Measuring Dynamic Transmission using Pass-Through Impulse Response Functions

Giorgi Nikolaishvili*
Wake Forest University

(Access the Most Recent Version)
This Version: January 5, 2025
First Version: November 23, 2022

Abstract

I propose the pass-through impulse response function (PT-IRF) as a novel reduced-form empirical approach to measuring transmission channel dynamics. In essence, a PT-IRF quantifies the propagation of a shock through the Granger causality of a specified set of endogenous variables within a dynamical system. This approach has fewer informational requirements than alternative methods, such as structural parameter and empirical policy counterfactual exercises. A PT-IRF only requires the specification of a reduced-form VAR and identification of a shock of interest, bypassing the need to either build a structural model or identify multiple shocks. I demonstrate the flexibility of PT-IRFs by empirically analyzing the indirect dynamic transmission of oil price shocks to inflation and output via interest rates, as well as the indirect dynamic effect of monetary policy shocks on output via changes in credit supply.

JEL Classifications: C10; C32; C50; E52

Keywords: Directed graph; dynamic propagation; Granger causality; vector autoregression

*A short version of this paper was formerly circulated as "Pass-Through Impulse Response Functions (PT-IRFs)". I am grateful to Jeremy Piger, Aeimit Lakdawala, David Evans, George Evans, Bruce McGough, Jose Carreno, and participants at the 2024 Liberal Arts Macroeconomics Conference, Western Economic Association 99th Annual Conference, 2023 Meeting of the Midwest Econometrics Group, Fall 2023 Midwest Macroeconomics Meeting, University of Oregon Macro Group, and Wake Forest University Junior Faculty Seminar Series for many helpful comments and suggestions. All errors are my own. E-mail: nikolag@wfu.edu.

1 Introduction

The total effect of any macroeconomic disturbance is the net result of its dynamic propagation through the economy via a set of transmission channels. Gauging the extent to which a given channel amplifies, delays, or influences the direction of the effect over different horizons is of general interest to researchers and policymakers alike. I propose the pass-through impulse response function (PT-IRF) as a new empirical method enabling flexible quantification of transmission channels with relatively few informational requirements.

The PT-IRF leverages the insight that in any dynamical data-generating process (DGP) captured by a reduced-form vector autoregression (VAR), the dynamic propagation of a shock is governed by the Granger (G-) causality of the endogenous variables in the system. If a given endogenous variable does not G-cause the collective set of endogenous variables, it does not influence the dynamic propagation of disturbances. Consequently, the contribution of a set of endogenous variables to the total effect of a shock is the difference between the total effect and the impulse response observed when the joint G-causality of these "intermediate" variables is shut off. The PT-IRF quantifies this difference over a range of impact horizons as a smooth nonlinear mapping of VAR parameters, allowing for straightforward estimation and inference using both parametric VAR-based and semi-parametric local projections (LP)-based methods. This approach extends to nonlinear dynamical DGPs, provided that the G-causality of the relevant intermediate variables is well-defined. A key appeal of the PT-IRF is its methodological parsimony – unlike alternative approaches to quantifying dynamic transmission channels (further discussed below), it does not require specifying a structural model or identifying multiple distinct structural shocks. The PT-IRF can be applied to yield provisional insights regarding the strength and dynamics of a channel when alternative causal approaches are infeasible due to strict informational barriers, such as the inability to identify the required number of shocks and/or lack of knowledge on the structure of the relevant transmission mechanism.

I illustrate the applicability of the PT-IRF by studying the effect of an unexpected change in monetary policy on output via its transmission through the lending channel – a subset of the greater monetary transmission mechanism. The empirical PT-IRF approach involves the following steps: (1) One must first specify a reduced-form VAR containing all relevant endogenous variables to capture macroeconomic dynamics. The endogenous vector must include the necessary response variable, along with at least one transmission variable capturing time-series variation in bank credit supply (the transmission medium of interest). (2) A monetary policy shock can be identified using a valid internal, external, or combined identification scheme – the PT-IRF does not restrict the researcher to any specific shock identification method. (3) The model parameters can be estimated using either frequentist or Bayesian methods. (4) The slope and impact parameter estimates are mapped to a standard IRF, as well as a restricted IRF based on a transformed companion matrix in which the lag coefficients of the credit supply variable(s) are mapped to zero without altering other parameter estimates. The restricted impulse response captures the effect of a monetary policy shock with the influence of its dynamic propagation via credit supply partialed out, quantified by the IRF of the shock in a subsystem of the full model with strict G-non-causality of credit supply variable(s). Since the standard impulse response quantifies the total effect of the shock, the difference between the standard and restricted IRF captures the partial effect of a monetary shock on output via its influence on bank credit supply – this is the PT-IRF. (5) The resulting PT-IRF is a smooth mapping of VAR parameters, therefore common bootstrap procedures for IRF estimation can be applied without any additional modifications beyond mapping the estimated VAR parameters to the PT-IRF instead of an IRF at each step of the bootstrap. In a Bayesian setting, the same mapping can be performed with parameters drawn at each step of a sampler to generate credible sets.

An alternative approach to measuring the strength and nature of a transmission channel involves estimating a structural model and comparing the effects of a shock under two different parameter calibrations: one in which a channel of interest is present, and another in which the channel is shut off through a corresponding set of structural parameter restrictions. This structural methodology leverages the "Lucas program" approach to counterfactual analysis, as described in Christiano et al. (2005) and McKay and Wolf (2023). Although this methodology is highly flexible, an inherent downside is the required commitment to a particular structural model and corresponding parameterization. With regards to specification, the PT-IRF relies solely on the more parametrically-agnostic assumption that the dynamics in a given DGP can be accurately captured by a reduced-form VAR – a common approach in empirical macroeconomics (Stock and Watson, 2016; Ramey, 2016). I demonstrate an equivalence between the Lucas program and PT-IRF approaches to dynamic transmission quantification. This equivalence occurs when a set of restrictions on the primitive parameters of a structural model (related to a transmission mechanism of interest) can eliminate the G-causality of a corresponding intermediate endogenous variable in the VAR representation without affecting other reduced-form slope or contemporaneous impact parameters.[1] An empirical PT-IRF with such a corresponding structural mapping has a causal interpretation. Otherwise, PT-IRFs provide a mechanical description of shock propagation within a single equilibrium captured by an empirical reduced-form VAR.

