable SVMA)**.** $\varepsilon_t=\mathrm{P}(\varepsilon_t\mid W_t,W_{t-1},\cdots)$*, for all* $t\geq 1$.

Assumption 5.5. *For all* $t\geq 1$*, suppose*

(i) $\mathbb{E}[\varepsilon_{i,t}\mid\varepsilon_{-i,t},\{\varepsilon_s\}_{s\neq t},\{\eta_s\}_{s\leq t}]=0$ *almost surely, for* $i\in\{x,r\}$.

(ii) $\mathbb{E}[\eta_t\mid\{\varepsilon_s\}_{s\neq t},\{\eta_s\}_{s<t}]=0$*, almost surely.*

Assumption 5.4 establishes a condition for recoverable SVMA, allowing the SVMA to be transformed into an SVAR model in which the structural shock ε_t spans the same space as the reduced-form VAR error. This assumption rules out the possibility that the projection errors $u^{\perp}_{t,h}$ and $v^{\perp}_{t,h}$ are correlated with any past structural shocks. This condition is crucial for proposing an alternative estimation method for long-run variance.

Assumption 5.5 (i) and (ii) imposes a more stringent restriction on the structural shocks than the typical assumptions of serial uncorrelation or martingale difference sequence. This condition can be satisfied in several scenarios, such as: (i) when ε_t is an independent and identically distributed random variable, (ii) when ε_t is a mean-independent process, or (iii) when ε_t belongs to certain types of conditionally heteroskedastic processes, such as an ARCH(1) process with an i.i.d. normally distributed shock variable.

Our objective is to propose an alternative heteroskedastic-robust/consistent (HR/HC) estimation method for the long-run variance $\Omega_{i,h}$, for $i \in \{r, x\}$. Specifically, this alternative method relies solely on the sample variance of the regression score function. Given that the sample variance is naturally positive semi-definite, our method can help researchers avoid HAC-type estimates, which depend on the choice of kernel functions and bandwidth. Since the regression score function is serially correlated, our HC/HR method requires a reordered regression score function.

$$
\begin{aligned}
s^*_{r,t,h} &:= z_t(u^\perp_{t,h}, u^\perp_{t-1,h}, \cdots, u^\perp_{t-h,h})', \\
s^*_{x,t,h} &:= \left[\varepsilon_{x,t} v^\perp_{t,h}, z_t(v^\perp_{t,h}, v^\perp_{t-1,h}, \cdots, v^\perp_{t-h,h})'\right]'.
\end{aligned}
\tag{5.19}
$$

Their sample counterpart is denoted as $\hat{s}^*_{r,t+h}, \hat{s}^*_{x,t+h}$ by replacing $\varepsilon_{x,t}, u^\perp_{t,h}, v^\perp_{t,h}$ with their sample estimates, such that

$$
\begin{aligned}
\hat{s}^*_{r,t+h} &= z_t(\hat{u}^\perp_{t,h}, \hat{u}^\perp_{t-1,h}, \cdots, \hat{u}^\perp_{t,h})' \\
\hat{s}^*_{x,t+h} &= [\hat{\varepsilon}_{x,t} \hat{v}^\perp_{t,h}, z_t(\hat{v}^\perp_{t,h}, \hat{v}^\perp_{t-1,h}, \cdots, \hat{v}^\perp_{t,h})']',
\end{aligned}
\tag{5.20}
$$

where $\hat{u}^\perp_{t,h} = \hat{Y}^\perp_{t+h} - \hat{\vartheta}'_{yr,h} \hat{R}^\perp_{t:t+H}$, $v^\perp_{t,h} = \hat{Y}^\perp_{t+h} - \hat{\vartheta}'_{yr,h} \hat{R}^\perp_{t:t+H} - \hat{\phi}_h \hat{\varepsilon}_{x,t}$, and $\hat{\varepsilon}_{x,t} = X_t - \hat{\gamma}' \overline{W}_{t-1}$.

Proposition 5.3. *Let Assumption 4.2*, 5.4, and 5.5 hold. Then*

$$
\operatorname{Var}(s^*_{i,t,h}) = \Omega_{i,h}, \ \text{for } i \in \{r, x\}. \tag{5.21}
$$

Furthermore, if Assumption 5.1 and 5.2 hold, then

$$
\widehat{\operatorname{Var}}(\hat{s}^*_{i,t,h}) \xrightarrow{p} \Omega_{i,h}, \ \text{for } i \in \{r, x\}. \tag{5.22}
$$

where $\widehat{\operatorname{Var}}(\hat{s}^*_{i,t,h}) = T^{-1} \sum_{t=1}^{T-h} \hat{s}^*_{i,t,h} \hat{s}^{*\prime}_{i,t,h}$.

The proof of Proposition 5.3 is presented in Appendix C.6. It offers an alternative approach for computing the long-run variance matrices $\Omega_{i,h}$, where $i \in r, x$. This method has the advantage of eliminating the need to correct for serial correlation in the projection residuals, as it relies on the sample covariance of the reordered score functions. Consequently, the estimates are inherently positive semidefinite. Moreover, this approach bypasses the need for selecting kernel functions and bandwidths. Monte Carlo simulations demonstrate that empirical tests using this method outperform those based on conventional HAC covariance matrix estimates, such as those reported by Dufour and Wang (2024a).

6 Delta-method statistical inference

In this subsection, we provide delta-method inference for these counterfactual parameters with explicit formulas for the covariance matrix. As we have derived the analytical solution in an SVMA framework, the necessary input would be the impulse response estimates. These could be recursive-VAR-based or LP-based estimates.

Assumption 6.1 (Asymptotic Normality of Impulse Response estimates)**.** *Suppose*

$$\sqrt{T}\begin{pmatrix} \hat{\theta}_{yx,h}-\theta_{yx,h} \\ \hat{\boldsymbol{d}}_H^{(po)}-\boldsymbol{d}_H^{(po)} \\ \mathrm{vec}(\hat{\Theta}_{re,H}-\Theta_{re,H}) \\ \hat{\boldsymbol{\theta}}_{ye,h}-\boldsymbol{\theta}_{ye,h} \end{pmatrix} \xrightarrow{d} N(\mathbf{0},\Omega_{h,H}). \tag{6.1}$$

Denote the the bottom-right block of $\Omega_{h,H}$ as $\Omega_{e,h,H}$, which represents the asymptotic variance of $(\mathrm{vec}(\hat{\Theta}_{re,H})', \hat{\boldsymbol{\theta}}'_{ye,h})'$. Assumption 6.1 assumes the convergence for the impulse response estimates. It is important to note that the covariance matrix $\Omega_{h,H}$ may not be full rank, as the joint distribution of impulse response estimates from the recursive-VAR method might yield a singular Jacobian matrix, or the dimension of the impulse responses may exceed the number of VAR slope coefficients, see, e.g., Inoue and Kilian (2016).

In Section 3, we derived the analytical solutions for the parameters of interest in the two counterfactuals. Let the estimates computed from these explicit formulas be denoted as

$$\begin{aligned} \check{\psi}_h &= \check{\beta}'_{h,H}\boldsymbol{d}_H^{(ht)}, \\ \check{\phi}_h &= \hat{\theta}_{yx,h}-\check{\beta}'_{h,H}\hat{\boldsymbol{d}}_H^{(po)}, \end{aligned} \tag{6.2}$$

where $\check{\beta}_{h,H}=\hat{\Theta}^{-'}_{re,H}\hat{\boldsymbol{\theta}}_{ye,h}$. Using the Delta method for linearization and Assumption 6.1, we can derive the asymptotic distribution for these two estimates. The results are presented in the following Proposition. Note that let $K_{n,m}$ be the commutation matrix, such that for any $m\times n$ matrix A, $\mathrm{vec}(A')=K_{n,m}\mathrm{vec}(A)$.

Proposition 6.1. *Suppose Assumption 6.1 holds, then*

$$\sqrt{T}(\check{\psi}_h-\psi_h)\xrightarrow{d} N(0,\mathrm{AVar}(\check{\psi}_h)), \tag{6.3}$$

$$\sqrt{T}(\check{\phi}_h-\phi_h)\xrightarrow{d} N(0,\mathrm{AVar}(\check{\phi}_h)), \tag{6.4}$$

where

$$\begin{aligned} \mathrm{AVar}(\check{\psi}_h) &= \boldsymbol{d}_H^{(ht)'}G\Omega_{e,h,H}G'\boldsymbol{d}_H^{(ht)}, \\ \mathrm{AVar}(\check{\phi}_h) &= J\Omega_{h,H}J', \end{aligned}$$

such that G,J are Jacobian matrices, $J=\left[1,-\beta'_{h,H},-\boldsymbol{d}_H^{(po)'}G\right]$, $G=[G_1,\Theta^{-'}_{re,H}]$, and the form of G_1 depends on the form of the Moore–Penrose inverse.

(i) If $\Theta^{-}_{re,H}=(\Theta'_{re,H}\Theta_{re,H})^{-1}\Theta'_{re,H}$, then

$$G_1=(\boldsymbol{\theta}'_{ye,h}(\Theta'_{re,H}\Theta_{re,H})^{-1})\otimes M_{\Theta}-(\boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H}\otimes\Theta^{-'}_{re,H})K_{(H+1),n_e};$$

(ii) If $\Theta^{-}_{re,H}=\Theta^{-1}_{re,H}$, then

$$G_1=-(\boldsymbol{\theta}'_{ye,h}\Theta^{-1}_{re,H}\otimes\Theta^{-1'}_{re,H})K_{(H+1),(H+1)};$$

(iii) If $\Theta^{-}_{re,H} = \Theta'_{re,H}(\Theta_{re,H}\Theta'_{re,H})^{-1}$*, then*

$$G_1 = (\boldsymbol{\theta}'_{ye,h} M_{\Theta'}) \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} - ((\boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H}) \otimes \Theta^{-'}_{re,H}) K_{(H+1),n_e};$$

where M_{Θ} *and* $M_{\Theta'}$ *are the residual maker matrices for* $\Theta_{re,H}$ *and* $\Theta'_{re,H}$*, respectively.*

The proof of Proposition 6.1 is given in Appendix C.1. Note that both estimators, ψ_h and ϕ_h, involve nonlinear transformations of impulse response functions. Consequently, compared to the projection-based estimates that yield nonsingular standard errors, as established in Proposition 5.1, the Delta-method inference may asymptotically produce standard errors that converge to zero.

7 Nonlinear model and simulation-based counterfactuals

In this section, we introduce a Monte Carlo simulation-based procedure for conducting macroeconomic counterfactual analysis within a general parametric framework.

Consider the output process described in equation (2.1) without assuming linearity. The model $\boldsymbol{g}_t(\cdot)$ is designed to accommodate time-varying dynamics, where the output variable is influenced by a series of current and past random shocks. For instance, $\boldsymbol{g}_t(\cdot)$ could represent a state-dependent model, a smooth transition VAR, or a time-varying VAR, as discussed in Kilian and Lütkepohl (2017), Gonçalves, Herrera, Kilian, and Pesavento (2024), and the references therein.

In contrast to the robust approach discussed in the previous section, relaxing the linearity assumption to account for nonlinearity offers a more flexible framework for capturing the potentially varied responses of the private sector to changes in economic conditions. However, a key drawback of this relaxation is that it requires a full specification of the structural model, as highlighted in Barnichon and Mesters (2023).

Historical scenario analysis: We present an algorithm for conducting historical scenario analysis within a fully specified nonlinear structural model.

In historical scenario analysis, nonlinearity implies that the economic conditions shaped by past structural shocks can affect the magnitude of shock deviations in counterfactual exercises. Therefore, accurately recovering these structural shocks is essential for the model's reliability. This approach nests the prevalent SVAR model, where the reduced-form model can be consistently estimated, and the structural shocks identified through a linear transformation of the reduced-form forecast errors.

A key distinction from our approach under linearity is that, in this method, we recover the sequence of historical shocks while manipulating only the selected policy shocks to closely replicate the counterfactual policy path. By combining these optimal hypothetical policy shocks with the remaining historical shocks, we construct the counterfactual economy based on the well specified model. The model's nonlinearity captures the private sector's response to the differing economic conditions in the counterfactual scenario, addressing these reactions in an econometric manner.

Future scenario analysis: We propose a simulation-based approach to assess the performance of macroeconomic aggregates under an alternative policy path. This method involves simulating all potential future trajectories of the structural shocks.

Algorithm 1 Historical scenario analysis

1: **Input:** $\boldsymbol{g}_t(\cdot)$, historical data $\boldsymbol{w}_{1:T}; \boldsymbol{w}_0$, and counterfactual policy path $\tilde{\boldsymbol{r}}_{t:t+H}$

2: Using the historical data $\boldsymbol{w}_{1:T}; \boldsymbol{w}_0$, recover the historical shocks $\hat{\epsilon}_{1:T}$

3: Determine the optimal shock deviation $\boldsymbol{\delta}_H$,

$$\boldsymbol{\delta}_H^{(ht)} = \underset{\boldsymbol{\delta}_H}{\operatorname{argmin}} \sum_{h=0}^{H} |\tilde{r}_{t+h} - g_{r,t+h}(\hat{\epsilon}_{t,h,(S)} - S_h \boldsymbol{\delta}_H, \hat{U}_{t+h})|^2$$

where S_h is defined in (5.3), $\hat{\epsilon}_{t,h,(S)}$ and $\hat{U}_{t+h}$ are estimated historical selected policy shocks and the remaining shocks up to horizon h.

4: **Output:** Compute the hypothetical the output variable,

$$\tilde{y}_{t+h} = g_{y,t+h}(\hat{\epsilon}_{t,h,(S)} - S_h \boldsymbol{\delta}_H^{(ht)}, \hat{U}_{t+h}),$$

and the output disparity,

$$\psi_h = y_{t+h} - \tilde{y}_{t+h}.$$

This procedure is comparable to the valuation of financial derivatives and options, where researchers simulate numerous future shock sequences, calculate the option price for each sequence, and then take the arithmetic average. In a similar vein, we simulate N sequences of future shocks and assess the shock deviation $\boldsymbol{\delta}_H$ for each sequence. Notably, due to the model's non-linearity, this shock deviation could be path-dependent. The economic performance across these scenarios forms the basis for generating the distribution used in the future scenario analysis.

The algorithm described involves a scenario in which the policy variable is fixed along a predetermined path, denoted as $\tilde{\boldsymbol{r}}_{t:t+H}$. For a given simulated sequence of future structural shocks that generate the baseline economy, the policymaker in the counterfactual scenario intervenes in the market through selected policy shocks of path-dependent magnitude $\boldsymbol{\delta}_H^{(i)}$. The path-dependence raises due to the non-linearity. The intervention of the policymaker is to ensure that the policy variable adheres to the predetermined path $\tilde{\boldsymbol{r}}_{t:t+H}$. Consequently, both the baseline and counterfactual economies can be computed using the simulated structural shocks and the path-dependent shock deviations.

After conducting a large number of simulations, we generate the distribution of future outcomes for both the baseline and counterfactual economies. The mean of the baseline economy, as the shocks are drawn from the unconditional empirical distribution, represents the unconditional mean of future performance. Conversely, the mean of the counterfactual economy, influenced by the selected policy shock deviations, reflects the conditional mean of future performance, given the predetermined policy path and the structural interpretation of the market intervention through the selected policy shocks. This interpretation implies that only the selected policy shocks deviate from their unconditional distribution to incorporate the conditional information from the policy path.

Policy intervention: In the previous discussion, we defined the policy intervention

Algorithm 2 Future scenario analysis

Input: $\boldsymbol{g}_t(\cdot), \boldsymbol{w}_{1:t-1}, \boldsymbol{w}_0$, and counterfactual policy path $\tilde{\boldsymbol{r}}_{t:t+H}$

Recover the historical shocks $\hat{\boldsymbol{\epsilon}}_{1:t-1}$,

Construct the empirical distribution for $\boldsymbol{\varepsilon}_t$ by assuming $\boldsymbol{\varepsilon}_t$ has i.i.d. distribution.

Set number of simulations N

for $i \leftarrow 1$ to N **do**

Draw a sequence of structural shocks $\boldsymbol{\epsilon}^{(i)}_{t:t+H}$ from the empirical distribution

Compute the economy performance, $\boldsymbol{w}^{(i)}_{t+h} = \boldsymbol{g}_{t+h}(\hat{\epsilon}_{1:t-1}, \epsilon^{(i)}_{t:t+h})$ for $0 \leq h \leq H$

Compute optimal shock deviation $\boldsymbol{\delta}^{(i)}_H$,

$$\boldsymbol{\delta}^{(i)}_H = \underset{\|\boldsymbol{\delta}_H\|}{\operatorname{argmin}} \sum_{h=0}^{H} |\tilde{r}_{t+h} - g_{r,t+h}(\epsilon^{(i)}_{t,h,(S)} - S_h \boldsymbol{\delta}_H, \hat{U}^{(i)}_{t+h})|^2$$

where $\tilde{r}_{t+h}$ is the given counterfactual policy variable, S_h is defined in (5.3), $\{\epsilon^{(i)}_{t,h,(S)}\} = \{\epsilon^{(i)}_{t,H,(S)}\} \backslash \{\epsilon^{(i)}_{t+h+1:t+H}\}$, and $\{\hat{U}^{(i)}_{t+h}\} = \{\epsilon_{1:t-1}, \epsilon^{(i)}_{t:t+h}\} \backslash \{\epsilon^{(i)}_{t,h,(S)}\}$.

