Nonparametric Time-Varying Local Projections*

Kenwin Maung
Rutgers University
Nathaniel Yoshida
Rutgers University

Abstract: Local projections are widely used to estimate impulse response functions, yet there is substantial evidence that these responses change over time in many macroeconomic and financial settings. To better capture such dynamics, we adopt a nonparametric time-varying local projection framework and estimate the coefficients using local linear regression. Our approach is agnostic to the source of time variation and imposes minimal structure on the underlying data-generating process, making it suitable for analyzing macroeconomic and financial series with model uncertainty. To accommodate flexible temporal dependence and heavy-tailed distributions, we adopt the τ-mixing framework and develop a computationally tractable estimator for the long-run variance that can be an alternative to standard HAC corrections. In high-dimensional settings with a large number of controls, we incorporate a post-double selection lasso procedure to address model selection uncertainty and enable uniformly valid inference. Simulation studies demonstrate strong finite-sample performance, and an empirical application to severe weather shocks reveals significant departures from estimates obtained using parameterized nonlinear VARs. Our results highlight the advantages of a flexible, data-driven strategy for characterizing dynamic responses in complex economic environments.

JEL Classifications: *C13, C14, C22, C32, C53*

Key words: Time-Varying, Local Projections, Local Linear Estimation, Weak Dependence, High Dimensional, Post-Lasso Inference.

---

*Email: kenwin.maung@rutgers.edu. We would like to thank Laura Liu, Elena Pesavento, Tatevik Sekhposyan, Christian Matthes for valuable comments. We would also like to thank all seminar participants at the University of Pittsburgh, 95th Southern Economic Association Conference, 2025 Midwest Econometrics Conference, 17th Greater New York Metropolitan Area Econometrics Colloquium, and the Society of Economic Measurement 2024 conference for helpful comments.

1. Introduction

Recent empirical macroeconomic literature has revealed substantial evidence of time variation in impulse responses to exogenous macroeconomic and financial shocks. In its simplest form, the local projection method entails a series of predictive regressions of a variable of interest on a shock across multiple forecast horizons. The impulse response is then traced out by the sequence of regression coefficients associated with the shock at each horizon (see Jordà, 2005). Time-varying impulse responses within the framework of local projections suggest the presence of parameter instability, specifically time-varying coefficients.

A common approach to incorporate time-variation is to allow the coefficients to depend on some state of the economy, whether observed or latent, resulting in impulse responses that vary across different economic regimes. For instance, Tenreyro and Thwaites (2016) employ a smooth transition local projection to study the macroeconomic effects of U.S. monetary policy during expansions and recessions, concluding that monetary policy is less effective during recessions[1]. Owyang et al. (2013) and Ramey and Zubairy (2018) analyze government spending multipliers over different economic regimes using regime-switching local projections, with mixed results depending on how the regimes are defined.

In many cases, time variation in the coefficients is modeled to depend on some binary regime variable that switches with time. In smooth transition models, impulse responses would be a weighted combination of coefficients from two or more regimes. This approach, however, introduces challenges as researchers may define regimes differently, and the timing of switches can vary. Consequently, estimated impulse responses may differ depending on the chosen regime specification. For instance, Ramey and Zubairy (2018) find that the government spending multiplier does not change significantly when comparing between 'high' and 'low' unemployment states, but notable differences emerge when comparing periods constrained by the zero lower bound in monetary policy to those that are not.

More general models of time-varying local projections have previously been considered by Ruisi (2019) within a Bayesian framework. However, this approach does require parametrically specifying the exact form of parameter instability which is typically modeled as a random walk. In constrast, Inoue et al. (2024) offers an alternative that avoids specifying the instability process by using the path estimator of Müller and Petalas (2010). This procedure is particularly well-suited to study moderate time variations, as it assumes that the magnitude of instability tapers off as the sample size increases (i.e. local-to-zero

[1]Using a similar local projection, Piger and Stockwell (2025) find that this conclusion can be flipped when using data in long-differences specification instead of log-levels.

time variation).

We take a different approach here and do not assume that the parameter instability is driven by either a step or smooth function determined by ex-ante specified states. Instead, we assume that the coefficients are nonparametric functions of time that vary gradually and estimate them with a local linear regression. This approach aims to capture time-varying impulse responses in a more general and flexible setting. In a recent, closely related paper, Gao et al. (2024) study the estimation and inference of a time-varying vector autoregression (VAR) via nonparametric local linear regression. While impulse responses from a VAR can be obtained through an iterative approach, local projections offer a direct forecasting alternative that has grown in popularity. Key appealing features of local projections include their robustness to misspecification in the underlying data-generating process (Montiel Olea et al., 2024) and computational simplicity. Unlike VARs, where impulse responses are derived iteratively, local projections allow impulse responses across forecast horizons to be estimated directly as a series of coefficients from linear least squares regressions. This simplicity also extends to inference: confidence intervals for VAR impulse responses are often obtained using the Delta method or multivariate bootstrap, which can be computationally intensive. In contrast, confidence intervals for local projections can be constructed using a heteroskedasticity and autocorrelation consistent (HAC) estimate from the asymptotic variance of the estimated coefficients, similar to a standard linear time-series regression.