Another approach to measuring dynamic transmission channel effects involves performing empirical counterfactual simulations inspired by Sims and Zha (2006). This method compares a standard impulse response to that of a counterfactual in which an intermediate variable is held constant by a sequence of corresponding exogenous shocks following the initial shock of interest. For example, Bernanke et al. (1997) quantifies the contribution of the systematic portion of monetary policy to the net dynamic effect of oil price shocks

[1] As an example, consider a model in which the lag coefficients of the intermediate endogenous variable(s) are all scaled by the same structural parameter representing the presence of some mechanism. Suppose this parameter does not function into any other reduced-form parameter in the model. A counterfactual in which the corresponding mechanism is not present involves setting the structural parameter equal to zero, which also eliminates the G-causality of the intermediate variable(s) from the model. The IRF to some shock in this counterfactual model is equivalent to the restricted IRF. Subtracting this counterfactual impulse response from the standard IRF yields a PT-IRF measuring the contribution of the mechanism to the total effect of the shock.

on key macroeconomic variables; Bachmann and Sims (2012) studies the role of consumer confidence in the dynamic transmission of fiscal policy shocks to the real economy. A recent refinement of this methodology in a policy counterfactual setting by McKay and Wolf (2023) shows the need to identify a potentially large set of news shocks in addition to a standard contemporaneous policy shock to ensure full robustness to the Lucas critique.[2] These strict requirements for establishing causality can be prohibitive in cases where it is difficult to identify multiple news shocks to the intermediate variable of interest in addition to a contemporaneous exogenous shock, in which case the traditional Sims-Zha approach loses its interpretability. By comparison, the PT-IRF does not necessitate the identification of any shocks to the intermediate variable(s) of interest – instead, it simply requires access to time-series data on (proxies of) those variables. The PT-IRF then leverages the estimates of the G-causality of the intermediate variable(s) to quantify the extent to which the corresponding channel of interest contributes to the propagation of some shock over a specified horizon.[3] In practice, the PT-IRF approach is descriptive of propagation dynamics within an equilibrium evidenced by a given sample, but can (1) be applied to a variety of intermediate variable beyond policy variables, (2) accommodate multiple intermediate variables simultaneously, and (3) accommodate potential nonlinearities, while being easily implementable.

Outline. Section 2 of the paper builds an intuitive interpretation of the parametric composition of a PT-IRF in the case of a single-lag linear VAR, extends this interpretation to a multi-lag linear VAR, and finally presents a general formulation of the PT-IRF for potentially multi-lag nonlinear VARs. Section 3 demonstrates a potential equivalence between the PT-IRF and Lucas program approaches to measuring transmission dynamics using a simplified DSGE model. Section 4 illustrates the empirical flexibility of PT-IRFs

[2] McKay and Wolf further demonstrate that the informational requirements of such counterfactual simulations increase in direct proportion to the length of the specified response horizon.

[3] In the case of measuring the contribution of the lending channel to total monetary transmission, it is notoriously challenging to identify exogenous shocks to credit supply. Therefore, the empirical policy counterfactual approach is infeasible. However, the PT-IRF can condition on a proxy for credit supply, such as the excess bond premium (EBP) (Gilchrist and Zakrajšek, 2012), to quantify the propagation of a monetary shock via the EBP as an endogenous variable in a VAR.

by estimating the transmission of various macroeconomic shocks via channels examined in the literature using linear VARs. Specifically, I study the transmission of oil price shocks to output and inflation via interest rate dynamics, and the transmission of monetary policy shocks to output via credit supply dynamics. Section 5 elaborates on methods of estimating and conducting inference on PT-IRFs using frequentist VAR- and LP-based bootstrap procedures. Section 6 concludes the paper with a brief discussion of potential applications of PT-IRFs beyond the illustrative examples.

2 Methodology

In this section I define the concept of a PT-IRF starting with the simple case of a linear VAR(1), proceeding to the more general case of a linear VAR(p), and finally generalizing to a stationary Markov process.

2.1 Linear VAR(1)

Consider the following VAR(1) process:

$$y_{t+1} = \alpha + A y_t + B \varepsilon_{t+1} \, , \tag{1}$$

where $y_t = (y_{1t}, \ldots, y_{Nt})'$ is a vector of N endogenous variables, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Kt})'$ is a vector of K structural shocks, α is an intercept vector, and $A \in \mathbb{R}^{N \times N}$ and $B \in \mathbb{R}^{N \times K}$ are the lag coefficient and contemporaneous impact matrices, respectively. Our goal is to interpret the linear VAR(1) as a directed weighted graph through which shock impulses travel over time, use this alternative interpretation of a linear VAR(1) to reinterpret the familiar IRF, and finally define the PT-IRF within the given context.

Firstly, notice that the ik-th entry of B represents the contemporaneous effect of the k-th structural shock on the i-th endogenous variable. Refer to Figure 1 for an illustration of the special case of a 3-dimensional VAR(1).

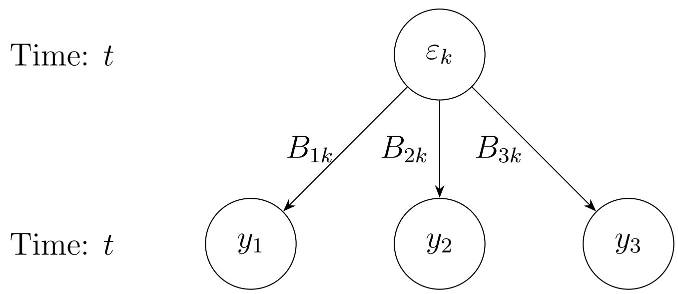

Figure 1: Contemporaneous effects of a structural shock ε_k on a 3-dimensional VAR as a weighted directed graph. **Note:** Notice that for each B_{ik}, i indexes the affected variable, while k indexes the shock of origin.

Next, notice that A_{ij} represents the one-period-ahead effect of the j-th variable on the i-th variable. If we think of A as the adjacency matrix in the context of a directed weighted graph, where each endogenous variable at a given point in time is a vertex, then A_{ij} may also be interpreted as the intensity of the travel path of a signal from variable j at time t to variable i at time $t + 1$. Once again, for an illustration of the above in the special case of a 3-dimensional VAR(1), refer to Figure 2.

Lastly, we can simply put the above interpretations of B and A together to formulate a VAR(1) as a directed weighted graph, which allows us to trace the propagation of a shock through the system and assess its impact on some variable of interest at a future point in time. Each possible path of a given shock ε_k has a corresponding weight equal to the product of the weights of each of its edges, determined by the contemporaneous impact and lag coefficient matrices. An IRF is simply the sum of the weights of all paths that ultimately reach a destination node corresponding to a variable of interest y_i at a given horizon h. Refer to Figure 3 for an illustration of the one-period-ahead propagation of a shock through a 3-dimensional VAR(1).

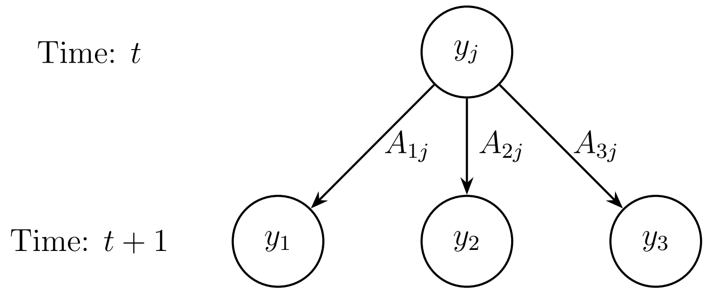

Figure 2: One-period-ahead effects of a change in the variable y_j of a 3-dimensional VAR as a weighted directed graph. **Note:** Notice that for each A_{ij}, i indexes the next period's destination variable, while j indexes the variable of origin.