Compute the performance of the output variable,

$$\begin{aligned} y^{(i)}_{t+h} &= g_{y,t+h}(\epsilon^{(i)}_{t,h,(S)}, \hat{U}^{(i)}_{t+h}) \\ \tilde{y}^{(i)}_{t+h} &= g_{y,t+h}(\epsilon^{(i)}_{t,h,(S)} - S_h \boldsymbol{\delta}^{(i)}_H, \hat{U}^{(i)}_{t+h}) \end{aligned}$$

end for

Output: The output disparity,

$$\psi^{(i)}_h = y^{(i)}_{t+h} - \tilde{y}^{(i)}_{t+h},$$

and the average output disparity is $\bar{\psi}_h = \frac{1}{N} \sum_{i=1}^{n} \psi^{(i)}_h$.

effect as:

$$\phi_h := Y_{t+h}\{1, R_{t:t+H}(0)\} - Y_{t+h}\{0, R_{t:t+H}(0)\}. \tag{7.1}$$

Algorithm 3 Zeroing-out policy intervention

ht

1: **Input:** $\boldsymbol{g}_t(\cdot)$

2: Set the number of simulations N

3: **for** $i \leftarrow 1$ to N **do**

3.1: Draw a sequence of structural shocks $\epsilon_{1:T}^{(i)}$ from the empirical distribution for ε_t.

3.2: Compute the term $Y_{t+h}\{0, R_{t:t+H}(0)\}$ as

$$Y_{t+h}\{0, R_{t:t+H}(0)\} = g_{y,t+h}(\epsilon_{x,t} = 0, \epsilon_{1:t+h}^{(i)} \backslash \epsilon_{x,t}).$$

3.3: Compute the performance of the policy variable in the absence of the shock

$$\tilde{r}_{t+h}^{(i)} = g_{r,t+h}(\epsilon_{x,t} = 0, \epsilon_{1:t+h}^{(i)} \backslash \epsilon_{x,t}),$$

for $h = 0, 1, \cdots, H$.

3.4: Compute the shock deviation $\boldsymbol{\delta}_H^{(i)}$,

$$\boldsymbol{\delta}_H^{(i)} = \underset{\|\boldsymbol{\delta}_H\|}{\operatorname{argmin}} \sum_{h=0}^{H} |\tilde{r}_{t+h}^{(i)} - g_{r,t+h}(\epsilon_{x,t} = 1, \epsilon_{t,h,(S)}^{(i)} - S_h \boldsymbol{\delta}_H, U_{t+h}^{(i)} \backslash \epsilon_{x,t})|^2.$$

3.5: Compute the term $Y_{t+h}\{1, R_{t:t+H}(0)\}$ as

$$Y_{t+h}\{1, R_{t:t+H}(0)\} = g_{y,t+h}(\epsilon_{x,t} = 1, \epsilon_{t,h,(S)}^{(i)} - S_h \boldsymbol{\delta}_H, U_{t+h}^{(i)} \backslash \epsilon_{x,t}).$$

3.6: Obtain $\phi_h^{(i)}$ by subtracting the result of step 3.5 from step 3.2,

$$\phi_h^{(i)} = g_{y,t+h}(\epsilon_{x,t} = 1, \epsilon_{t,h,(S)}^{(i)} - S_h \boldsymbol{\delta}_H, U_{t+h}^{(i)} \backslash \epsilon_{x,t}) - g_{y,t+h}(\epsilon_{x,t} = 0, \epsilon_{1:t+h}^{(i)} \backslash \epsilon_{x,t}).$$

4: **end for**

5: **Output:** The average policy intervention effect,

$$\bar{\phi}_h = \frac{1}{N} \sum_{i=1}^{N} \phi_h^{(i)}.$$

Policy intervention consists of two key steps. First, given a sequence of realized structural shocks, we compute the benchmark $Y_{t+h}\{0, R_{t:t+H}(0)\}$. Next, we calculate the policy shock deviation by minimizing the distance between the non-responsive policy path and the policy path in the counterfactual economy, where only selected policy shocks deviate. We then assess the performance of the output variable in this deliberately constructed counterfactual economy. The policy intervention effect is obtained by subtracting the benchmark output from the counterfactual output. It is important to note that the policy intervention effect can be a random variable in a nonlinear setting. Therefore, we take the average of the N simulated effects to estimate the average policy intervention.

In practice, policymakers generally have access to historical data reflecting the current state of the economy. With a well-specified parametric model, structural shocks can be recovered, making the state-dependent and path-dependent effects particularly relevant. The key distinction is that the state-dependent effect relies solely on past information, treating future shocks as random, whereas the path-dependent effect incorporates all information available up to the period when the effect is realized. For example, a central bank might find the state-dependent effect especially useful when formulating future monetary policy to stimulate the economy under current conditions, while the path-dependent effect may be more applicable when evaluating the impact of historical policies. Accordingly, we introduce these two extended definitions: the state-dependent effect and the path-dependent policy intervention effect,

$$\mathbb{E}[\bar{\phi}_h \mid \epsilon_{1:t-1}], \qquad \mathbb{E}[\bar{\phi}_h \mid \epsilon_{1:t-1}, \epsilon_{-x,t}, \epsilon_{t+1:t+h}], \tag{7.2}$$

In empirical macroeconomics, state-dependent and path-dependent causal effects are widely employed, although the average causal effect is generally more prevalent in microeconometrics for representing population averages. The path-dependent causal effect is particularly valuable in scenario analysis and historical policy evaluation. By capturing the impact of a shock given the entire sequence of the shocks, this effect is well-suited for addressing research questions such as evaluating the effectiveness of monetary policy during certain historical episode.

8 Empirical applications

In this section, we apply our approach to evaluate monetary policy in three counterfactual exercises: (1) historical scenario analysis, (2) future scenario analysis, and (3) zero-out policy intervention.

8.1 Historical scenario analysis

In the post-pandemic era, individuals have endured persistently high inflation rates. In March 2022, the Federal Reserve (Fed) initiated its first interest rate hike, signalling the end of the zero-interest-rate policy. Throughout 2022 and 2023, the Fed continued to incrementally raise interest rates as part of its strategy to counter the rising inflation. However, considerable scholarly debate persists over whether the Fed should have acted sooner and less aggressively in raising rates to more effectively curb inflation while reducing the burden of high interest costs, such as those related to residential mortgage payments.

A review of the data reveals that the Fed's first post-pandemic interest rate hike occurred in March 2022, at which point the inflation rate had reached 8%. Notably, by mid-2021, inflation had already surged to 5%, a significant increase from approximately 1% just six months earlier. Given the persistent nature of inflation as a macroeconomic indicator, it is reasonable to hypothesize that inflation would have continued to escalate if monetary

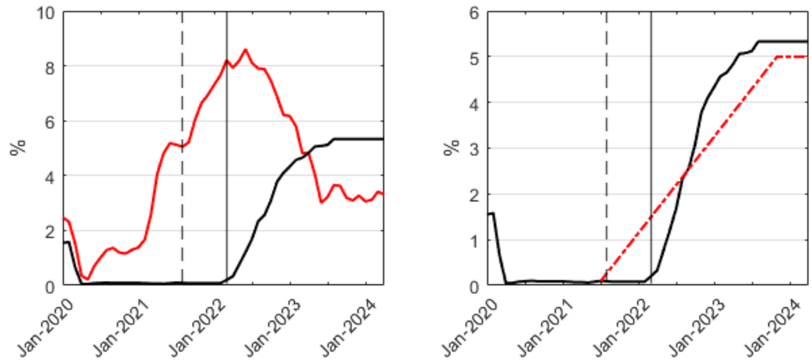

Figure 1: The left figure presents the historical trajectories of the Federal Funds Rate (FFR) (solid black) and the Consumer Price Index (CPI) inflation rate (solid red) from January 2020 to April 2024. The right figure depicts the historical trajectory of the FFR (solid black) and a counterfactual trajectory of the FFR (dot-dashed red). The solid vertical lines in both figures denote March 2022, indicating the timing of the first interest rate hike in the post-pandemic era within the actual economy. The dashed vertical lines, on the other hand, indicate August 2021, marking the period for the first interest rate hike in a counterfactual scenario.

policy had remained at near-zero interest rates. This raises an important question: should the Fed have considered raising interest rates as early as mid-2021, when inflation had already reached 5%?

To explore this question, we conduct a hypothetical trajectory analysis, simulating a counterfactual scenario in which interest rate hikes begin in August 2021, when inflation reached 5%. In this scenario, interest rates are gradually increased at a slower pace, reaching 3.5% by mid-2023. Subsequently, the rate is pushed to 5% over six months and maintained at that level until April 2024. This analysis aims to evaluate the potential impact of such a policy shift on inflation and industrial production indices, addressing two key research questions: (1) whether an earlier response by the Fed could have mitigated high inflation, and (2) whether a less aggressive response would have yielded better outcomes. We employ the hypothetical trajectory method proposed in this paper, utilizing the identification and estimation method for the causal effect of monetary policy shocks with an external instrumental variable, specifically the high-frequency rate of return of U.S. Treasury yields within the time window of Federal Open Market Committee (FOMC) meetings (Bauer and Swanson (2023), hereafter BS)[13]. The primary output variables are monthly U.S. industrial production and the CPI inflation rate. Data sources for these variables are obtained from the Federal Reserve Bank of St. Louis.

The data spans from January 1973 to December 2019 on a monthly basis. We identify and estimate the causal effect of monetary policy using a 4-variate SVAR-IV model with fourteen lags, including the FFR, industrial production, inflation rate, and excess bond premium. The number of lags is determined using the R command "VARselect" with lag.max = 18 and type = c("both"). The output from "VARselect" suggests 14, 2, 2, and 14 lags for AIC, HQ, SC, and FPE, respectively. The instrumental variable for monetary policy is the orthog-

[13]See data source at https://www.michaeldbauer.com/publication/mps/.

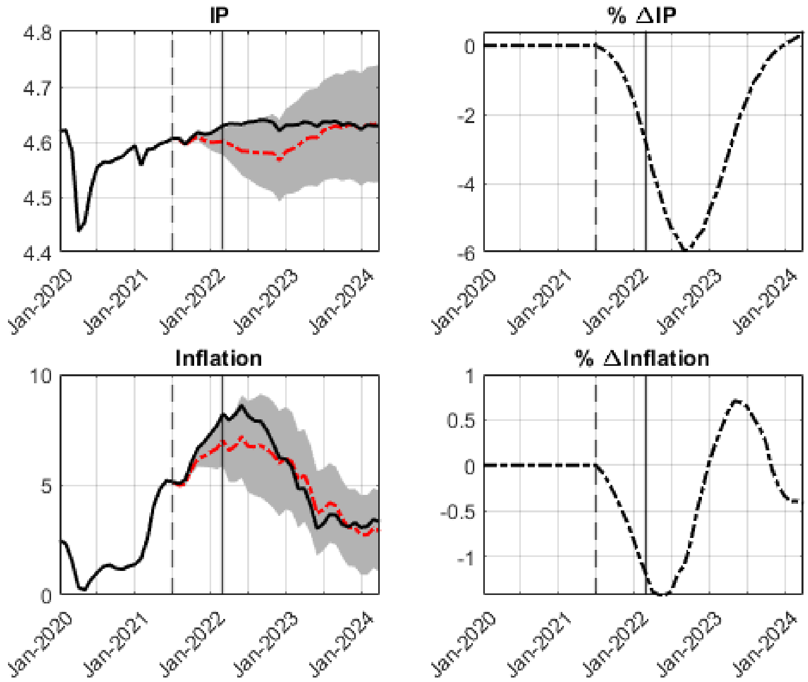

Figure 2: The figures illustrate both the historical and a hypothetical trajectory of industrial production (IP, log) and the CPI inflation rate. The upper-left panel shows the historical IP (black solid line) alongside the counterfactual IP (red dot-dashed line), based on the hypothetical policy path outlined in Figure 1. The lower-left panel displays the historical inflation rate (black solid line) and the counterfactual inflation rate (red dot-dashed line) under the same hypothetical policy path. The upper-right panel highlights the percentage difference, showing how much lower the counterfactual IP is compared to the historical IP. The lower-right panel presents the percentage difference in inflation rates. In all panels, the solid vertical lines denote March 2022, marking the first interest rate hike in the post-pandemic economy, while the dashed vertical lines indicate August 2021, the timing of the first interest rate hike in the counterfactual scenario.

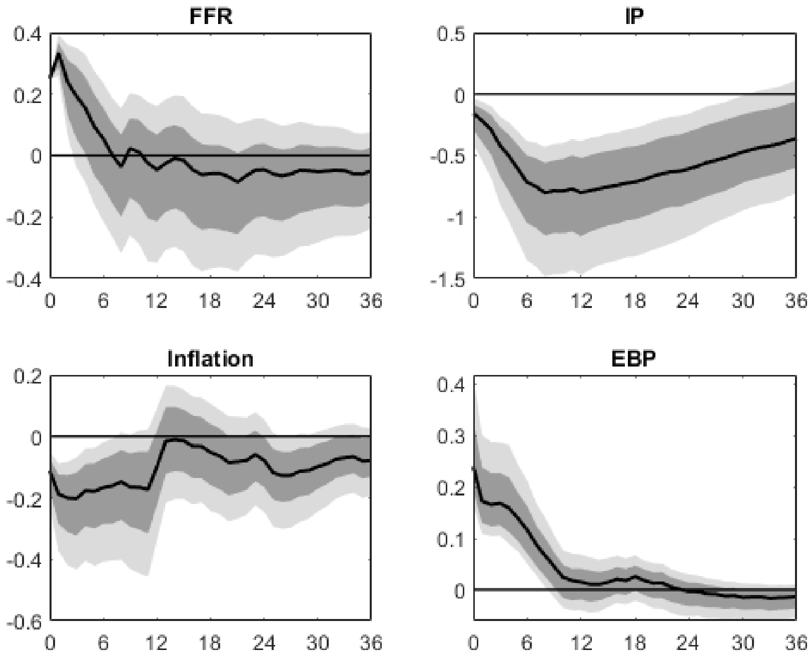

Figure 3: The structural impulse response to a 25 basis point monetary policy shock is presented, with the x-axis representing months and the y-axis indicating percentages. The dark grey shading represents the 68% confidence interval, while the light grey shading denotes the 90% confidence interval.

onalized monetary policy surprise (MPS ORTH) measure published by BS. The structural impulse response to a 25 bps monetary policy shock on the FFR is presented in Figure 3. The confidence interval is computed using the Wild bootstrap Gonçalves and Kilian (2004), as descried in page 1225 of Mertens and Ravn (2013).

We compute the hypothetical trajectory parameters for the desired output variables. As depicted in Figure 2, the hypothetical path involves an earlier interest rate hike compared to the actual timeline. In this scenario, we predict that industrial production would have been impacted differently. Specifically, the earlier rate hike would have lowered the inflation rate from its peak of 8.6% to 7% by June 2022. The cost of this intervention, however, would have been a 5% reduction in industrial production by June 2022 compared to the actual outcome. Nonetheless, this decrease in industrial production is temporary, as our analysis indicates that industrial output would rebound to actual levels by 2024, with the inflation rate potentially falling to 2.5% compared to 3% in reality.

In early 2020, the entire society experienced an exogenous shock from the COVID-19 pandemic, prompting the Fed to lower interest rates to near-zero levels to support the economy. This accommodative monetary policy, coupled with fiscal stimulus measures, led to an unsustainable rise in inflation, with individuals bearing the brunt of increased prices for goods and energy. The Fed's subsequent intervention through interest rate hikes further burdened households with higher costs on debt services. Our counterfactual analysis suggests that a more proactive monetary policy might have been beneficial. Indeed, the high inflation rate post-pandemic was likely unavoidable due to the necessary monetary and fiscal policies implemented to sustain livelihoods. However, the 5% inflation rate observed in mid-2021 should have served as an early warning sign for the Fed, and the decision to delay rate hikes until March 2022 warrants criticism for its tardy response. Our analysis reveals that the cost of reducing the inflation rate by 1.5% would have been a 5% reduction in industrial production. This result is primarily attributed to the impulse response estimates derived from the SVAR-IV model, consistent with the findings in BS for the sample period from January 1973 to December 2019. While it is reasonable to debate these results—specifically, the statement that a 25 bps interest rate hike could cause a 0.8% decrease in industrial production—these estimates are dependant upon the sample period, the number of VAR lags, the choice of FFR or 2-year interest rate proxy variables, and whether the MPS instrument is orthogonalized or not. This empirical estimation variation on impulse responses is also discussed in Ramey (2016). Ultimately, our analysis may stimulate a broader discussion on the historical trajectory of monetary policy path and its implications for interest rate decisions.

8.2 Future scenario analysis

In this exercise of future scenario analysis, our primary objective is to assess the economy's performance under an alternative interest rate trajectory, as opposed to the unconditional forecast. The data set comprises four key variables: industrial production, the CPI inflation rate, the federal funds rate (FFR), and the excess bond premium (EBP). This monthly data spans from January 1973 to December 2019. The instrumental variable for monetary policy is the orthogonalized monetary policy surprise (MPS ORTH) measure, as published by BS, covering the period from February 1988 to December 2019.

The unconditional forecast for industrial production (IP) is derived from the most recent observable value of IP (April 2024) and its growth rate. The forecast of the growth rate starts Q2, 2024 until Q2, 2025 is 1.6, 1.2, 1.7, 1.5, and 1.8 percentage, and the beyond Q2, 2025, we take growth rate 1.7 percentage. These growth rate is based on the Second Quarter 2024 Survey of Professional Forecasters. For the FFR, we utilize the projected appropriate policy path from the same economic projections, as reported in the Survey of Professional Forecasters (SPF) conducted by the Federal Reserve Bank of Philadelphia. As forecasts are provided on a quarterly or annual basis, we apply linear extrapolation methods. Our primary focus in this scenario is to evaluate the economy's performance, with particular attention to industrial production, when the Federal Reserve implements to an interest rate path that diverges from the unconditional forecast. This analysis holds profound implications for policymakers, providing insights into the potential costs and benefits associated with different monetary policy approaches.