A PT-IRF is the sum of weights associated with the subset of the above-mentioned paths that pass at least once through some intermediate variable of interest y_j – if a given path never passes through y_j, then it is irrelevant in gauging the role of y_j as a transmission medium for a shock in the system. For example, the one-period-ahead pass-through response of y_i with respect to y_1 as an intermediate variable for the transmission of some shock ε_k in the case illustrated by Figure 3 is equal to $A_{i1}B_{1k}$ – the weight of the only path that allows for the shock to pass through y_1 at least once before reaching its destination. If we were interested in the union of y_1 and y_2 as a transmission medium for ε_k, then the PT-IR would be $A_{i1}B_{1k} + A_{i2}B_{2k}$ – the sum of the weights of the two paths that allow the shock to pass through either y_1 or y_2 at least once before reaching its destination. The same logic can be extended to h-period-ahead impulse responses, for h strictly greater than zero.

It can be shown that in the case of a VAR(1), the h-period-ahead impulse response (IR) with respect to some vector of shocks $\overline{\varepsilon}$ may be expressed as

$$\mathrm{IR}(h, \overline{\varepsilon}) = A^h \, B \, \overline{\varepsilon} \,. \tag{2}$$

It can also be shown that for $h > 0$, the corresponding pass-through impulse response

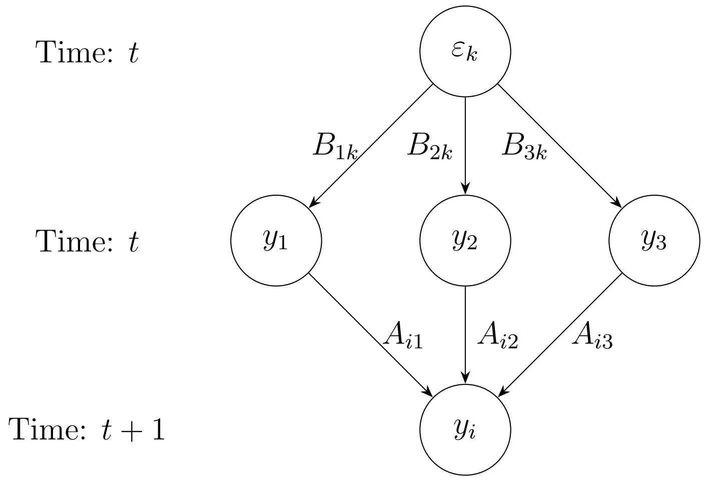

Figure 3: The propagation of an impulse originating at the k-th shock with the i-th variable as its destination, one period ahead. **Note:** The one-period-ahead impulse response of y_i with respect to a unit shock to ε_k equals the sum of the weights of all three paths leading to y_{it+1}: $A_{i1}B_{1k} + A_{i2}B_{2k} + A_{i3}B_{3k}$.

(PT-IR) with intermediate variable y_j is algebraically equivalent to

$$\text{PT-IR}(h, j, \bar{\varepsilon}) \equiv \left(A^h - \widetilde{A}^h_{-j}\right) B\,\bar{\varepsilon}\,, \tag{3}$$

where $\widetilde{A}_{-j}$ is identical to A across all but the j-th column, which is set equal to the zero vector. In the case that $h = 0$, the PT-IR always equals to zero due to the fact that contemporaneous pass-through of any given shock occurs purely through the impact matrix:

$$\text{PT-IR}(0, j, \bar{\varepsilon}) \equiv 0\,. \tag{4}$$

2.2 Linear VAR(p)

Consider the following VAR(p) process:

$$y_t = \alpha + A(L)y_t + B\varepsilon_t\,, \tag{5}$$

where all familiar objects are defined as before, and $A(L)$ is a lag polynomial of the form

$$A(L) = \sum_{i=1}^{p} A_i L^i \,, \tag{6}$$

such that each A_i is a lag coefficient matrix corresponding to y_{t-i}. Suppose we aim to derive PT-IR$(h, j, \overline{\varepsilon})$ for this system. The goal is once again to sum the weights associated only with those paths that originate at the shock of interest, pass through y_j at least once over the given horizon, and end at the response variable of interest h periods ahead.

Suppose we represent a linear VAR(p) in state-space form as a VAR(1) with companion matrix Φ and augmented contemporaneous impact matrix $\Gamma = \begin{bmatrix} B' & \mathbf{0} \end{bmatrix}'$. Then for $h \geq 0$ the corresponding PT-IR with intermediate variable y_j may be expressed as

$$\text{PT-IR}(h, j, \overline{\varepsilon}) \equiv \left(\Phi^h - \widetilde{\Phi}^h_{-j} \right) \Gamma \, \overline{\varepsilon} \,, \tag{7}$$

where $\widetilde{\Phi}_{-j}$ is the companion matrix of a modified version of the process described in Eq. (5) with the i-th lag coefficient matrix restricted to equaling

$$\widetilde{A}_{i,-j} \equiv \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_{j-1} & \vec{0} & \vec{a}_{j+1} & \dots & \vec{a}_N \end{bmatrix} \,, \tag{8}$$

where $\vec{a}_m$ denotes the m-th column of A_i. Notice that $\widetilde{\Phi}^h_{-j} \, \Gamma \, \overline{\varepsilon}$ captures the impulse response of a shock for a restricted version of the given linear VAR(p) in which the G-causality of the j-th endogenous variable is completely removed (Kilian and Lütkepohl, 2017) – all paths passing through the j-th variable are assigned a weight of zero. Therefore, PT-IR$(\cdot)$ sums the weights of only those paths that pass through the j-th variable, which can be interpreted as the impulse response of the system attributable to the G-causality of the j-th endogenous variable.

2.3 General Formulation

Let $y_t = (y_{1t}, \ldots, y_{Nt})' \in \mathbb{R}^N$ and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Kt})' \in \mathbb{R}^K$ such that y_t is determined by the stationary Markov process

$$y_t = G(\varepsilon_t, y_{t-1}; \theta) , \tag{9}$$

where $t \in \mathbb{N}^+$ denotes time, $G : \mathbb{R}^K \times \mathbb{R}^N \to \mathbb{R}^N$ is a mapping conditioned on a set of parameters θ, and ε_t is a vector of zero-mean i.i.d. shocks. We may define the h-step impulse response of the given system with respect to some shock vector $\bar{\varepsilon} \in \mathbb{R}^K$ as the following difference between two forecasts $\forall h \in \mathbb{N}$:

$$\mathrm{IR}(t, h, \bar{\varepsilon}) \equiv \mathbb{E}\left[y_{t+h} \,|\, \varepsilon_t = \bar{\varepsilon},\, y_{t-1},\, \theta\right] - \mathbb{E}\left[y_{t+h} \,|\, \varepsilon_t = 0,\, y_{t-1},\, \theta\right] , \tag{10}$$

where the conditional expectation operator $\mathbb{E}[\cdot \,|\, \cdot]$ represents the best mean squared error predictor.

The pass-through impulse response of the system with respect to the same shock is once again defined as

$$\text{PT-IR}(t, h, j, \bar{\varepsilon}) \equiv \mathrm{IR}(t, h, \bar{\varepsilon}) - \widetilde{\mathrm{IR}}(t, h, j, \bar{\varepsilon}) , \tag{11}$$

where $\widetilde{\mathrm{IR}}$ denotes an object similar to that in Eq. (10), but applied to a transformed version of the process expressed in Eq. (9) in which the G-causality of the j-th variable in the system is removed. More specifically, we may define

$$\widetilde{\mathrm{IR}}(t, h, j, \bar{\varepsilon}) \equiv \mathbb{E}\left[\widetilde{Y}_{t+h} \,|\, \varepsilon_t = \bar{\varepsilon},\, Y_{t-1},\, \theta\right] - \mathbb{E}\left[\widetilde{Y}_{t+h} \,|\, \varepsilon_t = 0,\, Y_{t-1},\, \theta\right] , \tag{12}$$

where $\widetilde{y}_t \equiv G(\varepsilon_t, \widetilde{I}_j y_{t-1}; \theta)$, such that $\widetilde{I}_j$ is the identity matrix with the j-th diagonal entry set equal to zero. In other words, $\widetilde{y}_t$ is generated by the same process as y_t, but with the influence of the lags of the j-th variable removed from the data generating process (DGP). Notice that $\widetilde{I}_j$ removes the G-causality of the j-th variable in the system, thus allowing for $\widetilde{\mathrm{IR}}(\cdot)$ to isolate the impulse response to a shock without accounting for its transmission through y_{jt}. Therefore, subtracting $\widetilde{\mathrm{IR}}(\cdot)$ from $\mathrm{IR}(\cdot)$ yields the impulse response associated with the transmission of a shock through y_{jt}.

3 Structural Equivalence to the Lucas Program

The structural interpretation of a PT-IRF depends entirely on the mapping between the primitive parameters of a model and the G-causality of the transmission medium of interest. $\widetilde{IR}(\cdot)$ can be obtained as a standard impulse response after restricting the structural parameters such that the G-causality of the intermediate variable(s) is eliminated from the model dynamics. The PT-IRF can then be obtained by subtracting this counterfactual impulse response from the true impulse response to a given shock in the full model. In some cases, the set of restrictions necessary to eliminate the G-causality of transmission media may be easily interpretable. I demonstrate one such case using a small-scale dynamic stochastic general equilibrium (DSGE) model, in which the dynamic responses of all endogenous variables with respect to demand, cost-push, and monetary policy shocks are determined by the stickiness of interest rates in the Taylor rule.

An and Schorfheide (2007) specify a canonical New Keynesian DSGE model that abstracts away from wage rigidities and capital accumulation. The model contains a representative household, a single final goods-producing firm, a continuum of intermediate goods producing firms, as well as monetary and fiscal authorities. The equilibrium of their model – linearized at the steady state – consists of the following six equations:

$$y_t = E_t y_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau}\left(r_t - E_t \pi_{t+1} - E_t z_{t+1}\right) \tag{13}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(y_t - g_t) \tag{14}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left(\psi_\pi \pi_t + \psi_y (y_t - g_t)\right) + \varepsilon_{rt} \tag{15}$$

$$c_t = y_t - g_t \tag{16}$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{zt} \tag{17}$$

$$g_t = \rho_g g_{t-1} + \varepsilon_{gt}\,. \tag{18}$$

These are the Euler equation, Phillips curve, Taylor rule, aggregate accounting equality, law of motion for TFP, and law of motion for government spending, respectively. Here y_t, π_t, r_t, c_t, z_t, and g_t denote output, inflation, interest rate, consumption, TFP, and government spending, respectively. All but one coefficient is a model primitive – κ is determined by the following mapping:

$$\kappa = \frac{\tau(1-\nu)}{\nu\phi\overline{\pi}^2}\,, \tag{19}$$

where τ is the constant relative risk aversion (CRRA) parameter, ν is the inverse elasticity of demand, ϕ is the level of price stickiness, and $\overline{\pi}$ is the level of steady-state inflation. The vector of innovations $\varepsilon_t = (\varepsilon_{zt}, \varepsilon_{gt}, \varepsilon_{zt})'$ is mean-zero *iid*.

I simplify the An-Schorfheide model and solve it analytically to yield a reduced-form VAR similar to Morris (2014) . Consider the case in which TFP is *iid*, $z_t = \varepsilon_{zt}$, and government spending is zero at all times. In this restricted case, the model has the following three equations:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left(\psi_\pi \pi_t + \psi_y (y_t - g_t)\right) + \varepsilon_{rt} \tag{20}$$

$$y_t = E_t y_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau}\left(r_t - E_t \pi_{t+1} - \varepsilon_{zt}\right) \tag{21}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(y_t - g_t)\,. \tag{22}$$

The solution of this model suggests the following functional form for the interest rate equation:

$$r_t = \phi_{rr} r_{t-1} + d_{rz}\varepsilon_{zt} + d_{rg}\varepsilon_{gt} + d_{rr}\varepsilon_{rt}\,. \tag{23}$$

We may define $\rho_r = \phi_{rr}$ and let $\varepsilon_{rt} = d_{rz}\varepsilon_{zt} + d_{rg}\varepsilon_{gt} + d_{rr}\varepsilon_{rt}$. Adding a cost-push shock, $\varepsilon_{\pi t}$, to the Phillips curve and defining $\varepsilon_{yt} = (1/\tau)\varepsilon_{zt}$ allows for the following expression for the TE of the three-equation model:

$$r_t = \rho_r r_{t-1} + \varepsilon_{rt} \tag{24}$$

$$y_t = E_t y_{t+1} - \frac{1}{\tau}\left(r_t - E_t \pi_{t+1}\right) + \varepsilon_{yt} \tag{25}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t + \varepsilon_{\pi t}\,. \tag{26}$$

By the method of undetermined coefficients, assuming $E_t y_{t+1} = \phi_{yr} r_t$ and $E_t \pi_{t+1} = \phi_{\pi r} r_t$, we obtain the following VAR(1) representation:

$$\underbrace{\begin{bmatrix} r_t \\ y_t \\ \pi_t \end{bmatrix}}_{Y_t} = \underbrace{\begin{bmatrix} \rho_r & 0 & 0 \\ \phi_{yr} & 0 & 0 \\ \phi_{\pi r} & 0 & 0 \end{bmatrix}}_{\Phi} \begin{bmatrix} r_{t-1} \\ y_{t-1} \\ \pi_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ \frac{\phi_{yr}}{\rho_r} & 1 & 0 \\ \frac{\phi_{\pi r}}{\rho_r} & \kappa & 1 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} \varepsilon_{rt} \\ \varepsilon_{yt} \\ \varepsilon_{\pi t} \end{bmatrix}}_{U_t}, \tag{27}$$

where

$$\phi_{yr} = \left(\frac{1 - \rho_r \beta}{\kappa}\right) \phi_{\pi r} \tag{28}$$

$$\phi_{\pi r} = -\left(\frac{1}{\tau} \cdot \frac{\rho_r}{1 - \rho_r}\right) \left(\frac{1 - \rho_r \beta}{\kappa} - \left[\frac{1}{\tau} \cdot \frac{\rho_r}{1 - \rho_r}\right]\right)^{-1}. \tag{29}$$