Recently, the Federal Reserve (Fed) announced at its March 2024 meeting that it would maintain the overnight federal funds rate within the current range of 5.25% to 5.5%, in response to higher-than-expected inflation data at the start of the year. This decision follows the January FOMC meeting, where the central bank emphasized the need for more evidence of decelerating prices before initiating rate cuts. This decision underscores the trade-off between controlling inflation and projected economic growth, prompting critical questions: Is this strategy effective? Would the economy perform better or worse if the central bank adopted a more or less confrontational stand in response to inflationary pressures?

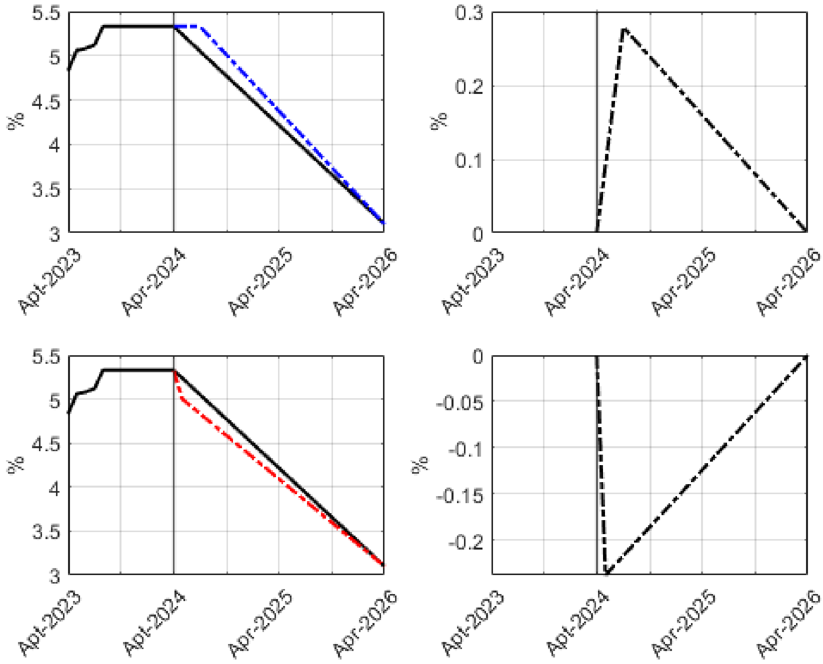

Figure 4: Two scenarios, dovish (red dot-dash) and hawkish (blue dot-dash), on future path of FFR, and the baseline path (black solid) is the unconditional forecast.

To explore these questions, we consider two alternative interest rate paths—one reflecting a dovish stance and the other a hawkish stance. Under the dovish policy, the interest rate is reduced more rapidly than the professional projection, decreasing by 50 basis points from 4.2% to 3.7% within a year before converging with the projected rate of 3.1% within two years. The hawkish policy involves maintaining the interest rate for an additional three months (one quarter) before initiating the planned rate cuts. We then evaluate the consequences of these two counterfactual scenarios on both inflation and industrial production.

Figure 5 illustrates that under the hawkish interest rate plan, where the rate cut is delayed by a quarter, industrial production decreases by 1.3% relative to the baseline within one year, while the CPI inflation rate is reduced by 0.35%, returning to the 2% target. Conversely, the dovish plan, with a 25 basis point more reduction in the first month than the SPF forecast, boosts industrial production by 1.2% but raises the inflation rate 0.3% up in one year. In 24 months, the hawkish plan results in a roughly 2% decrease in industrial production, with the inflation rate stabilizing at the 2% benchmark, while the dovish plan leads to a 4% increase in industrial output but leaves inflation at 2.5%. Given Chairman Jerome Powell emphasized the Fed remains "fully committed" to bringing inflation down to its 2% target, perhaps the Fed is expected to maintain the high interest rate for a few more months or have a slow than the expected plan on interest rate cut. Our quantitative analysis is designed to equip policymakers and investors with critical insights as they consider future policy rate path.

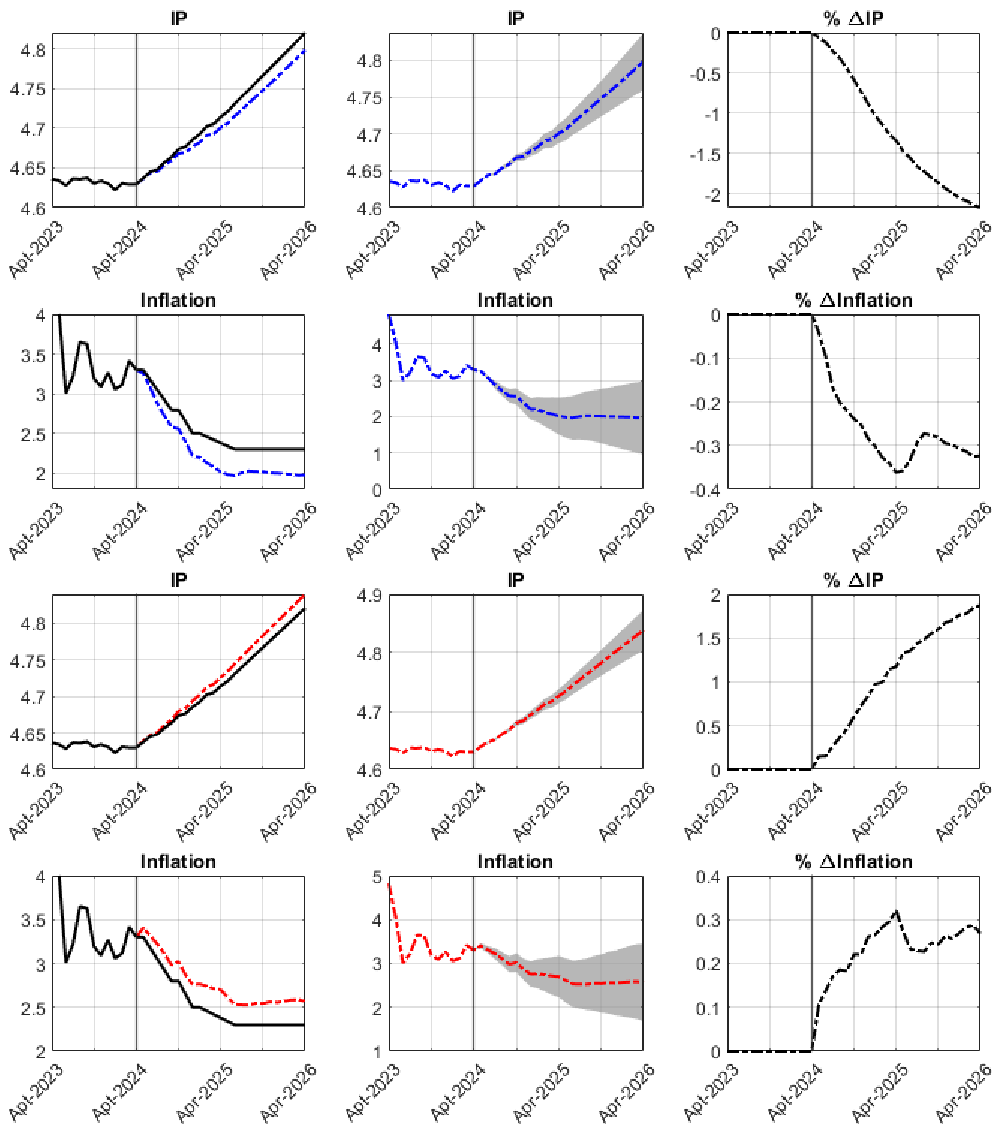

Figure 5: Economy performance under counterfactual monetary policy path. The left panel presents the SPF forecast (black solid) alongside the counterfactual performance (red dot-dash) under the dovish or hawkish monetary policy path. The middle panel displays the counterfactual performance, with the grey shading representing the 68% confidence interval. The right panel shows the percentage difference between the counterfactual performance and the SPF forecast. The blue dotted line represents the counterfactual outcome under the hawkish monetary policy path, while the red dotted line indicates the counterfactual outcome under the dovish monetary policy path.

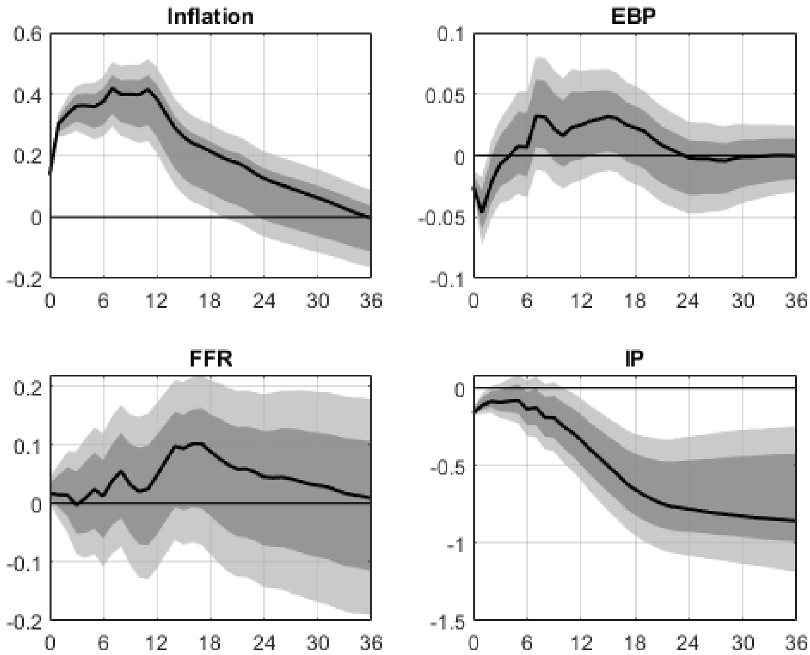

Figure 6: Impulse responses to 10% increase of WTI oil price shock, with the x-axis representing months and the y-axis indicating percentages. The dark grey shading represents the 68% confidence interval, while the light grey shading denotes the 90% confidence interval.

8.3 Zeroing out policy intervention

In this subsection, we undertake a thought experiment where the Federal Reserve (Fed) remains unresponsive to an oil price shock. Our objective is to evaluate the causal effect of this shock on industrial production and the inflation rate within this counterfactual scenario in which interest rates remain unresponsive. Through this analysis, we aim to quantitatively examine the central bank's role in the causal transmission of an oil price shock and emphasize the Fed's critical role in maintaining economic stability by systematically adjusting interest rates in response to anticipated inflation pressure.

The impulse responses to a positive 10% oil price shock are illustrated in Figure 6. It is computed through a five-variable SVAR-IV model with twelve lags, including the WTI oil price (log), industrial production (log), the federal funds rate (FFR), the CPI inflation rate, and the excess bond premium (EBP), spanning from January 1975 to December 2019. The external instrument for the oil price shock is the monthly "Oil Supply News Shocks" provided by Känzig (2021), a high-frequency instrument constructed from changes in oil futures prices within the daily window of OPEC meetings. It is observed in Figure 6 that the inflation rate is observed to rise by 0.4% within 12 months, with a gradual decline thereafter. The raised oil prices lead to a reduction in industrial production by approximately 0.4% within 12 months and 0.8% within 24 months. The FFR remains nearly unresponsive during the first six months following the oil price shock, subsequently increasing by 0.1% over the next 15 months before gradually declining. This pattern suggests that consumer inflation quickly reacts to rising oil prices, while the Fed's response is delayed, activating only after sustained

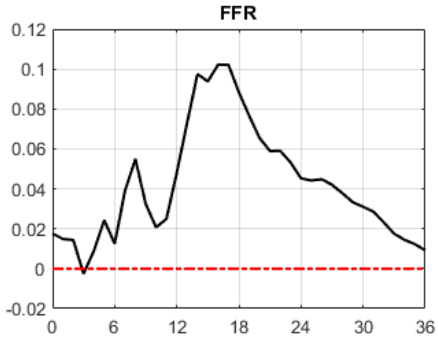

Figure 7: Zeroing-out impulse responses of FFR to 10% oil price shock. The red dot line represents the counterfactual zero response of FFR to oil price shock.

high inflation over several months. The chronological sequence of events following the oil price shock reflects the causal transmission of the shock through the economy and the Fed's primary focus on controlling inflation other than being direct responsible to high oil price.

In this thought experiment, the central bank deviates from its policy commitment and maintains a constant interest rate, effectively not responding to the oil price shock. Figure 8 shows that in the absence of policy intervention, the FFR remains at zero, and the industrial production and inflation rates in the first six months mirror the impulse response without policy intervention. It is due to the reason that the FFR maintain close to zero line in the first 6 months. However, as inflation gradually increases due to the zero interest rate, which should rise endogenously by 0.1%, the inflation rate surpasses the baseline by 0.1% at 12 months and approximately 0.2% at 36 months. Conversely, industrial production experiences a 0.2% increase in output at 12 months and approximately 0.4% at 36 months. The trade-off between output and inflation in response to systematic interest rate adjustments following an oil price shock is evident: a 0.2% reduction in output corresponds to a 0.1% higher inflation rate at 12 months, and a 0.4% reduction in output corresponds to a 0.2% higher inflation rate at 36 months. Generally, the Fed appears to trade two units of output for each unit of inflation to uphold its commitment to maintaining a stable inflation rate.

This ratio is derived from a counterfactual analysis within a linear framework, and while our study is subject to Lucas' critique, we argue that our counterfactual remains valid to some extent, particularly within a 12-month horizon. This validity stems from the fact that the Fed's deviation from its systematic response is relatively minor, likely going unnoticed by private sector players. However, the reliability of the counterfactual output at 24 months is less certain, as the Fed's significant deviation from its historical behavior may alter private sector responses. Nevertheless, our analysis provides valuable insights into the trade-offs between output and inflation when the Fed faces an oil price shock and highlights the central bank's essential role in maintaining price stability within the economy.

8.4 Empirical implications

We emphasize several key points relevant to practical applications. First, the counterfactual parameters under consideration are nonlinear transformations of impulse response functions, meaning their accurate estimation heavily depends on the identification and

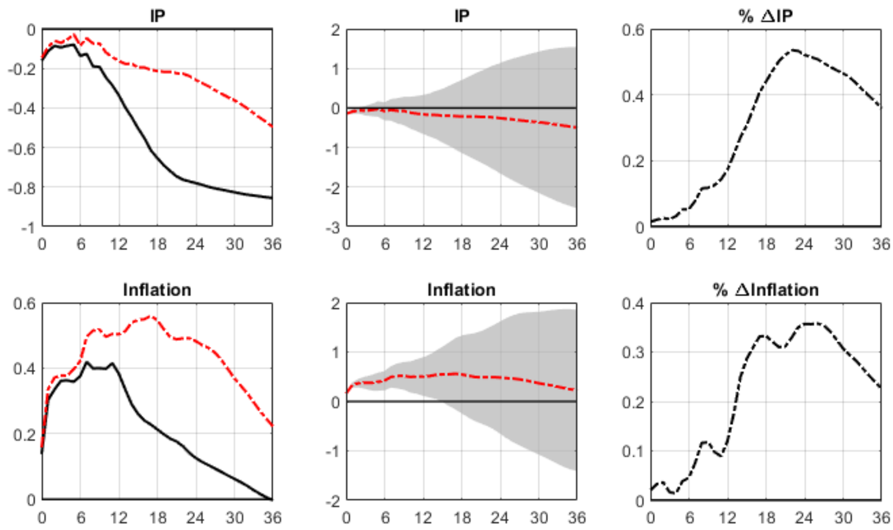

Figure 8: Zero-out policy intervention. The left panel displays the impulse response (black solid) alongside the policy intervention response (red dot-dashed), assuming the interest rate remains unresponsive to oil price changes. The middle panel presents the policy intervention response, with the 68% confidence interval shaded in grey. The right panel illustrates the difference between the policy intervention response and the corresponding impulse response.

estimation of these responses in practice. For instance, Ramey (2016) outlines various identification methods for monetary policy shocks, highlighting the variability in impulse response patterns across different methods and subsamples. As shown in Figure 3.1 on page 115 of Ramey (2016), the recursive identification method of Christiano et al. (1999) yields notably different impulse response patterns for industrial production between the 1965-1995 subsample (U-shaped) and the 1983-2007 subsample (flat). Researchers who strongly believe that GDP contracts following a positive monetary policy shock may select identification and estimation methods that align with their economic assumptions. Thus, the choice of identification strategies, estimation techniques, and sample periods plays a critical role in obtaining accurate counterfactual impulse response estimates. However, these issues lie beyond the scope of this paper, which focuses specifically on counterfactual analysis.

A common challenge in practical applications is achieving estimation precision, particularly given the well-known trade-off between LP-IV and SVAR-IV methods. While LP-IV is robust to non-recoverable SVMA components, it tends to produce less efficient estimates compared to SVAR-IV. This trade-off becomes especially pertinent in the context of counterfactual impulse response estimation. For example, when estimating policy intervention impulse responses over a 36-period horizon (equivalent to three years in monthly data), 37 control variables are required, as all contemporaneous and future variables must remain constant. In a typical macroeconomic setup, employing 37 control variables along with a few lags in a Local Projection equation designed to address data persistence can significantly reduce the degrees of freedom, potentially resulting in imprecise estimates of policy intervention impulse responses.

Furthermore, similar issues arise with hypothetical trajectory parameters when the deviation path spans a long time horizon. The limited number of observations in macroeconomic data presents a persistent challenge for Local Projection estimation. To address these concerns, Li et al. (2024) propose penalized Local Projection as an alternative, demonstrating that it effectively navigates the bias-variance trade-off compared to standard Local Projection and SVAR-IV methods. Practitioners seeking improved estimation precision should consider incorporating penalization techniques into their estimation equations.