Notice that y and π do not G-cause the process in Eq. (27). Therefore, PT-IRFs that condition on these two endogenous variables as the transmission medium (intermediate variables) must equal to zero at all horizons. r is the only variable allowing for the dynamic propagation of structural shocks via its G-causality – PT-IRFs that condition on r as the intermediate variable equal their corresponding IRFs beyond the initial period. The G-causality of r is determined by the structural parameter ρ_r and reduced-form parameters ϕ_{yr} and $\phi_{\pi r}$, both of which are scaled by ρ_r. If the interest rate stickiness in Eq. (15) is removed by setting $\rho_r = 0$, then the G-causality of r is also shut off and all dynamic effects in the model disappear. So, the PT-IRF of an endogenous variable in response to a shock with r as the intermediate variable can be obtained by computing the response of the system under the counterfactual $\rho_r = 0$, and subtracting it from the true IRF over a set horizon. Therefore, the dynamic transmission of shocks via the interest rate occurs due to the stickiness of the Taylor rule.[4]

To demonstrate the relationship between the theoretical model and empirics, I calibrate the simplified model – parameter values are presented in Table 1. I compare the empirical and theoretical PT-IRFs of all three endogenous variables in response to an innovation in government spending (which affects the system through the monetary policy shock ε_r), with the interest rate as the intermediate variable. The theoretical PT-IRFs are generated using the above-mentioned counterfactual exercise involving ρ_r, whereas the empirical PT-IRF confidence intervals are obtained by estimating a VAR(2) on 1,000 observations simulated using the structural model.[5] Figure 4 presents the results of this exercise.

The empirical PT-IRF confidence intervals presented in Figure 4 are equivalent to a Lucas program-style structural counterfactual exercise that compares the responses to a

[4]One caveat in this example is that since ρ_r also scales the contemporaneous effects of monetary policy shocks, the instantaneous effect of a monetary shock in the counterfactual model will not match that of the true DGP. Notice that the contemporaneous effect of ε_r on π and y equals to $-\kappa/\tau$ and $-1/\tau$, respectively, given $\rho_r = 0$.

[5]I generate the empirical PT-IRFs using a VAR(2) instead of a VAR(1) to emulate minor model misspecification.

shock under two alternative structural parameterizations.[6] Comparatively, the PT-IRF confidence intervals are generated using transformations of reduced-form VAR parameter estimates. These results are also equivalent to a policy counterfactual exercise in the style of Sims and Zha (2006) that compares the effects of a government spending shock under two alternative monetary policy rules. Unlike the Sims-Zha approach, the PT-IRF only leverages the identification of the government spending shock and nothing else.

Parameter	Description	Value
β	Discount factor	0.9975
ρ_r	Taylor rule stickiness	0.75
τ	CRRA	2
κ	Composite parameter	0.33

Table 1: Simplified An-Schorfheide model parameter calibration following that of Morris (2014).

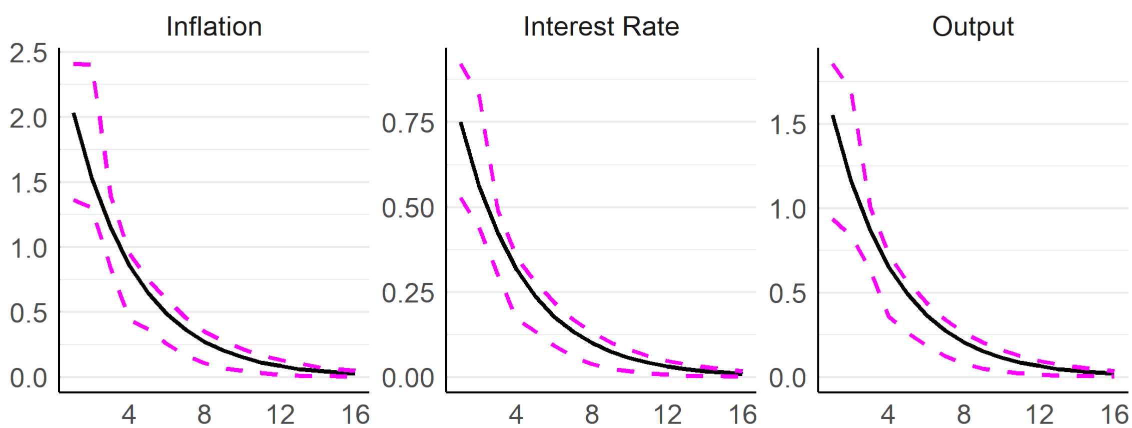

Figure 4: Structural vs. empirical PT-IRF confidence intervals of all endogenous variables in the simplified An-Schorfheide model in response to a government spending shock, with the interest rate as the intermediate variable. **Note:** The solid black line is the theoretical PT-IRF, produced using the $\rho_r = 0$ counterfactual exercise described earlier in Section 3. The dashed magenta lines are the bootstrapped 95% confidence intervals for matching PT-IRFs obtained using a VAR(2) estimated on a sample of 1,000 observations generated using the structural model.

[6]The bootstrap algorithm used to generate these PT-IRF confidence intervals is presented in Section 5.

4 Illustrative Empirical Applications

I demonstrate the empirical flexibility of the PT-IRF by applying it study oil price shock transmission to inflation and output via monetary policy, as well as the transmission of monetary policy shocks to output via credit supply.

4.1 Oil Price Shock Transmission to Output via Monetary Policy

Bernanke et al. (1997) use the Sims-Zha empirical counterfactual methodology to study the contribution of the systematic portion of monetary policy to the effect of oil price shocks on business cycles. Bernanke et al. construct a sequence of monetary policy shocks that keeps the federal funds rate unchanged in the face of an oil price shock – intuitively, this is equivalent to shutting down the systematic monetary policy response. They then compare the dynamic responses of key macroeconomic variables under scenarios with and without systematic monetary responses to conclude how much it contributes to the propagation of oil price shocks. They find that a significant portion of the economic downturn following a positive oil price shock can be attributed to systemic monetary policy responses, rather than the direct impact of the oil price shock. In other words, Bernanke et al. find that monetary policy amplifies the contractionary effects of an oil price shock.

Kilian and Lewis (2011) carry out a similar analysis using a modified counterfactual framework – they propose constructing a sequence of monetary policy shocks that offset the contemporaneous and lagged effects of including the real price of oil in the policy reaction function. Kilian and Lewis argue that this alternative exercise yields a more accurate picture of the contributions of systematic monetary policy to oil price shock propagation. Their results challenge the notion that systematic monetary policy responses to oil price shocks have been a major source of aggregate fluctuations in the U.S. economy.

I replicate the VAR estimated by Kilian and Lewis (2011) to study the transmission of oil price shocks to output and inflation via monetary policy using PT-IRFs. The model consists of the following five endogenous variables, constructed in precisely the same way as in Kilian and Lewis (2011): (1) the percentage change in the real price of imported commodities; (2) the percentage change in the real price of imported crude oil; (3) the Chicago Fed National Activity Index (CFNAI) measure of US real activity; (4) the U.S. CPI inflation rate; and (5) the federal funds rate (FFR). The sample runs from May 1967 until July 1987. All data are obtained from the replication package for Chen (2023), which successfully reproduces the results of Kilian and Lewis (2011). The oil price shock in this VAR is recursively-identified as the shock corresponding to the real price of imported crude oil variable (the second shock yielded by the Cholesky decomposition of the error covariance matrix). This identification scheme exploits the conventional assumption that oil prices are predetermined with respect to domestic macroeconomic aggregates.

Figure 5 shows the net dynamic response of output and inflation (black lines), matching those of Kilian and Lewis (2011). The same Figure also shows the PT-IRFs representing the response of output and inflation to oil price shocks via their propagation through the FFR (dashed lines) – these capture the contributions of monetary policy to the dynamic transmission of oil price shocks to business cycles. Although the majority of confidence intervals contain zero, the signs of point estimates of both PT-IRFs shown in Figure 5 generally match the directions of their respective IRFs at most horizons. This result loosely suggests that monetary policy amplifies the effects of oil price shocks on the macroeconomy, as argued by Bernanke et al. (1997). However, it is once again worth noting that PT-IRFs do not necessarily capture the same notion of "dynamic transmission" as does the Sims and Zha (2006) counterfactual method, though there can exist an equivalence in special cases such as the one discussed in Section 3.

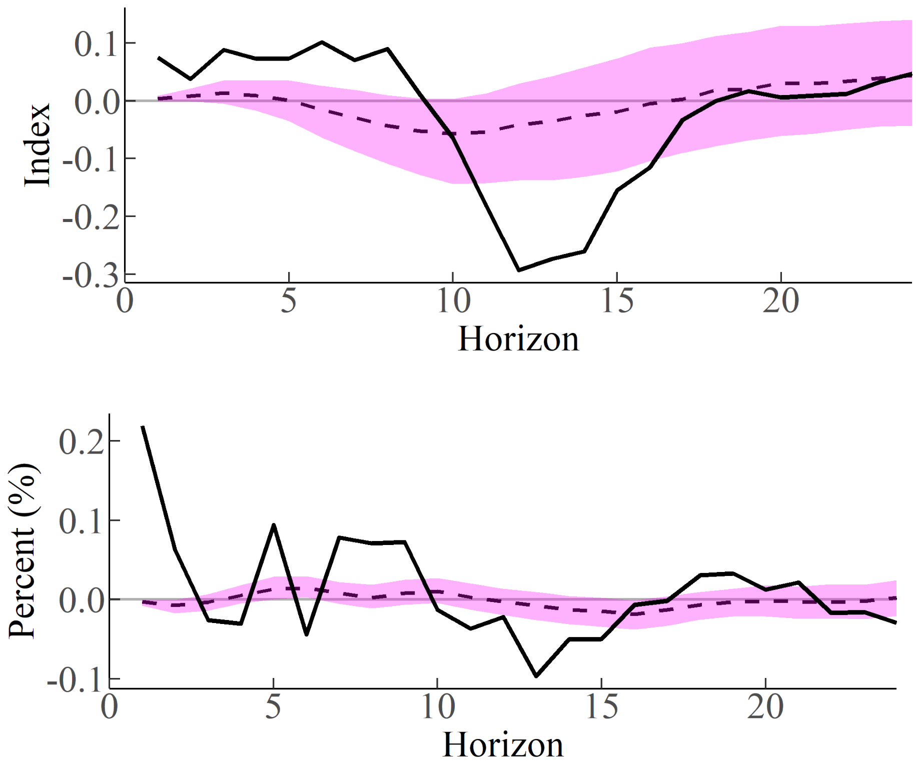

Figure 5: Responses of output and inflation to a positive real oil price shock, respectively. **Note:** The solid black lines represent IRs, while the dashed lines are the PT-IRs of the two response variables with the federal funds rate as the intermediate variable. The magenta bands are bootstrapped 68% PT-IRF confidence using 1,000 runs.

4.2 Monetary Transmission to Output via Credit Supply

The credit channel of monetary transmission refers to the mechanism through which changes in monetary policy affect the macroeconomy by influencing the supply and demand for credit. The credit channel consists of the bank lending and balance sheet (sub-)channels (Bernanke and Gertler, 1995). The bank lending channel (BLC) describes the influence of monetary policy on banks' ability to supply loans, which affects macroeconomic activity beyond the direct effects of monetary policy on interest rates (Bernanke and Blinder, 1988). The balance sheet channel (BSC) describes changes in credit demand brought on by changes in creditworthiness and borrowing costs due to monetary policy actions. Like the BLC, the BSC produces indirect effects on macroeconomic activity beyond the direct interest rate channel.

In this illustrative application, I study the credit channel by applying PT-IRFs to the VAR used by Bu et al. (2021). Gilchrist and Zakrajšek (2012) develop the excess bond premium (EBP) index, an increase in which is associated with a contraction in the supply of credit. Bu et al. (2021) include the EBP as an endogenous variable in a monthly VAR, which they use to examine the dynamic effects of (externally-identified) monetary policy shocks on key macroeconomic variables. Their justification for including the EBP in the VAR focused largely on its strong predictive power for future macroeconomic activity. I replicate the VAR and monetary policy shock identification scheme used by Bu et al. (2021), and estimate the transmission of unanticipated changes in monetary policy to output via changes in credit supply using the EBP as the proxy intermediate variable.[7]

The model consists of the following five endogenous variables: (1) the cumulated Bu-Rogers-Wu (BRW) externally-identified monetary policy shock series; (2) log-transformed industrial production; (3) log-transformed consumer price index for all items in the U.S; (4) log-transformed producer price index for all commodities; and (5) the EBP. The sample runs from January 1994 until December 2019. The monetary policy shock is identified recursively as having a contemporaneous effect on all endogenous variables in the system – in other words, it is the innovation to the BRW series (ordered first in the system).