This paper does not rule out the use of SVAR-IV for estimation and inference. If practitioners are willing to assume invertibility in the SVAR model and impose conditions like partial invertibility on the shock of interest, the precision of impulse response estimation can improve, particularly in cases with moderate SVAR dimensions and shorter VAR orders. However, longer SVAR orders, such as 12 or 18, commonly used in studies of oil price shocks, may still present challenges in terms of estimation precision. In these instances, researchers might consider high-dimensional SVAR models with LASSO methods, incorporating sparsity assumptions, as demonstrated in recent studies by Adamek et al. (2023) and Krampe et al. (2023).

9 Conclusion

In this paper, we introduce a robust approach to address two common counterfactual scenarios in macroeconomics: hypothetical trajectory and policy intervention within a linear framework. Our methodology eliminates the need for a fully specified structural model or policy rule equation to assess the economy under hypothetical scenarios, focusing instead on the identification of policy shocks. Specifically, we derive unique analytical solutions, along with statistical inference, for the parameters governing these counterfactuals by interpreting policy path deviation as manipulations of selected policy shocks. Our results establish a connection between macroeconomic counterfactual analysis and the expanding literature on shock identification and impulse response estimation.

In addition to providing analytical solutions for the counterfactuals, we introduce a linear projection-based method for parameter identification. These counterfactual parameters are identified and estimated using external instruments, extending the Local Projection-Instrumental Variable (LP-IV) framework. This approach highlights the potential of Local Projection techniques in counterfactual analysis. Furthermore, we derive statistical inference for the linear projections by employing two methods for estimating the covariance matrix: the HAC/HAR method and the Eicker-Huber-White HR method, with the latter requiring slightly stronger assumptions on the disturbances. Our contribution to covariance matrix estimation advances the existing literature on statistical inference in time series models.

We apply our innovative methods to analyze U.S. monetary policy through a series of counterfactual exercises. Specifically, we forecast the hypothetical economy's performance under alternative interest rate paths across three distinct scenarios. The first scenario examines a hypothetical post-pandemic case, evaluating how the economy would have evolved had the Federal Reserve implemented earlier and less aggressive interest rate hikes. The second scenario explores potential future dovish or hawkish monetary policy, assessing the economy's response if the Federal Reserve deviates from the SPF interest rate path. The third scenario focuses on policy intervention, analyzing the Federal Reserve's role in the causal transmission of oil price shock by suppressing the systematic response of interest rate. This analysis provides valuable insights into the central bank's influence on oil shock transmission and its commitment to maintaining inflation stability.

References

R. Adamek, S. Smeekes, and I. Wilms. Lasso inference for high-dimensional time series. *Journal of Econometrics*, 235(2):1114–1143, 2023.

J. D. Angrist and G. M. Kuersteiner. Causal effects of monetary shocks: Semiparametric conditional independence tests with a multinomial propensity score. *Review of Economics and Statistics*, 93(3):725–747, 2011.

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.

J. Antolin-Diaz, I. Petrella, and J. F. Rubio-Ramirez. Structural scenario analysis with SVARs. *Journal of Monetary Economics*, 117:798–815, 2021.

M. Bańbura, D. Giannone, and M. Lenza. Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of forecasting*, 31 (3):739–756, 2015.

R. Barnichon and G. Mesters. A sufficient statistics approach for macro policy. *American Economic Review*, 113(11):2809–2845, 2023.

R. M. Baron and D. A. Kenny. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6):1173, 1986.

M. D. Bauer and E. T. Swanson. A reassessment of monetary policy surprises and high-frequency identification. *NBER Macroeconomics Annual*, 37(1):87–155, 2023.

C. Baumeister and L. Kilian. Do oil price increases cause higher food prices? *Economic Policy*, 29(80):691–747, 2014.

P. Beaudry, F. Portier, and A. Preston. *Some Inference Perils of Imposing a Taylor Rule*. Centre for Economic Policy Research, 2023.

M. Beraja. A semistructural methodology for policy counterfactuals. *Journal of Political Economy*, 131(1):190–201, 2023.

B. S. Bernanke, M. Gertler, and M. Watson. Systematic monetary policy and the effects of oil price shocks. *Brookings papers on economic activity*, 1997(1):91–157, 1997.

O. J. Blanchard and D. Quah. The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, pages 655–673, 1989.

I. Bojinov and N. Shephard. Time series experiments and causal estimands: exact randomization tests and trading. *Journal of the American Statistical Association*, 114(528): 1665–1682, 2019.

G. E. Box and G. C. Tiao. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70(349):70–79, 1975.

M. Bruns and H. Lütkepohl. Comparison of local projection estimators for proxy vector autoregressions. *Journal of Economic Dynamics and Control*, 134:104277, 2022.

V. Celli. Causal mediation analysis in economics: Objectives, assumptions, models. *Journal of Economic Surveys*, 36(1):214–234, 2022.

R. Chetty. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, 1(1):451–488, 2009.

T. E. C. A. M. Christian and K. Wolf. Evaluating policy counterfactuals: A "var-plus" approach. *Research Paper*, 2024.

L. J. Christiano, M. Eichenbaum, and C. L. Evans. Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, 1:65–148, 1999.

L. J. Christiano, M. Eichenbaum, and C. L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, 113(1):1–45, 2005.

J. S. Davis and A. Zlate. Monetary policy divergence and net capital flows: Accounting for endogenous policy responses. *Journal of International Money and Finance*, 94:15–31, 2019.

M. Del Negro and F. Schorfheide. Monetary policy analysis with potentially misspecified models. *American Economic Review*, 99(4):1415–1450, 2009.

E. Dettaa and E. Wang. Inference in high-dimensional linear projections: Multi-horizon granger causality and network connectedness. *Technical Report, McGill University, Economic Department*, 2024.

T. Doan, R. Litterman, and C. Sims. Forecasting and conditional projection using realistic prior distributions. *Econometric reviews*, 3(1):1–100, 1984.

J.-M. Dufour. Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics*, 133(2):443–477, 2006.

J.-M. Dufour and D. Pelletier. Practical methods for modeling weak VARMA processes: Identification, estimation and specification with a macroeconomic application. *Journal of Business & Economic Statistics*, 40(3):1140–1152, 2022.

J.-M. Dufour and E. Renault. Short run and long run causality in time series: theory. *Econometrica*, 66(5):1099–1125, 1998.

J.-M. Dufour and E. Wang. Simple robust two-stage estimation and inference for generalized impulse responses and multi-horizon causality. *Technical Report, McGill University, Economic Department*, 2024a.

J.-M. Dufour and E. Wang. Causal mechanism and mediation analysis for macroeconomics dynamics. *Technical Report, McGill University, Economic Department*, 2024b.

J.-M. Dufour, D. Pelletier, and É. Renault. Short run and long run causality in time series: inference. *Journal of Econometrics*, 132(2):337–362, 2006.

R. A. Fisher. *Statistical methods for research workers*. Oliver and Boyd, 1st ed. edition, 1925.

M. Gertler and P. Karadi. Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76, 2015.

S. Gonçalves and L. Kilian. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*, 123(1):89–120, 2004.

S. Gonçalves, A. M. Herrera, L. Kilian, and E. Pesavento. Impulse response analysis for structural dynamic models with nonlinear regressors. *Journal of Econometrics*, 225(1): 107–130, 2021.

S. Gonçalves, A. M. Herrera, L. Kilian, and E. Pesavento. State-dependent local projections. *Journal of Econometrics*, page 105702, 2024.

J. D. Hamilton. *Time Series Analysis*. Princeton: Princeton University Press, 1994.

J. D. Hamilton and A. M. Herrera. Oil shocks and aggregate macroeconomic behavior: the role of monetary policy. *Journal of Money, Credit and Banking*, 36(2):265–286, 2004.

J. D. Hamilton, S. Pruitt, and S. Borger. Estimating the market-perceived monetary policy rule. *American Economic Journal: Macroeconomics*, 3(3):1–28, 2011.

E. P. Herbst and B. K. Johannsen. Bias in local projections. *Journal of Econometrics*, 240(1): 105655, 2024.

K. Imai, L. Keele, and T. Yamamoto. Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1):51–71, 2010.

K. Imai, D. Tingley, and T. Yamamoto. Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 176(1):5–51, 2013.

A. Inoue and L. Kilian. Inference on impulse response functions in structural VAR models. *Journal of Econometrics*, 177(1):1–13, 2013.

A. Inoue and L. Kilian. Joint confidence sets for structural impulse responses. *Journal of Econometrics*, 192(2):421–432, 2016.

A. Inoue and L. Kilian. The uniform validity of impulse response inference in autoregressions. *Journal of Econometrics*, 215(2):450–472, 2020.

A. Inoue and B. Rossi. A new approach to measuring economic policy shocks, with an application to conventional and unconventional monetary policy. *Quantitative Economics*, 12(4):1085–1138, 2021.

M. Jarociński. Conditional forecasts and uncertainty about forecast revisions in vector autoregressions. *Economics Letters*, 108(3):257–259, 2010.

M. Jarociński. Estimating the fed's unconventional policy shocks. *Journal of Monetary Economics*, 144:103548, 2024.

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

Ò. Jordà. Local projections for applied economics. *Annual Review of Economics*, 15(1): 607–631, 2023.

D. R. Känzig. The macroeconomic effects of oil supply news: Evidence from opec announcements. *American Economic Review*, 111(4):1092–1125, 2021.

L. Kilian and Y. J. Kim. How reliable are local projection estimators of impulse responses? *The Review of Economics and Statistics*, 93(4):1460–1466, 2011.

L. Kilian and T. K. Lee. Quantifying the speculative component in the real price of oil: The role of global oil inventories. *Journal of International Money and Finance*, 42:71–87, 2014.

L. Kilian and L. T. Lewis. Does the fed respond to oil price shocks? *The Economic Journal*, 121(555):1047–1072, 2011.

L. Kilian and H. Lütkepohl. *Structural Vector Autoregressive Analysis*. Cambridge: Cambridge University Press, 2017.

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

J. Krampe, E. Paparoditis, and C. Trenkler. Structural inference in sparse high-dimensional vector autoregressions. *Journal of Econometrics*, 234(1):276–300, 2023.

E. M. Leeper and T. Zha. Modest policy interventions. *Journal of Monetary Economics*, 50 (8):1673–1700, 2003.

R. Lewis and G. C. Reinsel. Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis*, 16(3):393–411, 1985.

D. Li, M. Plagborg-Møller, and C. K. Wolf. Local projections vs. vars: Lessons from thousands of dgps. *Journal of Econometrics*, page 105722, 2024.

R. E. Lucas Jr. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. North-Holland, 1976.

A. Lusompa. Local projections, autocorrelation, and efficiency. *Quantitative Economics*, 14 (4):1199–1220, 2023.

A. McKay and C. K. Wolf. What can time-series regressions tell us about policy counterfactuals? *Econometrica*, 91(5):1695–1725, 2023.

K. Mertens and M. O. Ravn. The dynamic effects of personal and corporate income tax changes in the united states. *American economic review*, 103(4):1212–1247, 2013.

S. Miranda-Agrippino and G. Ricco. Identification with external instruments in structural VARs. *Journal of Monetary Economics*, 135:1–19, 2023.

J. L. Montiel Olea and M. Plagborg-Møller. Local projection inference is simpler and more robust than you think. *Econometrica*, 89(4):1789–1823, 2021.

E. Nakamura and J. Steinsson. High-frequency identification of monetary non-neutrality: the information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330, 2018.

H. H. Pesaran and Y. Shin. Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1):17–29, 1998.

M. Plagborg-Møller and C. K. Wolf. Local projections and VARs estimate the same impulse responses. *Econometrica*, 89(2):955–980, 2021.

M. Plagborg-Møller and C. K. Wolf. Instrumental variable identification of dynamic variance decompositions. *Journal of Political Economy*, 130(8):2164–2202, 2022.

A. Rambachan and N. Shephard. When do common time series estimands have nonparametric causal meaning. *URL: https://scholar. harvard. edu/shephard/publications/nonparametric-dynamic-causal-modelmacroeconometrics*, 2021.

V. A. Ramey. Macroeconomic shocks and their propagation. *Handbook of Macroeconomics*, 2:71–162, 2016.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.

C. A. Sims and T. Zha. Does monetary policy generate recessions?: Using less aggregate price data to identify monetary policy. *Paper, Yale University*, 782, 1995.

C. A. Sims and T. Zha. Does monetary policy generate recessions? *Macroeconomic Dynamics*, 10(2):231–272, 2006.

C. A. Sims, J. H. Stock, and M. W. Watson. Inference in linear time series models with some unit roots. *Econometrica*, 58(1):113–144, 1990.

J. H. Stock and M. W. Watson. Vector autoregressions. *Journal of Economic perspectives*, 15 (4):101–115, 2001.

J. H. Stock and M. W. Watson. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128(610):917–948, 2018.

D. F. Waggoner and T. Zha. Conditional forecasts in dynamic multivariate models. *The Review of Economics and Statistics*, 81(4):639–651, 1999.

H. White. *Asymptotic theory for econometricians, revised edition*. Academic press, San Diego, Florida, 1999.

I. Wilms, S. Basu, J. Bien, and D. S. Matteson. Sparse identification and estimation of large-scale vector autoregressive moving averages. *Journal of the American Statistical Association*, pages 1–12, 2021.

K.-L. Xu. Local projection based inference under general conditions. *Technical Report, Indiana University*, 2023.

A An alternative identification method

The alternative approach involves replacing equation (4.8) with the following specification:

$$Y_{t+h} = \phi_h X_t + \beta'_{h,H} R_{t:t+H} + v^*_{t,h}, \tag{A.1}$$

where $v^*_{t,h} = v_{t,h} - \phi_h(X_t - \varepsilon_{x,t})$.

The advantage of equation (A.1) over (4.8) is that it does not require the partial identification assumption for $\varepsilon_{x,t}$. However, there are two key disadvantages: (1) it necessitates a contemporaneous orthogonality assumption, specifically

$$\mathrm{P}(X_t \mid \varepsilon_{r,t}) = 0, \tag{A.2}$$

and (2) it requires a valid instrument to identify the shock to $\varepsilon_{x,t}$.

The contemporaneous orthogonality assumption is necessary because, similar to (4.11), the identification of the parameter ϕ_h requires the following zero projection equation:

$$\mathrm{P}(v^*_{t,h} \mid \varepsilon_{x,t}, \varepsilon_{t,H,(S)}) = 0. \tag{A.3}$$

Given that $v^*_{t,h} = u_{t,h} - \phi_h X_t$ and $u_{t,h}$ is orthogonal to the policy shocks $\varepsilon_{t,H,(S)}$, the equation holds only if

$$\mathrm{P}(X_t \mid \varepsilon_{t,H,(S)}) = 0. \tag{A.4}$$

Since the selected policy shocks $\varepsilon_{t,H,(S)}$ may include contemporaneous policy shocks, it is required that the structural shock $\varepsilon_{r,t}$ has no contemporaneous causal effect on the variable x_t. This assumption can be satisfied, for instance, if a lower-triangular identification method is used and X_t is ordered before the policy variable R_t. If this assumption is violated, then X_t may respond to changes in the policy variable $R_{t:t+H}$. In some empirical applications, X_t refers to the oil price, and $\epsilon_{r,t}$ is the monetary policy shock. If researchers believe that monetary policy has no contemporaneous causal effect on oil prices (as in Bernanke et al. (1997) and Kilian and Lewis (2011)), it is valid to use X_t as the dependent variable, and the coefficient matrix will still be $\Theta_{r,x}$.

Given the contemporaneous orthogonality assumption and a valid instrument for the shock $\varepsilon_{x,t}$, denoted as $z_{x,t}$, we obtain the following moment equation:

$$\mathbb{E}[(y_{t+h} - \phi_h X_t - \beta'_{h,H} R_{t:t+H})(\boldsymbol{z}'_{t,H}, z_{x,t})'] = \mathbf{0} \tag{A.5}$$

This is because both regressors, X_t and $R_{t:t+H}$, are endogenous to the residual. In contrast to the first identification method, where equation (4.8) does not impose a restriction on the contemporaneous causal effect but requires the identification of $\varepsilon_{x,t}$, equation (A.1) relaxes the partial identification condition while imposing the requirement of no contemporaneous causal effect and the existence of a valid instrument.

B Lemmas

Lemma B.1. *Let W_t follows equation (2.3) and Assumption 4.2* and 5.1 are satisfied. Then*

$$T^{-1}\sum_{t=1}^{T-h} \boldsymbol{z}_{t,H} R^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h} \tag{B.1}$$

$$T^{-1}\sum_{t=1}^{T-h} \boldsymbol{z}_{x,t,H} R^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h} \tag{B.2}$$

$$T^{-1/2}\sum_{t=1}^{T-h} \boldsymbol{z}_{t,H} u^{\perp}_{t,h} \xrightarrow{d} N(0, \Omega_{r,h}) \tag{B.3}$$

$$T^{-1/2}\sum_{t=1}^{T-h} \boldsymbol{z}_{x,t,H} v^{\perp}_{t,h} \xrightarrow{d} N(0, \Omega_{x,h}) \tag{B.4}$$

Lemma B.2. *Let W_t follows equation (2.3) and Assumption 4.2* and 5.1 are satisfied. Then*

$$T^{-1}\sum_{t=1}^{T-h} \boldsymbol{z}_{t,H} \hat{R}^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h} \tag{B.5}$$

$$T^{-1}\sum_{t=1}^{T-h} \hat{\boldsymbol{z}}_{x,t,h} \hat{R}^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h} \tag{B.6}$$

The proof of the Lemma B.1-B.2 is in Appendix C.7.