The resulting PT-IRF point estimates and confidence intervals are presented in Figure 6. Notice that the PT-IRF point estimates have the same sign as the IRF of a monetary tightening on output at all but the last couple periods over a 30-month horizon. Furthermore, notice that some of the point estimates are statistically significant at the 68% level when tested separately. Therefore, if the EBP is indeed a good proxy of variation in credit supply, then these findings show aggregate evidence of monetary transmission through the credit

[7] Nikolaishvili (2024) applies PT-IRFs to study the transmission of monetary policy shocks to output via community versus noncommunity bank lending in a factor-augmented VAR with hierarchical bank lending factors.

channel amplifying the effect of monetary policy on economic activity. This finding is in line with the theoretical foundation for the supply-side credit channel, which hypothesizes that micro-level frictions in lenders' ability to issue credit amplify the effects of monetary shocks.[8]

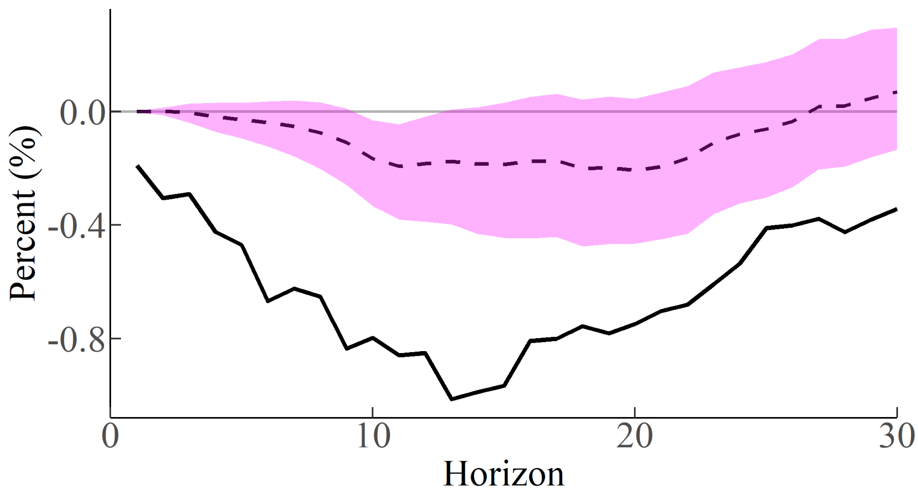

Figure 6: Response of industrial production to an unexpected monetary tightening. **Note:** The solid black line is the IR, the dashed line is a PT-IR representing the transmission of the monetary shock via the EBP (proxy for credit supply), and the magenta bands are bootstrapped 68% PT-IRF confidence using 1,000 runs.

5 Estimation and Inference

PT-IRF point estimates, much like those of a VAR-based IRF, can be obtained by estimating a VAR and mapping its parameters to the PT-IRF.[9] A model that explicitly approximates the dynamic structure and G-causality of the transmission media, such as a VAR, is necessary to estimate PT-IRFs. Non-parametric or semi-parametric alternatives like local projections (LPs) cannot be used to generate point estimates, unless realizations of the counterfactual DGP in which the G-causality of the transmission media is shut off can be either directly observed or perfectly inferred using available data. On the other hand,

[8]See Bernanke et al. (1996) for a seminal treatment of the topic of macrofinancial amplification.

[9]This approach applies to nonlinear VAR PT-IRFs as well. For example, in the case of time-varying parameter (TVP) VARs, PT-IRF mappings can be done at different sample periods of interest.

confidence intervals for a PT-IRF may be obtained using both parametric methods like VARs and semi-parametric ones like LPs. Frequentist VAR-based PT-IRF inference can be done via bootstrapping, such that an accompanying PT-IRF is generated using the synthetically-estimated VAR parameters at each step of the bootstrap.[10] Refer to Algorithm 1 for a standard VAR-based recursive-design PT-IRF bootstrapping procedure based on Runkle (1987). Bias adjustment and wild bootstrap modifications for the given procedure to account for possible biases and heteroskedasticity in the errors are straight-forward.[11]

Algorithm 1 Bootstrap Procedure for PT-IRF VAR-based Confidence Intervals

Estimate a VAR and save its parameters and residuals.
Set the initial conditions using the first p observations, where p is the number of lags.
for each bootstrap iteration **do**
  Generate new series by bootstrapping:
  for each replication **do**
    Select a residual at random (from the original series of residuals).
    Use the selected residual and past observations to generate a new observation.
    Repeat to form a complete bootstrapped series.
  end for
  Re-estimate the VAR parameters using the bootstrapped series.
  Compute $\mathrm{IR} - \widetilde{\mathrm{IR}}$ using the bootstrapped VAR parameters and store the result.
end for
Compute the empirical confidence intervals from a sufficiently large sample of stored PT-IRFs.

LP-based PT-IRF confidence intervals may also be bootstrapped with the use of synthetic samples simulated using an empirical VAR. At each iteration of the bootstrap, LPs are used to estimate the net response, as well as the response of the counterfactual DGP in which the G-causality of the transmission media is shut off. Refer to Algorithm 2 for a standard LP-based recursive-design PT-IRF bootstrapping procedure that combines elements of Runkle (1987) and the lag-augmented LP framework of Montiel Olea and Plagborg-Møller (2021). Once again, various modifications to account for bias and deviations from the identical and

[10]Similarly, Bayesian VAR-based PT-IRF posterior distributions can be obtained by mapping the VAR parameter estimates to the PT-IRF at each step of the sampler. This is possible since PT-IRFs are mappings of VAR parameters.

[11]For instance, popular bias-correcting and heteroskedasticity-robust bootstrap algorithms developed in Kilian (1998) and Gonçalves and Kilian (2004), respectively, can easily accommodate the PT-IRF framework.

independent distribution of the errors are easy to make.

Algorithm 2 Bootstrap Procedure for PT-IRF LP-based Confidence Intervals

Estimate a VAR and save its parameters and residuals.
Set the initial conditions using the first p observations, where p is the number of lags.
for each bootstrap iteration **do**
    Generate new series by bootstrapping:
    for each replication **do**
        Select a residual at random (from the original series of residuals).
        Generate a new obs. for both Y and $\widetilde{Y}$ using the selected residual and past obs.
        Repeat to form complete bootstrapped series of Y and $\widetilde{Y}$.
    end for
    Estimate IR using Y and $\widetilde{\mathrm{IR}}$ using $\widetilde{Y}$.
    Compute $\mathrm{IR} - \widetilde{\mathrm{IR}}$ and store the result.
end for
Compute the empirical confidence intervals from a sufficiently large sample of stored PT-IRFs.

I test the performance of the VAR estimator in a Monte Carlo simulation study. I use the VAR estimated in Section 4.1 as the population DGP with which I simulate a large number of samples with a realistic range of samples sizes. For each sample, I estimate and generate 95% confidence intervals for the PT-IRF measuring the effect of an oil price shock on output via the FFR (see top panel of Figure 5). For each sample size, I check the performance of the estimator by computing the average bias, variance, and mean squared error (MSE) at each horizon using correctly-specified and misspecified VAR models. The results of this exercise are presented in Figure A.1. The same Figure also shows the coverage rate of the bootstrapped confidence intervals for each combination of sample size and model specification. The coverage of the estimator seems to deteriorate for longer horizons relative to the sample size for under-specified VARs, which is not the case for correctly-specified and over-specified models. Overall, according to the evidence yielded by this simulation study, the VAR estimator appears consistent and the empirical coverage rate matches the nominal coverage probability when the VAR model is not under-specified.