C Proofs

C.1 Proof of Proposition 6.1

Since the parameter ψ_h and ϕ_h are functions of $(\theta_{yx,h}, \boldsymbol{d}_H^{(po)'}, \mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h})'$, given Assumption (6.1), we apply Delta-method (linearization) and obtain the asymptotic statistical inference for $\hat{\psi}_h$ and $\hat{\phi}_h$. Now, we show the derivation for the explicit form of the Jacobian matrices, G, J case-by-case due to the Moore–Penrose inverse depends on the column(row) linear independence.

Recall $\boldsymbol{\beta}_{h,H} = (\Theta^{-'}_{re,H})\boldsymbol{\theta}_{ye,h}$, $\psi_h = \boldsymbol{\beta}'_{h,H}(\boldsymbol{r}_{t:t+H} - \tilde{\boldsymbol{r}}_{t:t+H})$, and $\phi_h = \theta_{yx,h} - \boldsymbol{\beta}'_{h,H}\boldsymbol{d}_H^{(po)}$, and G, J are Jacobian matrices,

$$G = \partial \boldsymbol{\beta}_{h,H} / \partial (\mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h}), \tag{C.1}$$

$$J = \partial \phi_h / \partial (\theta_{yx,h}, \boldsymbol{d}_H^{(po)'}, \mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h}). \tag{C.2}$$

Write G as

$$G = \frac{\partial \Theta^{-'}_{re,H}\boldsymbol{\theta}_{ye,h}}{\partial (\mathrm{vec}(\Theta^{-'}_{re,H})', \boldsymbol{\theta}'_{ye,h})} \frac{\partial (\mathrm{vec}(\Theta^{-'}_{re,H})', \boldsymbol{\theta}'_{ye,h})'}{\partial (\mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h})}, \tag{C.3}$$

and J as

$$\begin{aligned} J &= \frac{\partial (\theta_{yx,h} - \boldsymbol{\beta}'_{h,H}\boldsymbol{d}_H^{(po)})}{\partial (\theta_{yx,h}, \boldsymbol{d}_H^{(po)'}, \boldsymbol{\beta}'_{h,H})} \frac{\partial (\theta_{yx,h}, \boldsymbol{d}_H^{(po)'}, \boldsymbol{\beta}'_{h,H})'}{(\theta_{yx,h}, \boldsymbol{d}_H^{(po)'}, \mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h})} \\ &= \begin{bmatrix} 1 & , & -\boldsymbol{\beta}'_{h,H} & , & \boldsymbol{d}_H^{(po)'} G \end{bmatrix}. \end{aligned} \tag{C.4}$$

Then, the explicit form of J depends on G and we obtain the explicit form of G by showing the value of two fractions. The first fraction equals

$$\frac{\partial (\Theta^{-'}_{re,H})\boldsymbol{\theta}_{ye,h}}{\partial (\mathrm{vec}(\Theta^{-'}_{re,H})', \boldsymbol{\theta}'_{ye,h})} = [\boldsymbol{\theta}'_{ye,h} \otimes I_{H+1}, \Theta^{-'}_{re,H}]; \tag{C.5}$$

and the second fraction equals

$$\frac{\partial (\mathrm{vec}(\Theta^{-'}_{re,H}), \boldsymbol{\theta}'_{ye,h})'}{\partial (\mathrm{vec}(\Theta_{re,H})', \boldsymbol{\theta}'_{ye,h})} = \begin{bmatrix} \partial \mathrm{vec}(\Theta^{-'}_{re,H}) / \partial \mathrm{vec}(\Theta_{re,H})' & 0 \\ 0 & I_{n_e} \end{bmatrix}. \tag{C.6}$$

The explicit form of $\partial \mathrm{vec}(\Theta^{-'}_{re,H}) / \partial \mathrm{vec}(\Theta_{re,H})'$ depends on the form of Moore-Penrose inverse. We derive its form on the case-by-case basis.

Case 1: $\Theta^{-}_{re,H} = (\Theta'_{re,H}\Theta_{re,H})^{-1}\Theta'_{re,H}$. The term $\partial \mathrm{vec}(\Theta^{-'}_{re,H}) / \partial \mathrm{vec}(\Theta_{re,H})'$ has explicit form of

$$
\begin{aligned}
&\partial \text{vec}(\Theta^{-'}_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&\partial \text{vec}(\Theta_{re,H}(\Theta'_{re,H}\Theta_{re,H})^{-1})/\partial \text{vec}(\Theta_{re,H})' \\
=&\left((\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes I_{H+1}\right) \partial \text{vec}(\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
&+ \left(I_{n_e} \otimes \Theta_{re,H}\right) \partial \text{vec}((\Theta'_{re,H}\Theta_{re,H})^{-1})/\partial \text{vec}(\Theta_{re,H})' \\
=&(\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes I_{H+1} \\
&+ \left(I_{n_e} \otimes \Theta_{re,H}\right) \partial \text{vec}((\Theta'_{re,H}\Theta_{re,H})^{-1})/\partial \text{vec}(\Theta'_{re,H}\Theta_{re,H})' * \partial \text{vec}(\Theta'_{re,H}\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&(\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes I_{H+1} \\
&- \left(I_{n_e} \otimes \Theta_{re,H}\right)\left((\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes (\Theta'_{re,H}\Theta_{re,H})^{-1}\right) \partial \text{vec}(\Theta'_{re,H}\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&(\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes I_{H+1} \\
&- \left(I_{n_e} \otimes \Theta_{re,H}\right)\left((\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes (\Theta'_{re,H}\Theta_{re,H})^{-1}\right)\left((\Theta'_{re,H} \otimes I_{n_e})K_{(H+1),n_e} + I_{n_e} \otimes \Theta'_{re,H}\right) \\
=&(\Theta'_{re,H}\Theta_{re,H})^{-1} \otimes (I_{H+1} - \Theta^{-'}_{re,H}\Theta'_{re,H}) - \left(\Theta^{-}_{re,H} \otimes \Theta^{-'}_{re,H}\right) K_{(H+1),n_e}
\end{aligned}
\tag{C.7}
$$

where we apply the formula $\partial \text{vec}(A^{-1})/\partial \text{vec}(A)' = -A^{'-1} \otimes A^{-1}$ for any square full rank matrix A, and the second last equality is because

$$
\begin{aligned}
&\partial \text{vec}(\Theta'_{re,H}\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&(\Theta'_{re,H} \otimes I_{n_e})\partial \text{vec}(\Theta'_{re,H})/\partial \text{vec}(\Theta_{re,H})' + (I_{n_e} \otimes \Theta'_{re,H})\partial \text{vec}\Theta_{re,H}/\partial \text{vec}(\Theta_{re,H})' \\
=&(\Theta'_{re,H} \otimes I_{n_e})K_{(H+1),n_e}\partial \text{vec}(\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' + (I_{n_e} \otimes \Theta'_{re,H})\partial \text{vec}\Theta_{re,H}/\partial \text{vec}(\Theta_{re,H})' \\
&(K_{n,m} \text{ is the commutation matrix, } \text{vec}(A') = K_{n,m}\text{vec}(A), \text{ for a } m \times n\ A \text{ matrix A.}) \\
=&(\Theta'_{re,H} \otimes I_{n_e})K_{(H+1),n_e} + I_{n_e} \otimes \Theta'_{re,H}.
\end{aligned}
\tag{C.8}
$$

Thus, the Jacobian matrix G has the form

$$
G = \left[(\boldsymbol{\theta}'_{ye,h}(\Theta'_{re,H}\Theta_{re,H})^{-1}) \otimes (I_{H+1} - \Theta^{-'}_{re,H}\Theta'_{re,H}) - \left(\boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H} \otimes \Theta^{-'}_{re,H}\right)K_{(H+1),n_e} \quad , \quad \Theta^{-'}_{re,H}\right]
\tag{C.9}
$$

Case 2: $\Theta^{-}_{re,H} = \Theta^{-1}_{re,H}$. This is a special case of Case 1 as $\Theta^{-}_{re,H} = \Theta^{-1}_{re,H}$ implies that $\Theta_{re,H}$ is a full rank square matrix. Thus,

$$
\partial \text{vec}(\Theta^{-'}_{re,H})/\partial \text{vec}(\Theta_{re,H})' = -\left(\Theta^{-1}_{re,H} \otimes \Theta^{-1'}_{re,H}\right)K_{H+1,H+1}.
\tag{C.10}
$$

Thus, the Jacobian matrix G has the form

$$
G = \left[-(\boldsymbol{\theta}'_{ye,h}\Theta^{-1}_{re,H} \otimes \Theta^{-1'}_{re,H})K_{H+1,H+1} \quad , \quad \Theta^{-'}_{re,H}\right].
\tag{C.11}
$$

Case 3: $\Theta^{-}_{re,H} = \Theta'_{re,H}(\Theta_{re,H}\Theta'_{re,H})^{-1}$. The term $\partial \text{vec}(\Theta^{-'}_{re,H})/\partial \text{vec}(\Theta_{re,H})'$ has explicit form of

$$
\begin{aligned}
&\partial \text{vec}(\Theta^{-'}_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&\partial \text{vec}((\Theta_{re,H}\Theta'_{re,H})^{-1}\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&\left(I_{n_e} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1}\right)\partial \text{vec}(\Theta_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
&+\left(\Theta'_{re,H} \otimes I_{H+1}\right)\partial \text{vec}((\Theta_{re,H}\Theta'_{re,H})^{-1})/\partial \text{vec}(\Theta_{re,H})' \\
=&I_{n_e} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} \\
&+\left(\Theta'_{re,H} \otimes I_{H+1}\right)\partial \text{vec}((\Theta_{re,H}\Theta'_{re,H})^{-1})/\partial \text{vec}(\Theta_{re,H}\Theta'_{re,H})' * \partial \text{vec}(\Theta_{re,H}\Theta'_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&I_{n_e} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} \\
&-\left(\Theta'_{re,H} \otimes I_{H+1}\right)\left((\Theta_{re,H}\Theta'_{re,H})^{-1} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1}\right)\partial \text{vec}(\Theta_{re,H}\Theta'_{re,H})/\partial \text{vec}(\Theta_{re,H})' \\
=&I_{n_e} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} \\
&-\left(\Theta'_{re,H} \otimes I_{H+1}\right)\left((\Theta_{re,H}\Theta'_{re,H})^{-1} \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1}\right)\left((\Theta_{re,H} \otimes I_{H+1}) + (I_{H+1} \otimes \Theta_{re,H})K_{(H+1),n_e}\right) \\
=&(I_{n_e} - \Theta^{-}_{re,H}\Theta_{re,H}) \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} - \left(\Theta^{-}_{re,H} \otimes \Theta^{-'}_{re,H}\right)K_{(H+1),n_e}
\end{aligned}
\tag{C.12}
$$

Thus, the Jacobian matrix G has the form

$$
G = \left[(\boldsymbol{\theta}'_{ye,h}(I_{n_e} - \Theta^{-}_{re,H}\Theta_{re,H})) \otimes (\Theta_{re,H}\Theta'_{re,H})^{-1} - \left((\boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H}) \otimes \Theta^{-'}_{re,H}\right)K_{(H+1),n_e} \quad , \quad \Theta^{-'}_{re,H}\right]. \tag{C.13}
$$

C.2 Proof of Proposition 4.1

We prove from the right-hand-side of (4.7),

$$
\begin{aligned}
&\mathbb{E}[\boldsymbol{z}_{t,H}R'_{t:t+H}]^{-}\mathbb{E}[\boldsymbol{z}_{t,H}Y_{t+h}] \\
=&\mathbb{E}[(\Pi\boldsymbol{\varepsilon}_{t,H,(S)}+\boldsymbol{\eta}_{t,H})R'_{t:t+H}]^{-}\mathbb{E}[(\Pi\boldsymbol{\varepsilon}_{t,H,(S)}+\boldsymbol{\eta}_{t,H})Y_{t+h}] \\
&\text{(by definition of } \boldsymbol{z}_{t,H}) \\
=&\left(\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}R'_{t:t+H}]\right)^{-}\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}Y_{t+h}] \\
&\text{(by Assumption 4.2(ii))} \\
=&\left(\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\Theta'_{re,H}\right)^{-}\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\boldsymbol{\theta}_{ye,h} \\
&\text{(by SVMA equation 2.3 and structural shocks mutually uncorrelated)} \\
=&\Theta_{re,H}\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\Pi'\left(\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\Theta'_{re,H}\Theta_{re,H}\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\Pi'\right)^{-1} \\
&\Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}]\boldsymbol{\theta}_{ye,h} \\
&\text{(by the condition of full column rank , Assumption 4.2(ii) and full rankness of covariance matrix of } \varepsilon_t) \\
=&\Theta_{re,H}\left(\Theta'_{re,H}\Theta_{re,H}\right)^{-1}\boldsymbol{\theta}_{ye,h} \\
=&\Theta^{-'}_{re,H}\boldsymbol{\theta}_{ye,h} \\
&(\Theta_{re,H} \text{ has full column rank by the condition of full column rank)} \\
=&\boldsymbol{\beta}_{h,H}
\end{aligned}
\tag{C.14}
$$

Thus, the proposition is proved.

C.3 Proof of Proposition 4.2

We prove from the right-hand-side of (4.16),

$$
\begin{aligned}
& v(\mu_\theta)'\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]^{-}\mathbb{E}[\boldsymbol{z}_{x,t,H}Y_{t+h}] \\
=& v(\mu_\theta)'\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]'\left(\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}]'\right)^{-1}\mathbb{E}[\boldsymbol{z}_{x,t,H}Y_{t+h}] \\
& (\mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}] \text{ has full row rank by Assumption 4.2 and full-rank covariance matrix of } \varepsilon_t\ ) \\
=& v(\mu_\theta)'\Theta_{rx,H}\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1}(\theta_{yx,h}, \boldsymbol{\theta}'_{ye,h})' \\
& \left(\text{since } \mathbb{E}[\boldsymbol{z}_{x,t,H}R'_{x,t,H}] = \begin{bmatrix} \mathbb{E}[\varepsilon^2_{x,t}] & 0 \\ 0 & \Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}] \end{bmatrix}\Theta'_{rx,H},\right. \\
& \quad \text{by defining of } \Theta_{rx,H} := \begin{bmatrix} 1 & \mathbf{0} \\ \boldsymbol{d}_H^{(po)} & \Theta_{re,H} \end{bmatrix}, \\
& \quad \left.\text{and } E[\boldsymbol{z}_{x,t,H}Y_{t+h}] = \begin{bmatrix} \mathbb{E}[\varepsilon^2_{x,t}] & 0 \\ 0 & \Pi\mathbb{E}[\boldsymbol{\varepsilon}_{t,H,(S)}\boldsymbol{\varepsilon}'_{t,H,(S)}] \end{bmatrix}(\theta_{yx,h}, \boldsymbol{\theta}'_{ye,h})'\right) \\
=& v(\mu_\theta)'\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1}(\theta_{yx,h}, \boldsymbol{\theta}'_{ye,h})' \\
& (v(\mu_\theta)'\Theta_{rx,H} = v(\mu_\theta)' \text{ as } \Theta_{rx,H} = \begin{bmatrix} 1 & \mathbf{0} \\ \boldsymbol{d}_H^{(po)} & \Theta_{re,H} \end{bmatrix})
\end{aligned}
\tag{C.15}
$$

Then, we derive the inverse matrix, $\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1}$. Apply the block matrix inverse formula, we obtain

$$
\begin{aligned}
& v(\mu_\theta)'\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1} \\
=& \mu_\theta v(1)'\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1} \\
& (\text{simply write } v(\mu_\theta) = v(1)\mu_\theta \text{ by its definition}) \\
=& \mu_\theta v(1)'\begin{bmatrix} 1 + \boldsymbol{d}_H^{(po)\prime}\boldsymbol{d}_H^{(po)} & \boldsymbol{d}_H^{(po)\prime}\Theta_{re,H} \\ \Theta'_{re,H}\boldsymbol{d}_H^{(po)} & \Theta'_{re,H}\Theta_{re,H} \end{bmatrix}^{-1} \\
=& \mu_\theta\left[\mu_\theta^{-1}, \quad -\mu_\theta^{-1}\boldsymbol{d}_H^{(po)\prime}\Theta_{re,H}(\Theta'_{re,H}\Theta_{re,H})^{-1}\right] \\
& \left(\text{apply block matrix inverse formula, and}\right. \\
& \quad \left.\mu_\theta = 1 + \boldsymbol{d}_H^{(po)\prime}M_\Theta\boldsymbol{d}_H^{(po)} = 1 + \boldsymbol{d}_H^{(po)\prime}\boldsymbol{d}_H^{(po)} - \boldsymbol{d}_H^{(po)\prime}\Theta_{re,H}(\Theta'_{re,H}\Theta_{re,H})^{-1}\Theta'_{re,H}\boldsymbol{d}_H^{(po)}\right) \\
=& \left[1, \quad -\boldsymbol{d}_H^{(po)\prime}\Theta^{-\prime}_{re,H}\right]
\end{aligned}
\tag{C.16}
$$

Plug in the above result to (C.15), it yields

$$\begin{aligned}
&\nu(\mu_\theta)'\left(\Theta'_{rx,H}\Theta_{rx,H}\right)^{-1}(\theta_{yx,h},\boldsymbol{\theta}'_{ye,h})' \\
=&\left[1, \quad -\boldsymbol{d}_H^{(po)'}\Theta^{-'}_{re,H}\right](\theta_{yx,h},\boldsymbol{\theta}'_{ye,h})' \\
=&\theta_{yx,h}-\boldsymbol{\theta}'_{ye,h}\Theta^{-}_{re,H}\boldsymbol{d}_H^{(po)} \\
=&\phi_h.
\end{aligned} \tag{C.17}$$

The, the proposition is proved.