6 Concluding Remarks

The PT-IRF can be a useful tool for quantifying transmission channels of a wide variety of economic shocks. Although the applications in this paper involve linear VARs, PT-IRFs can be applied to arbitrarily nonlinear settings, such as TVP-VARs. Furthermore, the ability of PT-IRFs to accommodate multiple intermediate variables allows for the simultaneous quantification of potentially overlapping contributions of multiple transmission channels.[12] Future work can examine the properties of VAR vs. LP-based methods of conducting inference on PT-IRF estimates. For example, one may be more efficient than the other at certain horizons and/or given certain levels of persistence. Finally, PT-IRFs can be used to study dynamic transmission mechanisms both empirically and theoretically in fields outside of macroeconomics and macrofinance that rely on dynamic modeling.

[12] Nikolaishvili (2024) studies the contribution of comovement among community versus noncommunity banks' lending to the transmission of monetary policy shocks. The PT-IRFs in this study simultaneously condition on multiple lending factors for each bank type as intermediate variables / transmission media.

References

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews*, 26(2-4):113–172.

Bachmann, R. and Sims, E. (2012). Confidence and the transmission of government spending shocks. *Journal of Monetary Economics*, 59(3):235–249.

Bernanke, B., Gertler, M., and Gilchrist, S. (1996). The financial accelerator and the flight to quality. *The Review of Economics and Statistics*, 78(1):1–15.

Bernanke, B., Gertler, M., and Watson, M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, 28(1):91–157.

Bernanke, B. S. and Blinder, A. S. (1988). Credit, money, and aggregate demand. *The American Economic Review*, 78(2):435–439.

Bernanke, B. S. and Gertler, M. (1995). Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives*, 9(4):27–48.

Bu, C., Rogers, J., and Wu, W. (2021). A unified measure of fed monetary policy shocks. *Journal of Monetary Economics*, 118:331–349.

Chen, S.-S. (2023). A direct approach to kilian–lewis style counterfactual analysis in vector autoregression models. *Journal of Applied Econometrics*, 38(7):1068–1076.

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

Gilchrist, S. and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, 102(4):1692–1720.

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

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *The Review of Economics and Statistics*, 80(2):218–230.

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

Kilian, L. and Lütkepohl, H. (2017). *Vector Autoregressive Models*, page 19–74. Themes in Modern Econometrics. Cambridge University Press.

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

McKay, A. and Wolf, C. (2023). What Can Time-Series Regressions Tell Us About Policy Counterfactuals? Staff Report 642, Federal Reserve Bank of Minneapolis.

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

Morris, S. (2014). The Statistical Implications of Common Identifying Restrictions for DSGE Models. 2014 Meeting Papers 738, Society for Economic Dynamics.

Nikolaishvili, G. (2024). Commercial Bank Heterogeneity and the Transmission of Monetary Policy Through Bank Lending. Working Paper.

Ramey, V. (2016). Chapter 2 - macroeconomic shocks and their propagation. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 71–162. Elsevier.

Runkle, D. E. (1987). Vector autoregressions and reality. *Journal of Business & Economic Statistics*, 5(4):437–442.

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

Stock, J. and Watson, M. (2016). Chapter 8 - dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 415–525. Elsevier.

A Appendix

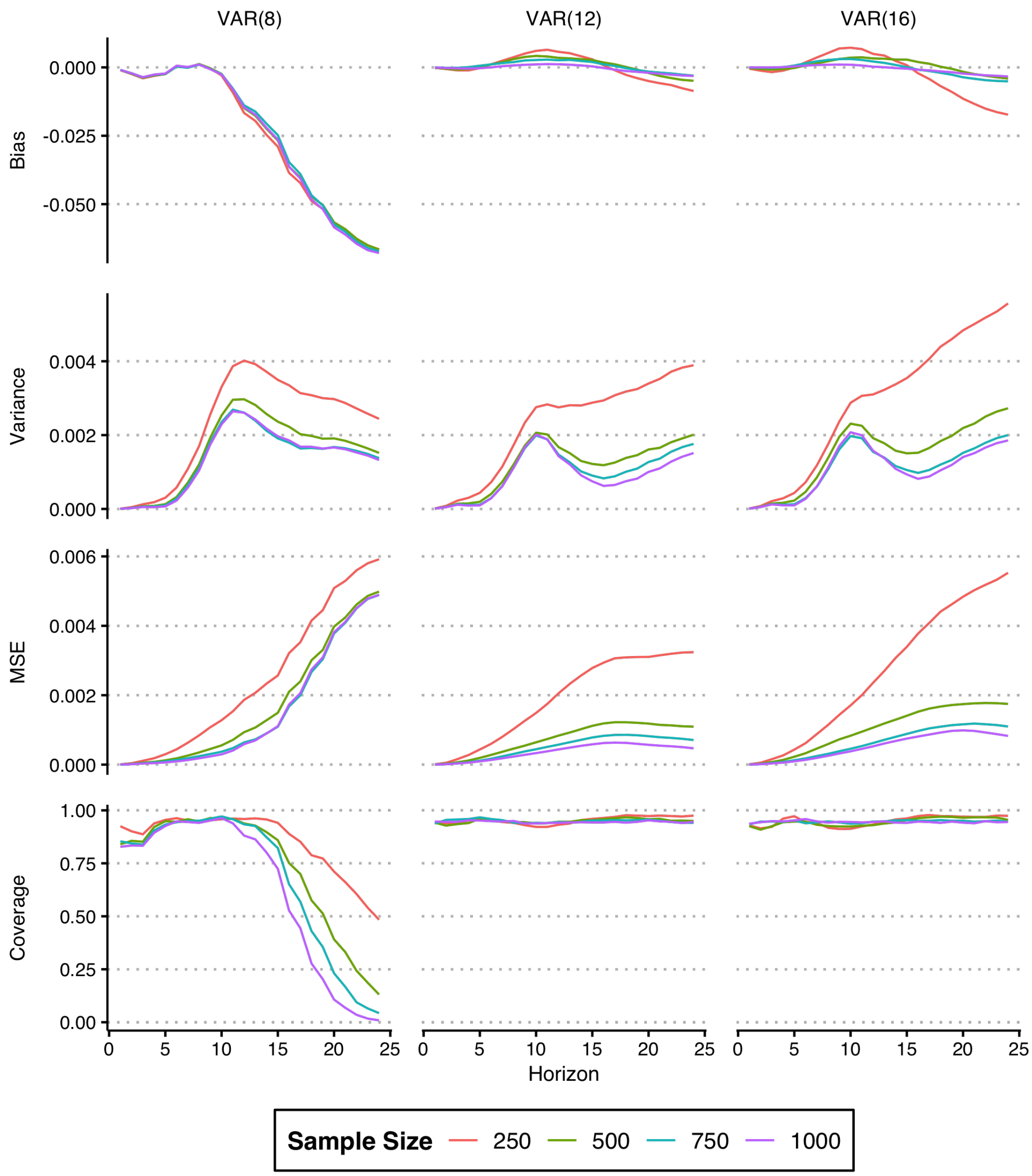

Figure A.1: Monte Carlo simulation analysis of the VAR estimator based on the Kilian-Lewis VAR as the population DGP. The nominal coverage probability is 95%. The simulation consists of 1000 samples for each given sample size.