C.4 Proof of Proposition 5.1

First, we prove the asymptotic normality for $\tilde{\beta}_h$,

$$\begin{aligned}
\tilde{\beta}_h - \beta_h &= \left(\sum_{t=1}^{T-h} z_{t,H} R^{\perp'}_{t:t+h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{t,H} Y^{\perp}_{t+h}\right) - \beta_h \\
&= \left(\sum_{t=1}^{T-h} z_{t,H} R^{\perp'}_{t:t+h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{t,H} (R^{\perp'}_{t:t+h}\beta_h + u^{\perp}_{t,h})\right) - \beta_h \\
&= \left(\sum_{t=1}^{T-h} z_{t,H} R^{\perp'}_{t:t+h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{t,H} u^{\perp}_{t,h}\right)
\end{aligned} \tag{C.18}$$

Lemma B.1 shows that $T^{-1}\sum_{t=1}^{T-h} z_{t,H} R^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$ and $T^{-1/2}\sum_{t=1}^{T-h} z_{t,H} u^{\perp}_{t,h} \xrightarrow{d} N(0, \Omega_{r,h})$. Matrix $\Omega_{r,h}$ is invertible by Assumption 5.1 (iv). The invertibility of matrix $\Sigma_{zr,h}$ can be shown as the follows. Since $z_{t,H} = \Pi\varepsilon_{t,H,(S)} + \eta_{t,H}$ and $\eta_{t,H}$ is uncorrelated with any structural shocks, then $\Sigma_{zr,h} = \Pi \mathrm{Var}(\varepsilon_{t,H,(S)})\Theta'_{re,H}$. Since Π is assumed to be full rank (Assumption 4.2*), $\mathrm{Var}(\varepsilon_{t,H,(S)})$ has full rank, $\Theta_{re,H}$ is a square matrix under Assumption 4.2* (i) and thereby has full rank (Assumption 4.1), then the matrix $\Sigma_{zr,h}$ is a square matrix with full rank. By Slutsky's theorem, the asymptotic normality of $\tilde{\beta}_h$ with full-rank covariance matrix $\Sigma^{-1}_{zr,h}\Omega_{r,h}\Sigma'^{-1}_{zr,h}$ is proved.

Similarly, the infeasible estimate $\tilde{\phi}_h$

$$\begin{aligned}
&\tilde{\phi}_h - \phi_h \\
=&\nu(1)'\left(\sum_{t=1}^{T-h} z_{x,t,h} R^{\perp'}_{x,t,h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{x,t,h} Y^{\perp}_{t+h}\right) - \phi_h \\
=&\nu(1)'\left(\sum_{t=1}^{T-h} z_{x,t,h} R^{\perp'}_{x,t,h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{x,t,h} (\phi_h\varepsilon_{x,t} + \beta'_h R^{\perp}_{t:t+h} + \nu^{\perp}_{t,h})\right) - \phi_h \\
=&\nu(1)'\left(\sum_{t=1}^{T-h} z_{x,t,h} R^{\perp'}_{x,t,h}\right)^{-1} \left(\sum_{t=1}^{T-h} z_{x,t,h} \nu^{\perp}_{t,h}\right)
\end{aligned} \tag{C.19}$$

Lemma B.1 shows that $T^{-1}\sum_{t=1}^{T-h} z_{x,t,h} R^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h}$ and $T^{-1/2}\sum_{t=1}^{T-h} z_{x,t,h} \nu^{\perp}_{t,h} \xrightarrow{d} N(0, \Omega_{x,h})$. Since matrix $\Omega_{r,h}$ is invertible by Assumption 5.1 (iv) and $\Sigma_{zx,h}$ is invertible due to the full rankness of matrix $\Sigma_{zr,h}$, the asymptotic normality of $\tilde{\phi}_h$ is proved.

C.5 Proof of Proposition 5.2

We prove the convergence in distribution by showing (1) the feasible estimates are asymptotic equivalent with the infeasible estimates, (2) the sample estimate of the covariance matrix/variance converge to its population counterpart in the limit.

First, we prove the convergence of $\hat{\beta}_h$ by showing the equivalence of $\hat{\beta}_h$ and $\tilde{\beta}_h$,

$$
\begin{aligned}
&\sqrt{T}\|\hat{\beta}_h - \tilde{\beta}_h\| \\
=&\sqrt{T}\left\|\left(\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}\hat{Y}^{\perp}_{t+h}\right) - \left(\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}Y^{\perp}_{t+h}\right)\right\| \\
=&\sqrt{T}\left\|\left(\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp'}_{t:t+h}\beta_h + (R^{\perp'}_{t:t+h} - \hat{R}^{\perp'}_{t:t+h})\beta_h + u^{\perp}_{t,h} + \hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right)\right. \\
&\left. - \left(\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}(R^{\perp'}_{t:t+h}\beta_h + u^{\perp}_{t,h})\right)\right\| \\
=&\sqrt{T}\left\|\left(\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}((R^{\perp'}_{t:t+h} - \hat{R}^{\perp'}_{t:t+h})\beta_h + u^{\perp}_{t,h} + \hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right)\right. \\
&\left. - \left(\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h}\right)^{-1}\left(\sum_{t=1}^{T-h} z_{t,H}u^{\perp}_{t,h}\right)\right\| \\
\leq&\left\|\left(T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1} - \left(T^{-1}\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h}\right)^{-1}\right\|\left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}u^{\perp}_{t,h}\right\| \\
&+\left\|\left(T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\right\|\left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right\| \\
&+\left\|\left(T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h}\right)^{-1}\right\|\left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(R^{\perp'}_{t:t+h} - \hat{R}^{\perp'}_{t:t+h})\right\|\|\beta_h\| \\
&\xrightarrow{p} 0
\end{aligned}
\tag{C.20}
$$

The convergence can be shown if all three terms in the last inequality converge to zero asymptotically. The convergence of the first term can be checked as $T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$ (Lemma B.2), $T^{-1}\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$ (Lemma B.1), the full rankness of the matrix $\Sigma_{zr,h}$ has been verified in the Appendix 5.1, and $T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}u^{\perp}_{t,h}$ converges to a normal distribution (Lemma B.1) and therefore $T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}u^{\perp}_{t,h}$ is bounded by a constant almost surely (Chebyshev's inequality). Thus, the first term converges.

The convergence of the second term can be checked as $T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$

(Lemma B.2), and

$$
\begin{aligned}
&\left\| T^{-1/2} \sum_{t=1}^{T-h} z_{t,H} (\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h}) \right\| \\
&\leq \sum_{h=0}^{k} \pi \left\| T^{-1/2} \sum_{t=1}^{T-h} \varepsilon_{x,t+h} (\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h}) \right\| + \sum_{h=0}^{k} \pi \left\| T^{-1/2} \sum_{t=1}^{T-h} \eta_{t+h} (\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h}) \right\| \\
&\xrightarrow{p} 0.
\end{aligned}
\tag{C.21}
$$

The convergence is due to Assumption 5.2 and $\hat{Y}^{\perp}_{t+h}$ is an element in $\hat{W}^{\perp}_{t+h}$. The convergence of the third term can be checked since $\left\| T^{-1/2} \sum_{t=1}^{T-h} z_{t,H} (R^{\perp'}_{t:t+h} - \hat{R}^{\perp'}_{t:t+h}) \right\| \xrightarrow{p} 0$. The convergence is due to the same argument as (C.21) that $\boldsymbol{R}^{\perp}_{t:t+h}$ is a stacked of finite number of $R^{\perp}_{t+h}$ (component of $W^{\perp}_{t+h}$). Thus, we proved the equivalence of $\hat{\boldsymbol{\beta}}_h$ and $\tilde{\boldsymbol{\beta}}_h$.

Then, we discuss the convergence of the covariance matrix of $\hat{\sigma}^{(hac)}_{\psi,h} \xrightarrow{p} \sigma_{\psi,h}$, where $\hat{\sigma}^{(hac)}_{\psi,h} = (\boldsymbol{d}^{(ht)'}_h \hat{\Sigma}^{-1}_{zr,h} \hat{\Omega}^{(hac)}_{r,h} \hat{\Sigma}^{-1'}_{zr,h} \boldsymbol{d}^{(ht)}_h)^{1/2}$ and $\hat{\Sigma}_{zr,h} = T^{-1} \sum_{t=1}^{T-h} z_{t,H} \hat{R}^{\perp'}_{t:t+h}$. Lemma B.2 shows the convergence of $T^{-1} \sum_{t=1}^{T-h} z_{t,H} \hat{R}^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$ and Assumption 5.3 supposes the consistency of $\hat{\Omega}^{(hac)}_{r,h}$. Thus, the covariance estimates is consistent. In addition, the matrix $\sigma_{\psi,h}$ is non-singular since both matrices $\Sigma_{zr,h}$ and $\Omega_{r,h}$ are non-singular. Thus, the convergence $\sqrt{T} \left(\hat{\sigma}^{(hac)}_{\psi,h} \right)^{-1} \left(\hat{\psi}_h - \psi_h \right) \xrightarrow{d} N(0,1)$ is proved.

Next, we prove the equivalence of $\hat{\phi}_h$ and $\tilde{\phi}_h$,

$$
\begin{aligned}
&\sqrt{T}|\hat{\phi}_h - \tilde{\phi}_h| \\
=&\sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{Y}^{\perp}_{t+h}\right) - \tilde{\phi}_h\right\| \\
=&\sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h} + Y^{\perp}_{t+h})\right) - \tilde{\phi}_h\right\| \\
=&\sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\underbrace{(v^{\perp}_{t,h} + R^{\perp'}_{x,t,h}(\phi_h, \beta_h')')}_{=Y^{\perp}_{t+h}}\right)\right. \\
&\left. + v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right) - \tilde{\phi}_h\right\| \\
\leq&\sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}v^{\perp}_{t,h}\right) + \phi_h - \tilde{\phi}_h\right\| \\
&+ \sqrt{T}\left\|\left(T^{-1}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1} T^{-1/2}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(R^{\perp'}_{x,t,h} - \hat{R}^{\perp'}_{x,t,h})(\phi_h, \beta_h')'\right\| \qquad \text{(C.22)} \\
&+ \sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right)\right\| \\
\leq&\sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}z_{x,t,H}R^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}(\hat{z}_{x,t,h} - z_{x,t,H})v^{\perp}_{t,h}\right)\right\| \\
&+ \sqrt{T}\left\| v(1)'\left\{\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1} - \left(\sum_{t=1}^{T-h}z_{x,t,H}R^{\perp'}_{x,t,h}\right)^{-1}\right\}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}v^{\perp}_{t,h}\right)\right\| \\
&+ \sqrt{T}\left\|\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(R^{\perp'}_{x,t,h} - \hat{R}^{\perp'}_{x,t,h})(\phi_h, \beta_h')'\right\| \\
&+ \sqrt{T}\left\| v(1)'\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h}\right)^{-1}\left(\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(\hat{Y}^{\perp}_{t+h} - Y^{\perp}_{t+h})\right)\right\| \\
\xrightarrow{p}&0
\end{aligned}
$$

The convergence can be shown if all four terms in the last inequality converge to zero asymptotically. Note that $\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h}$ (Lemma B.2), $\sum_{t=1}^{T-h}z_{x,t,H}R^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h}$ (Lemma B.1), and $T^{-1/2}\sum_{t=1}^{T-h}z_{x,t,H}v^{\perp}_{t,h}$ converges to a normal distribution (Lemma B.1).

The convergence of the first term in the last equality of (C.22) can be checked by

$$\begin{aligned}
&T^{1/2}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h}-z_{x,t,h})v_{t,h}^{\perp}\\
=&T^{1/2}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})v_{t,h}^{\perp}\\
&\text{(due to the explicit form } \hat{z}_{x,t,h}=(\hat{\varepsilon}_{x,H},z'_{t:t+H})')\\
=&T^{1/2}\sum_{t=1}^{T-h}(\hat{X}_t^{\perp}-X_t^{\perp})v_{t,h}^{\perp}\\
&\text{(due to Assumption 4.3 } \varepsilon_{x,t}=\mathrm{P}(X_t\mid \overline{W}_{t-1}),\ \text{then } \hat{\varepsilon}_{x,t}-\varepsilon_{x,t}=\hat{X}_t^{\perp}-X_t^{\perp})\\
\xrightarrow{p}&0\\
&\text{(due to Assumption 5.2 (i) and } v_{t,h}^{\perp}=v_{t,h}^{\perp}-\phi_h\varepsilon_{x,t}\ )
\end{aligned}\tag{C.23}$$

The convergence of the second term in the last equality of (C.22) can be checked by the convergence in distribution of $T^{1/2}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}v_{t,h}^{\perp}$, which can be shown by (C.23) and the convergence in distribution of $T^{1/2}\sum_{t=1}^{T-h}z_{x,t,h}v_{t,h}^{\perp}$ by Lemma B.1.

The convergence of the third term in the last equality of (C.22) can be shown if $\|T^{1/2}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(R_{x,t,h}^{\perp'}-\hat{R}_{x,t,h}^{\perp'})\|\xrightarrow{p}0$. It is sufficient to prove $\|T^{1/2}\sum_{t=1}^{T-h}z_{x,t,h}(R_{x,t,h}^{\perp'}-\hat{R}_{x,t,h}^{\perp'})\|\xrightarrow{p}0$ and $\|T^{1/2}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h}-z_{x,t,h})(R_{x,t,h}^{\perp'}-\hat{R}_{x,t,h}^{\perp'})\|\xrightarrow{p}0$, that is

$$\begin{aligned}
&\|T^{1/2}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h}-z_{x,t,h})(R_{x,t,h}^{\perp'}-\hat{R}_{x,t,h}^{\perp'})\|\\
=&\|T^{1/2}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})(R_{x,t,h}^{\perp'}-\hat{R}_{x,t,h}^{\perp'})\|\\
&\text{(due to the explicit form } \hat{z}_{x,t,h}=(\hat{\varepsilon}_{x,H},z'_{t:t+H})')\\
=&\|T^{1/2}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})(R_{t:t+h}^{\perp'}-\hat{R}_{t:t+h}^{\perp'})\|+\|T^{1/2}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2\|\\
&\text{(due to the explicit form } \hat{R}_{x,t,H}=(\hat{\varepsilon}_{x,H},\hat{R}_{t:t+H})')\\
=&\|T^{1/2}\sum_{t=1}^{T-h}(\hat{X}_t^{\perp}-X_t^{\perp})(R_{t:t+h}^{\perp'}-\hat{R}_{t:t+h}^{\perp'})\|+\|T^{1/2}\sum_{t=1}^{T-h}(\hat{X}_t^{\perp}-X_t^{\perp})^2\|\\
&\text{(due to Assumption 4.3 } \varepsilon_{x,t}=\mathrm{P}(X_t\mid \overline{W}_{t-1}),\ \text{then } \hat{\varepsilon}_{x,t}-\varepsilon_{x,t}=\hat{X}_t^{\perp}-X_t^{\perp})\\
\xrightarrow{p}&0\\
&\text{(due to Assumption 5.2 (ii))}
\end{aligned}\tag{C.24}$$

and

$$
\begin{aligned}
&\left\|T^{-1/2}\sum_{t=1}^{T-h} z_{x,t,h}(R^{\perp'}_{x,t,h}-\hat{R}^{\perp'}_{x,t,h})\right\| \\
=&\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(R^{\perp'}_{x,t,h}-\hat{R}^{\perp'}_{x,t,h})\right\| + \left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(R^{\perp'}_{x,t,h}-\hat{R}^{\perp'}_{x,t,h})\right\| \\
&\text{(due to } z_{x,t,h}=(\varepsilon_{x,t}, z'_{t,H})'\text{)} \\
=&\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\right\| + \left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{R}_{t:t+H}-R_{t:t+H})\right\| \\
&+\left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\right\| + \left\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}_{t:t+H}-R_{t:t+H})'\right\| \\
&\text{(due to } \hat{R}^{\perp}_{x,t,H}=(\hat{\varepsilon}_{x,t},\hat{R}'_{t:t+H})'\text{)} \\
\leq&\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\right\| + \left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{R}_{t:t+H}-R_{t:t+H})\right\| \\
&+\sum_{h=0}^{k}\pi\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t+h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\right\| + \sum_{h=0}^{k}\pi\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t+h}(\hat{R}_{t:t+H}-R_{t:t+H})'\right\| \\
&+\sum_{h=0}^{k}\left\|T^{-1/2}\sum_{t=1}^{T-h} \eta_{t+h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\right\| + \sum_{h=0}^{k}\left\|T^{-1/2}\sum_{t=1}^{T-h} \eta_{t+h}(\hat{R}_{t:t+H}-R_{t:t+H})'\right\| \\
&\text{(due to the explicit form of } z_{t,H}\text{)} \\
=&\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{X}^{\perp}_{t}-X^{\perp}_{t})\right\| + \left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t}(\hat{R}_{t:t+H}-R_{t:t+H})\right\| \\
&+\sum_{h=0}^{k}\pi\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t+h}(\hat{X}^{\perp}_{t}-X^{\perp}_{t})\right\| + \sum_{h=0}^{k}\pi\left\|T^{-1/2}\sum_{t=1}^{T-h} \varepsilon_{x,t+h}(\hat{R}_{t:t+H}-R_{t:t+H})'\right\| \\
&+\sum_{h=0}^{k}\left\|T^{-1/2}\sum_{t=1}^{T-h} \eta_{t+h}(\hat{X}^{\perp}_{t}-X^{\perp}_{t})\right\| + \sum_{h=0}^{k}\left\|T^{-1/2}\sum_{t=1}^{T-h} \eta_{t+h}(\hat{R}_{t:t+H}-R_{t:t+H})'\right\| \\
&\text{(due to Assumption 4.3 } \varepsilon_{x,t}=\mathrm{P}(X_t\mid \overline{W}_{t-1}),\text{ then } \hat{\varepsilon}_{x,t}-\varepsilon_{x,t}=\hat{X}^{\perp}_{t}-X^{\perp}_{t}\text{)} \\
\xrightarrow{p}&0.
\end{aligned}
\tag{C.25}
$$

The convergence of each element in the last step is the immediate results of Assumption 5.2 (i).

The convergence of the fourth term in the last equality of (C.22) is due to

$$
\begin{aligned}
&\left\| T^{-1/2}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
=&\left\| T^{-1/2}\sum_{t=1}^{T-h}\hat{\varepsilon}_{x,H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| + \left\| T^{-1/2}\sum_{t=1}^{T-h}z_{t:t+H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
&\text{(due to the explicit form } \hat{z}_{x,t,h}=(\hat{\varepsilon}_{x,H},z'_{t:t+H})'\text{)} \\
\leq&\left\| T^{-1/2}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,H}-\varepsilon_{x,H})(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| + \left\| T^{-1/2}\sum_{t=1}^{T-h}\varepsilon_{x,H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
&+\left\| T^{-1/2}\sum_{t=1}^{T-h}z_{t:t+H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
=&\left\| T^{-1/2}\sum_{t=1}^{T-h}(\hat{X}^{\perp}_{t}-X^{\perp}_{t})(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| + \left\| (T^{-1/2}\sum_{t=1}^{T-h}\varepsilon_{x,H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
&+\left\| T^{-1/2}\sum_{t=1}^{T-h}z_{t:t+H}(\hat{Y}^{\perp}_{t+h}-Y^{\perp}_{t+h})\right\| \\
&\text{(due to Assumption 4.3 } \varepsilon_{x,t}=X_t-\mathrm{P}(X_t\mid \overline{W}_{t-1}),\text{ then } \hat{\varepsilon}_{x,t}-\varepsilon_{x,t}=\hat{X}^{\perp}_t-X^{\perp}_t\text{)} \\
\xrightarrow{p}&0
\end{aligned}
\tag{C.26}
$$

In the last equality, the first two terms in the last equality converge to zero due to Assumption 5.2 and the third term converges to zero shown in (C.21).

Thus, we proved the equivalence of $\hat{\beta}_h$ and $\tilde{\beta}_h$. the convergence of the asymptotic variance estimate $\hat{\sigma}^{(hac)}_{\phi,h}$ can be readily checked that $\hat{\Sigma}_{zx,h}=T^{-1}\sum_{t=1}^{T-h}z_{x,t,h}\hat{R}^{\perp\prime}_{x,t,h}\xrightarrow{p}\Sigma_{zx,h}$ by Lemma B.2 and Assumption 5.3 supposes the consistency of $\hat{\Omega}^{(hac)}_{x,h}$. In summary, the proof of $\sqrt{T}\left(\hat{\sigma}^{(hac)}_{\phi,h}\right)^{-1}\left(\hat{\phi}_h-\phi_h\right)\xrightarrow{d}N(0,1)$ is completed.

C.6 Proof of Proposition 5.3

First, we prove the variance of each sequence is identical to its long run variance. It is sufficient to show, $s^*_{r,t,h}$ and $s^*_{x,t,h}$ are serially correlated process.

For process $s^*_{r,t,h}$, we show its serial uncorrelation by assuming $t < s$,

$$\begin{aligned}
&\mathbb{E}\left[s^*_{r,t,h} s^{*'}_{r,s+h} \right] \\
=&\mathbb{E}\left[z_t z_s (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
=&\mathbb{E}\left[(\pi\varepsilon_{r,t} + \eta_t)(\pi\varepsilon_{r,s} + \eta_s)(u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
&\text{(by definition of } z_t \text{ in Assumption 4.2}^* \text{)} \\
=&\pi^2 \mathbb{E}\left[\varepsilon_{r,t}\varepsilon_{r,s} (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
&+ \mathbb{E}\left[\eta_t \eta_s (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
&\text{(since } \eta_t \text{ is exogenous to all } \varepsilon_\tau,\ \tau \geq 1 \text{ by Assumption 5.5, and } u^{\perp}_{t,h} \text{ contains only future} \\
&\text{shocks of period } t-1,\ \forall k \geq 0 \text{ by Assumption 5.4 due to the partialling out "}\perp\text{")} \\
=&\pi^2 \mathbb{E}\left[\varepsilon_{r,t}\varepsilon_{r,s} (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
&\text{(since } \eta_t \text{ is exogenous to all } \varepsilon_\tau \text{ and } \eta_s,\ \tau \geq 1, s < t \text{ by Assumption 5.5)} \\
=&\pi^2 \mathbb{E}\left[\mathbb{E}\left[\varepsilon_{r,t}\varepsilon_{r,s} (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \mid \varepsilon_{-r,t}, \{\varepsilon_\tau\}_{\tau \neq t} \right] \right] \\
&\text{(apply Law of Iterated Expectation)} \\
=&\pi^2 \mathbb{E}\left[\mathbb{E}\left[\varepsilon_{r,t} \mid \varepsilon_{-r,t}, \{\varepsilon_\tau\}_{\tau \neq t} \right] \varepsilon_{r,s} (u^{\perp}_{t,h}, u^{\perp}_{t-1,h}, \cdots, u^{\perp}_{t-h,h})'(u^{\perp}_{s,h}, u^{\perp}_{s-1,h}, \cdots, u^{\perp}_{s-h,h})' \right] \\
=&0 \\
&\text{(since } \mathbb{E}[\varepsilon_{r,t} \mid \varepsilon_{-r,t}, \{\varepsilon_\tau\}_{\tau \neq t}] = 0 \text{ by Assumption 5.5)}
\end{aligned} \tag{C.27}$$

where the second last equality needs to show that $u^{\perp}_{t-i,h}$ and $u^{\perp}_{s-i,h}$ for $i \in [0, H]$ are measurable by that information set. This the information set contains all shocks but $\varepsilon_{r,t}$. It is sufficient to show that $u^{\perp}_{t-i,h}$ and $u^{\perp}_{s-i,h}$ contains no $\varepsilon_{r,t}$. By the structure of $u^{\perp}_{t-i,h}$, it contains no $\varepsilon_{r,t-i}, \cdots, \varepsilon_{r,t-i+k}$, for $i \in [0, H]$. As $t - i \leq t \leq t - i + k$, $u^{\perp}_{t-i,h}$ contains no $\varepsilon_{r,t}$. Also, $u^{\perp}_{s-i,h}$ contains only shocks since period $s - i$ because of the partialling out and it contains no $\varepsilon_{r,s-i}, \cdots, \varepsilon_{r,s-i+k}$, for $i \in [0, H]$. Since $s > t$, then $s - i + k > t$. Thus, $u^{\perp}_{s-i,h}$ contains no $\varepsilon_{r,\tau}$ for all $\tau \leq s - i + k$. In summary, we show that $u^{\perp}_{t-i,h}$ and $u^{\perp}_{s-i,h}$ contains no $\varepsilon_{r,t}$, and in turn they are measurable by that information set. Thus, the serial uncorrelation of $s^*_{r,t,h}$ is proved. In turns, $\text{Var}(s^*_{r,t,h}) = \Omega_{r,h}$.

Similarly, we show its serial uncorrelation for process $s^*_{x,t,h}$. Since $s^*_{x,t,h} = \left[\varepsilon_{x,t} v^{\perp}_{t,h}, z_t (v^{\perp}_{t,h}, v^{\perp}_{t-1,h}, \cdots, v^{\perp}_{t-h,h})' \right]'$, it is sufficient to prove the serial uncorrelation by checking three covariances: $\mathbb{E}[\varepsilon_{x,t} v^{\perp}_{t,h} \varepsilon_{x,s} v^{\perp}_{s,h}] = 0$,
$\mathbb{E}[\varepsilon_{x,t} v^{\perp}_{t,h} z_s (v^{\perp}_{s,h}, v^{\perp}_{s-1,h}, \cdots, v^{\perp}_{s-h,h})] = 0$,
and $\mathbb{E}[z_t (v^{\perp}_{t,h}, v^{\perp}_{t-1,h}, \cdots, v^{\perp}_{t-h,h}) z_s (v^{\perp}_{s,h}, v^{\perp}_{s-1,h}, \cdots, v^{\perp}_{s-h,h})] = 0$.

Without loss of generosity, suppose $t < s$. The first covariance,

$$
\begin{aligned}
&\mathbb{E}\left[\varepsilon_{x,t} v_{t,h}^{\perp} \varepsilon_{x,s} v_{s,h}^{\perp}\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[\varepsilon_{x,t} v_{t,h}^{\perp} \varepsilon_{x,s} v_{s,h}^{\perp} \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right]\right] \\
&\text{(Law of Iterated Expectation)} \\
=&\mathbb{E}\left[\mathbb{E}\left[\varepsilon_{x,t} \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right] v_{t,h}^{\perp} \varepsilon_{x,s} v_{s,h}^{\perp}\right] \\
=&0 \\
&\text{(Assumption 5.5, } \mathbb{E}\left[\varepsilon_{x,t} \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right] = 0\text{)}
\end{aligned} \tag{C.28}
$$

The second last equality is because $v_{s,h}^{\perp}$ is measurable in the information set. The partialling out induces that it contains no shock before periods s. As $s > t$, $v_{s,h}^{\perp}$ contains no $\varepsilon_{x,t}$ and thereby it is measurable.

The second covariance,

$$
\begin{aligned}
&\mathbb{E}\left[\varepsilon_{x,t} v_{t,h}^{\perp} z_s(v_{s,h}^{\perp}, v_{s-1,h}^{\perp}, \cdots, v_{s-h,h}^{\perp})\right] \\
=&\pi\mathbb{E}\left[\varepsilon_{x,t} v_{t,h}^{\perp} \varepsilon_{r,s}(v_{s,h}^{\perp}, v_{s-1,h}^{\perp}, \cdots, v_{s-h,h}^{\perp})\right] \\
&\text{(since } \eta_t \text{ is exogenous to all structural shocks by Assumption 5.5)} \\
=&\pi\mathbb{E}\left[\mathbb{E}\left[\varepsilon_{x,t} v_{t,h}^{\perp} \varepsilon_{r,s}(v_{s,h}^{\perp}, v_{s-1,h}^{\perp}, \cdots, v_{s-h,h}^{\perp}) \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right]\right] \\
&\text{(Law of Iterated Expectation)} \\
=&\pi\mathbb{E}\left[\mathbb{E}\left[\varepsilon_{x,t} \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right] v_{t,h}^{\perp} \varepsilon_{r,s}(v_{s,h}^{\perp}, v_{s-1,h}^{\perp}, \cdots, v_{s-h,h}^{\perp})\right] \\
=&0 \\
&\text{(Assumption 5.5, } \mathbb{E}\left[\varepsilon_{x,t} \mid \varepsilon_{-x,t}, \{\varepsilon_{\tau}\}_{\tau\neq t}\right] = 0\text{)}
\end{aligned} \tag{C.29}
$$

The second last equality is because $v_{t,h}^{\perp}, (v_{s,h}^{\perp}, v_{s-1,h}^{\perp}, \cdots, v_{s-h,h}^{\perp})$ are measurable in the information set. The information set contains all shocks except $\varepsilon_{x,t}$. Then, it is sufficient to show all these components contain no $\varepsilon_{x,t}$. The term $v_{t,h}^{\perp}$ contains no $\varepsilon_{x,t}$ by its structure. The term $v_{s-i,h}^{\perp}$ for $i \in [0, H]$ contains no shock before periods $s + h - i$ due to the partialling out and Assumption 5.4. Since $s > t$ and $i \in [0, H]$, $v_{s-i,h}^{\perp}$ has no shock $\varepsilon_{x,t}$. Thus, it is measurable in the information set.

The third covariance equal to zero can be checked in the same manner as (C.27) since $v_{t,h}^{\perp} = u_{t,h}^{\perp} - \phi_h \varepsilon_{x,t}$. For any $u_{t,h}^{\perp}$. Thus, we have shown that all three covariances equal to zero and in turn $s_{x,t,h}^{*}$ is proved to be serially uncorrelated. Thus, $\mathrm{Var}(s_{x,t,h}^{*}) = \Omega_{x,h}$.

Next, we prove the sample variance of each estimate sequence has consistency property.

For process $\hat{\mathbf{s}}^*_{r,t,h} = z_t(\hat{u}^{\perp}_{t,h}, \hat{u}^{\perp}_{t-1,h}, \cdots, \hat{u}^{\perp}_{t-h,h})'$, where $\hat{u}^{\perp}_{t,h} = \hat{Y}^{\perp}_{t+h} - \hat{\vartheta}'_{yr,h}\hat{R}^{\perp}_{t:t+H}$. We are about to show $T^{-1}\sum_{t=1}^{T-h} \hat{\mathbf{s}}^*_{r,t,h}\hat{\mathbf{s}}^{*\prime}_{r,t,h} \xrightarrow{p} \Omega_{r,h}$. It is sufficient to show that for any $0 \leq i, j \leq k$, the convergence of $T^{-1}\sum_{t=1}^{T-h} z_t^2 \hat{u}^{\perp}_{t-i,h}\hat{u}^{\perp}_{t-j,h}$,

$$\begin{aligned}
&T^{-1}\sum_{t=1}^{T-h} z_t^2 \hat{u}^{\perp}_{t-i,h}\hat{u}^{\perp}_{t-j,h} \\
=&T^{-1}\sum_{t=1}^{T-h} z_t^2 u^{\perp}_{t-i,h}u^{\perp}_{t-j,h} + T^{-1}\sum_{t=1}^{T-h} z_t^2 u^{\perp}_{t-i,h}(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h}) \\
&+ T^{-1}\sum_{t=1}^{T-h} z_t^2 (\hat{u}^{\perp}_{t-i,h} - u^{\perp}_{t-i,h})u^{\perp}_{t-j,h} + T^{-1}\sum_{t=1}^{T-h} z_t^2 (\hat{u}^{\perp}_{t-i,h} - u^{\perp}_{t-i,h})(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h}) \\
\xrightarrow{p}&\mathbb{E}[z_t^2 u^{\perp}_{t-i,h}u^{\perp}_{t-j,h}]
\end{aligned} \tag{C.30}$$

The convergence holds if $T^{-1}\sum_{t=1}^{T-h} z_t^2 u^{\perp}_{t-i,h}u^{\perp}_{t-j,h} \xrightarrow{p} \mathbb{E}[z_t^2 u^{\perp}_{t-i,h}u^{\perp}_{t-j,h}]$ and the remaining three terms is an asymptotically negligible.

The convergence is proved by Assumption 5.1 with Corollary 3.48 (page 49) in White (1999). Each component $z_t, u^{\perp}_{t-i,h}, u^{\perp}_{t-j,h}$ is strong mixing process with size $-r/(r-2)$ for $r > 2$. The existence of the moment existence condition beyond the first moment for $z_t^2 u^{\perp}_{t-i,h}u^{\perp}_{t-j,h}$ can be checked with the Hölder's inequality and the boundedness of the fourth moment $(4+\delta)$ in Assumption 5.1, for some constant $\delta > 0$.

The asymptotic negligiblity can be proved through the Hölder's inequality.

$$\begin{aligned}
&|T^{-1}\sum_{t=1}^{T-h} z_t^2 u^{\perp}_{t-i,h}(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h})| \\
\leq&(T^{-1}\sum_{t=1}^{T-h} z_t^4)^{1/2}(T^{-1}\sum_{t=1}^{T-h}(u^{\perp}_{t-i,h})^4)^{1/4}(T^{-1}\sum_{t=1}^{T-h}(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h})^4)^{1/4} \\
\xrightarrow{p}&0
\end{aligned} \tag{C.31}$$

The term $T^{-1}\sum_{t=1}^{T-h} z_t^4, T^{-1}\sum_{t=1}^{T-h}(u^{\perp}_{t-i,h})^4$ can be shown to converge in probability by Assumption 5.1 with Corollary 3.48 (page 49) in White (1999). The convergence of the term $T^{-1}\sum_{t=1}^{T-h}(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h})^4$ can be shown by explicitly expanding $\hat{u}^{\perp}_{t-j,h} = u^{\perp}_{t-j,h} + \hat{Y}^{\perp}_{t+h-j} -$

$Y^{\perp}_{t+h-j} + \beta'_h(\hat{R}^{\perp}_{t:t+H} - R^{\perp}_{t:t+H}) + (\hat{\beta}'_h - \beta'_h)R^{\perp}_{t:t+H} + (\hat{\beta}'_h - \beta'_h)(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h})$ and

$$
\begin{aligned}
&T^{-1}\sum_{t=1}^{T-h}(\hat{u}^{\perp}_{t-j,h} - u^{\perp}_{t-j,h})^4 \\
=&T^{-1}\sum_{t=1}^{T-h}(\hat{Y}^{\perp}_{t+h-j} - Y^{\perp}_{t+h-j} + \beta'_h(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}) + (\hat{\beta}'_h - \beta'_h)R^{\perp}_{t:t+h} \\
&+ (\hat{\beta}'_h - \beta'_h)(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}))^4 \\
\leq&16\Bigg\{T^{-1}\sum_{t=1}^{T-h}(\hat{Y}^{\perp}_{t+h-j} - Y^{\perp}_{t+h-j})^4 + ||\beta_h||^4 T^{-1}\sum_{t=1}^{T-h}||\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}||^4 \\
&+ ||\hat{\beta}_h - \beta_h||^4 T^{-1}\sum_{t=1}^{T-h}||R^{\perp}_{t:t+h}||^4 \\
&+ ||\hat{\beta}_h - \beta_h||^4 T^{-1}\sum_{t=1}^{T-h}||\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}||^4\Bigg\} \\
&\xrightarrow{p} 0
\end{aligned}
\tag{C.32}
$$

The convergence relies on the conditions that $\sum_{t=1}^{T-h}(\hat{Y}^{\perp}_{t+h-j} - Y^{\perp}_{t+h-j})^4$, $\sum_{t=1}^{T-h}||\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}||^4$ are $O_p(1)$ in the limit by Assumption 5.2, the parameter β_h is constant, $||\hat{\beta}_h - \beta_h||^4$ converges to zero, and $T^{-1}\sum_{t=1}^{T-h}||R^{\perp}_{t:t+h}||^4$ converges to $\mathbb{E}[(R^{\perp}_{t:t+h})^4]$ due to the regularity conditions in Assumption 5.1 and Corollary 3.48 (page 49) in White (1999).

For process $\hat{s}^*_{x,t,h} = [\hat{\varepsilon}_{x,t}\hat{v}^{\perp}_{t,h}, z_t(\hat{v}^{\perp}_{t,h}, \hat{v}^{\perp}_{t-1,h}, \cdots, \hat{v}^{\perp}_{t-h,h})']'$, where $\hat{v}^{\perp}_{t,h} = \hat{Y}^{\perp}_{t+h} - \hat{\vartheta}'_{yr,h}\hat{R}^{\perp}_{t:t+h} - \hat{\phi}_{yr,h,H}\hat{\varepsilon}_{x,t}$. We are about to show $T^{-1}\sum_{t=1}^{T-h}\hat{s}^*_{x,t,h}\hat{s}^{*'}_{x,t,h} \xrightarrow{p} \Omega_{x,h}$. It is sufficient to show that the convergence,

1. $T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}\hat{v}^{\perp}_{t,h})^2 - T^{-1}\sum_{t=1}^{T-h}(\varepsilon_{x,t}v^{\perp}_{t,h})^2 \xrightarrow{p} 0$
2. $T^{-1}\sum_{t=1}^{T-h}\hat{\varepsilon}_{x,t}\hat{v}^{\perp}_{t,h}z_t\hat{v}^{\perp}_{t-i,h} - T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}v^{\perp}_{t,h}z_t v^{\perp}_{t-i,h} \xrightarrow{p} 0$
3. $T^{-1}\sum_{t=1}^{T-h}z_t\hat{v}^{\perp}_{t-i,h}z_t\hat{v}^{\perp}_{t-j,h} - T^{-1}\sum_{t=1}^{T-h}z_t v^{\perp}_{t-i,h}z_t v^{\perp}_{t-j,h} \xrightarrow{p} 0$

for any $0 \leq i, j \leq k$. It is because the convergence $T^{-1}\sum_{t=1}^{T-h}(\varepsilon_{x,t}v^{\perp}_{t,h})^2$, $T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}v^{\perp}_{t,h}z_t v^{\perp}_{t-i,h}$, and $T^{-1}\sum_{t=1}^{T-h}z_t v^{\perp}_{t-i,h}z_t v^{\perp}_{t-j,h}$ can be proved by Assumption 5.1 with Corollary 3.48 (page 49) in White (1999). Each component $z_t, v^{\perp}_{t-i,h}, \varepsilon_{x,t}$ is strong mixing process with size $-r/(r-2)$ for $r > 2$. The existence of the moment existence condition beyond the first moment for each process can be checked with the Hölder's inequality and the boundedness of the fourth moment $(4+\delta)$ in Assumption 5.1, for some constant $\delta > 0$.

To prove the three results of convergence to zero, it is sufficient to show $T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t} - \varepsilon_{x,t})^4$ and $T^{-1}\sum_{t=1}^{T-h}(\hat{v}^{\perp}_{t-i,h} - v^{\perp}_{t-i,h})^4$ converge to zero. We explicitly write $\hat{\varepsilon}_{x,t} - \varepsilon_{x,t} = X^{\perp}_t - \hat{X}^{\perp}_t$ by Assumption 4.3, and $\hat{v}^{\perp}_{t-j,h} = v^{\perp}_{t-j,h} + \hat{Y}^{\perp}_{t+h-j} - Y^{\perp}_{t+h-j} + \beta'_h(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h}) + (\hat{\beta}'_h - \beta'_h)R^{\perp}_{t:t+h} +$

$\hat{\phi}_h(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})+(\hat{\phi}_h-\phi_h)\varepsilon_{x,t}$. Then,

$$
\begin{aligned}
& T^{-1}\sum_{t=1}^{T-h}(\hat{v}^{\perp}_{t-j,h}-v^{\perp}_{t-j,h})^4 \\
=& T^{-1}\sum_{t=1}^{T-h}\Big(\hat{Y}^{\perp}_{t+h-j}-Y^{\perp}_{t+h-j}+\beta_h'(\hat{R}^{\perp}_{t:t+h}-R^{\perp}_{t:t+h})+(\hat{\beta}_h'-\beta_h')R^{\perp}_{t:t+h} \\
& +(\hat{\beta}_h-\beta_h)'(\hat{R}^{\perp}_{t:t+h}-R^{\perp}_{t:t+h})+\phi_h(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t}) \\
& +(\hat{\phi}_h-\phi_h)\varepsilon_{x,t}+(\hat{\phi}_h-\phi_h)(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\Big)^4 \\
\le& 49\Big\{T^{-1}\sum_{t=1}^{T-h}(\hat{Y}^{\perp}_{t+h-j}-Y^{\perp}_{t+h-j})^4+\|\beta_h\|^4T^{-1}\sum_{t=1}^{T-h}\|\hat{R}^{\perp}_{t:t+h}-R^{\perp}_{t:t+h}\|^4 \\
& +\|\hat{\beta}_h-\beta_h\|^4T^{-1}\sum_{t=1}^{T-h}\|R^{\perp}_{t:t+h}\|^4+\|\hat{\beta}_h-\beta_h\|^4T^{-1}\sum_{t=1}^{T-h}\|\hat{R}^{\perp}_{t:t+h}-R^{\perp}_{t:t+h}\|^4 \\
& +\phi_h^4T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^4+(\hat{\phi}_h-\phi_h)^4T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}^4 \\
& +(\hat{\phi}_h-\phi_h)^4T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^4\Big\} \\
\xrightarrow{p}& 0
\end{aligned}
\tag{C.33}
$$

The convergence relies on the conditions that $\sum_{t=1}^{T-h}(\hat{Y}^{\perp}_{t+h-j}-Y^{\perp}_{t+h-j})^4$, $\sum_{t=1}^{T-h}\|\hat{R}^{\perp}_{t:t+h}-R^{\perp}_{t:t+h}\|^4$, $\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^4=\sum_{t=1}^{T-h}(\hat{X}^{\perp}_t-X^{\perp}_t)^4$ are $O_p(1)$ in the limit by Assumption 5.2, the parameter β_h,ϕ_h are constant, $\|\hat{\beta}_h-\beta_h\|^4$ and $\|\hat{\phi}_h-\phi_h\|^4$ converge to zero, and $T^{-1}\sum_{t=1}^{T-h}\|R^{\perp}_{t:t+h}\|^4$ and $T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}^4$ respectively converge to $\mathbb{E}[(R^{\perp}_{t:t+h})^4]$ and $\mathbb{E}[\varepsilon_{x,t}^4]$ due to the regularity conditions in Assumption 5.1 and Corollary 3.48 (page 49) in White (1999). Thus, we have shown the estimation error of the covariance matrix estimate is asymptotically negligible.

In summary, the proof is complete.

C.7 Proofs of Lemmas

Lemma B.1

Proof. We uses Corollary 3.48 (page 49) and Theorem 5.20 (Wooldridge-White, page 130) in White (1999) to show the convergence in probability and convergence in distribution, respectively.

The convergence in probability will be proved by verifying the conditions of Corollary 3.48 in White (1999), (1) the process is strong mixing of size $-r/(1-r)$ for $r>1$, and (2) the $(r+\delta)$-th moment of the process is finite for some $\delta>0$.

By Assumption 5.1, the following terms $R^{\perp}_{t:t+H}, z_{t,H}, z_{x,t,H}, R^{\perp}_{x,t,H}$ are all strong mixing process of size $-r/(r-2)$ for $r>2$. It is due to Proposition 3.50 in White (1999), if two elements are strong mixing of size $-a$, the product of two are also strong mixing of size $-a$. It is readily to check $R^{\perp}_{t:t+H}=(R^{\perp}_{t},R^{\perp}_{t+1},\cdots,R^{\perp}_{t+H})'$, $z_{t,H}=(z_t,z_{t+1},\cdots,z_{t+H})'$ (by Assumption 4.2*), $z_t=\pi\varepsilon_{x,t}+\eta_t$ (Assumption 4.2*), $z_{x,t,H}=(\varepsilon_{x,t},z'_{t,H})'$, $R^{\perp}_{x,t,H}=(\varepsilon_{x,t},R^{\perp'}_{t:t+h})'$. Then, since $-r/(r-2)=-r/(r-1)+r/((r-2)(r-1))$, $-r/(r-2)<-r/(r-1)$ for $r>2$. It implies these terms are strong mixing of size $-r/(r-1)$, since the strong mixing process of size $-r/(r-2)$ has shorter memory than the strong mixing process of size $-r/(r-1)$.

Next, we check the moment of two processes, $z_{t,H}R'_{t:t+H}$, $z_{x,t,H}R'_{x,t,H}$. We apply Cauchy–Schwarz inequality that $\mathbb{E}\|XY'\|^r\le(\mathbb{E}\|X\|^{2r}\mathbb{E}\|Y\|^{2r})^{1/2}$. Then, we could write the bounded moment, $\mathbb{E}\|z_{t,H}R'_{t:t+H}\|^{r+\delta/2}\le(\mathbb{E}\|z_{t,H}\|^{2r+\delta}\mathbb{E}\|R_{t:t+H}\|^{2r+\delta})^{1/2}$. Given Assumption 5.1 (iii), The right-hand-size of the inequality is bounded above by some constant. Similarly for $z_{x,t,H}R'_{x,t,H}$.

Thus, two conditions are satisfied. Since η_t is uncorrelated with any structural shocks (Assumption 4.2* (ii)) and ε_t has constant unconditional covariance matrix assumed in (2.3), then $\mathbb{E}[z_{t,H}R'_{t:t+H}]$ is a constant, defined as $\mathbb{E}[z_{t,H}R'_{t:t+H}]:=\Sigma_{zr,h}$. Similarly, $\mathbb{E}[z_{x,t,H}R'_{x,t,H}]$ is a constant, defined as $\mathbb{E}[z_{x,t,H}R'_{x,t,H}]:=\Sigma_{zx,h}$. Thus, the convergence in probability into the expectation is proved.

The convergence in distribution will be proved by verifying the conditions of Theorem 5.20
(Wooldridge-White, page 130) in White (1999), (1) the process is strong mixing of size $-r/(r-2)$ for $r>2$, and (2) the (r)-th moment of the process is finite.

As illustrated, $z_{t,H}, z_{x,t,H}$ are all strong mixing process of size $-r/(r-2)$ for $r>2$. Since the term $u^{\perp}_{t,h}$ is a linear function of two strong mixing processes $Y^{\perp}_{t+h}$ and $R^{\perp}_{t:t+H}$, then $u^{\perp}_{t,h}$ is strong mixing process of size $-r/(r-2)$ for $r>2$. Similarly, $v^{\perp}_{t,h}$ is strong mixing process of size $-r/(r-2)$ for $r>2$.

The moment condition is verified in the similar manner. By Cauchy–Schwarz inequality, $\mathbb{E}\|z_{t,H}u^{\perp}_{t,h}\|^r\le(\mathbb{E}\|z_{t,H}\|^{2r}\mathbb{E}|u^{\perp}_{t,h}|^{2r})^{1/2}$. The boundedness of $\mathbb{E}\|z_{t,H}\|^{2r}$ has been verified in the proof of consistency; the boundedness of $\mathbb{E}|u^{\perp}_{t,h}|^{2r}$ is because $u^{\perp}_{t,h}$ is a linear function of $Y^{\perp}_{t+h}$ and $R^{\perp}_{t:t+H}$, and Assumption 5.1 (iii) imposes $\mathbb{E}\|W^{\perp}_{t+h}\|^{2r}$ is bounded by a constant.

In addition, by Assumption 4.2*, $\mathbb{E}[z_{t,H}u^{\perp}_{t,h}]=0$, and its long run variance is defined in (5.9). Thus, the convergence in distribution is proved. Similarly for the process $z_{x,t,H}v^{\perp}_{t,h}$. The proof is complete.

□

Lemma B.2

Proof. First we proof the consistency $T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h} \xrightarrow{p} \Sigma_{zr,h}$.

$$
\begin{aligned}
&T^{-1}\sum_{t=1}^{T-h} z_{t,H}\hat{R}^{\perp'}_{t:t+h} \\
=&T^{-1}\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h} + T^{-1}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h})' \\
=&T^{-1}\sum_{t=1}^{T-h} z_{t,H}R^{\perp'}_{t:t+h} + o_p(1) \\
\xrightarrow{p}&\Sigma_{zr,h} \\
&\text{(by Lemma B.1)}
\end{aligned}
\tag{C.34}
$$

The term $T^{-1}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h})'$ is negligible since

$$
\begin{aligned}
&\|T^{-1}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp}_{t:t+h} - R^{\perp}_{t:t+h})'\| \\
\leq&T^{-1/2}\sum_{i=0}^{h}\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp}_{t+i} - R^{\perp}_{t+i})'\| \\
\xrightarrow{p}&0 \\
&(\|T^{-1/2}\sum_{t=1}^{T-h} z_{t,H}(\hat{R}^{\perp}_{t+h} - R^{\perp}_{t+h})'\| \xrightarrow{p} 0, \text{ similarly to (C.21)})
\end{aligned}
\tag{C.35}
$$

Then, we proof the consistency $T^{-1}\sum_{t=1}^{T-h} \hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h} \xrightarrow{p} \Sigma_{zx,h}$.

$$
\begin{aligned}
&T^{-1}\sum_{t=1}^{T-h} \hat{z}_{x,t,h}\hat{R}^{\perp'}_{x,t,h} \\
=&T^{-1}\sum_{t=1}^{T-h} z_{x,t,h}R^{\perp'}_{x,t,H} + T^{-1}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h} - z_{x,t,h})R^{\perp'}_{x,t,H} \\
&+T^{-1}\sum_{t=1}^{T-h} z_{x,t,h}(\hat{R}^{\perp'}_{x,t,h} - R^{\perp'}_{x,t,H}) + T^{-1}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h} - z_{x,t,h})(\hat{R}^{\perp'}_{x,t,h} - R^{\perp'}_{x,t,H}) \\
=&T^{-1}\sum_{t=1}^{T-h} z_{x,t,h}R^{\perp'}_{x,t,H} + o_p(1) \\
\xrightarrow{p}&\Sigma_{zx,h} \\
&\text{(by Lemma B.1).}
\end{aligned}
\tag{C.36}
$$

The term $T^{-1}\sum_{t=1}^{T-h} z_{x,t,h}(\hat{R}^{\perp'}_{x,t,h} - R^{\perp'}_{x,t,H})$ and $T^{-1}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h} - z_{x,t,h})(\hat{R}^{\perp'}_{x,t,h} - R^{\perp'}_{x,t,H})$ are asymptotically negligible as argued in (C.24) and (C.25). The term $T^{-1}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h} - z_{x,t,h})R^{\perp'}_{x,t,H}$

is asymptotically negligible as

$$
\begin{aligned}
&\|T^{-1}\sum_{t=1}^{T-h}(\hat{z}_{x,t,h}-z_{x,t,h})R_{x,t,H}^{\perp'}\| \\
=&\|T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})R_{x,t,H}^{\perp'}\| \\
=&\|T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})\varepsilon_{x,t}\|+\sum_{i=0}^{h}\|T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})R_{t+i}^{\perp}\| \\
\leq&\|T^{-1}(\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2)^{1/2}(\sum_{t=1}^{T-h}\varepsilon_{x,t}^2)^{1/2}\|+\sum_{i=0}^{h}\|T^{-1}(\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2)^{1/2}(\sum_{t=1}^{T-h}(R_{t+i}^{\perp})^2)^{1/2}\| \\
\leq&\|(T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2)^{1/2}\|\|(T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}^2)^{1/2}\| \\
&+\sum_{i=0}^{h}\|(T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2)^{1/2}\|\|(T^{-1}\sum_{t=1}^{T-h}(R_{t+i}^{\perp})^2)^{1/2}\| \\
\xrightarrow{p}&0
\end{aligned}
\tag{C.37}
$$

The term $T^{-1}\sum_{t=1}^{T-h}(\hat{\varepsilon}_{x,t}-\varepsilon_{x,t})^2 \xrightarrow{p} 0$ by Assumption 5.2; the term $T^{-1}\sum_{t=1}^{T-h}\varepsilon_{x,t}^2$ converges to the variance of $\varepsilon_{x,t}$ and $T^{-1}\sum_{t=1}^{T-h}(R_{t+i}^{\perp})^2$ converges to the variance of $R_{t+i}^{\perp}$ by Corollary 3.48 (page 49) in White (1999) and Assumption 5.1. Then, the convergence is proved. Thus, the consistency of $T^{-1}\sum_{t=1}^{T-h}\hat{z}_{x,t,h}\hat{R}_{x,t,h}^{\perp'} \xrightarrow{p} \Sigma_{zx,h}$ is proved.

□