This difference in inference becomes even more relevant in high-dimensional settings, where estimation can be infeasible without regularization, such as ℓ_1 penalization (Lasso) or its extensions. One approach is to assume sparsity, where only a small subset of coefficients are truly non-zero, and estimate the high-dimensional VAR with Lasso penalization, as demonstrated by Kock and Callot (2015). However, impulse responses are highly nonlinear transformations of the post-Lasso estimated coefficient matrices, making inference via the Delta method vulnerable to variable selection errors and potentially a failure of the normality approximation (Pötscher, 2009; Leeb and Pötscher, 2008). Here, it is possible to extend the post-double selection procedure of Belloni et al. (2014), which provides uniformly valid inference for a subset of post-selection coefficients, to VAR models. Hecq et al. (2023) explore this approach in the context of Granger causality testing. However, direct application of this approach to VAR impulse response functions is challenging due to their nonlinear dependence on, typically, the full set of estimated VAR coefficients.

This problem can be mitigated in the local projection framework as impulse responses can be directly estimated without first needing to estimate the full set of parameters governing the dynamics of all the endogenous variables. Furthermore, we are typically interested only in the impulse responses themselves

and specifically, the few coefficients associated with the (structural) shock variables. The remaining control variables, which often include lags of the endogenous variables, are not estimation targets. This makes the framework particularly well-suited to the post-selection procedure of Belloni et al. (2014).

In addition to its modeling flexibility, our estimator accommodates macroeconomic time series data that may be heavy-tailed and exhibit complex dependence structures. We develop inference under a general class of weak dependence known as τ-mixing, which allows for broader forms of temporal dependence than standard mixing conditions and is well-suited to processes with heavier tails. To facilitate inference in this setting, we propose a simple and computationally efficient estimator of the long-run variance that serves as a practical alternative to Newey-West-type corrections.

We apply this high-dimensional selection procedure to study the time-varying impact of severe weather shocks on the macroeconomy. Here, the use of variable selection is almost necessary due to the large number of control variables involved (primarily large number of lags). We replicate the analysis by Kim et al. (2024) but instead of using the parameteric non-linear VAR they suggest, we estimate the impulse responses directly via local projection. Our approach is fully nonparametric, while the cited paper uses a smooth-transition VAR, and therefore impulse responses are a function of the convex combination of two regimes' worth of coefficients. The main goal of the analysis is to compare how the impact of weather shocks on the economy have changed over time from an early sample period of 1970s, to the 2010s. Throughout our analysis we find significant differences across both early and recent sub-samples. Arguably, the different results we arrive at can be attributable to the more general estimation procedure in our nonparametric approach and/or to the difference between direct and indirect forecasting. Specifically, the original authors find that severe weather shocks have a significantly negative impact on national industrial production in the latter period, while we observe some evidence of a positive impact which echoes the 'build back better' effect of Tran et al. (2020), for local economies, at a national level.

The innovations of this paper are three-fold. First, we propose a local linear estimator of time-varying impulse responses in a fully nonparametric framework. Unlike much of the existing literature, we do not impose a specific parametric form or local-to-zero restrictions on the evolution of the coefficients. Instead, we allow for gradual time variation, subject only to mild smoothness conditions. This enhances modeling flexibility and enables a data-driven approach to uncovering temporal heterogeneity in impulse responses. Additionally, our estimator can be implemented using standard econometric software and operates at speeds comparable to ordinary least squares which may have advantages over more computationally intensive Bayesian approaches.

Second, we develop inference procedures under weak assumptions that explicitly accommodate heavy-

tailed data and complex temporal dependence structures. In particular, we establish asymptotic normality under τ-mixing, a flexible class of dependence that nests many commonly used time series models. To facilitate inference, we also propose a simple, computationally tractable estimator of the long-run variance that serves as a practical alternative to HAC-type corrections in this setting.

Finally, we extend the post-double selection procedure of Belloni et al. (2014) to our time-varying, nonparametric local projection framework. This approach enables valid inference on impulse responses in high-dimensional settings with many control variables, such as lagged outcomes. The structure of local projections makes them particularly amenable to post-selection inference, and we provide corresponding asymptotic theory for the nonparametric time-varying case. We illustrate the usefulness of this procedure in our empirical application, where a large number of lags are required to potentially control for persistence in the macroeconomic variables of interest.

The rest of the paper is organized as follows. Section 2 introduces our local projection framework and compares it with existing impulse response estimators to highlight the broad applicability of our approach. We also detail the local linear estimation of the time-varying coefficients. Section 3 presents asymptotic normality results for the proposed estimator under potentially heavy-tailed and weakly dependent time series, and introduces a computationally tractable estimator of the long-run variance. In Section 4, we extend our framework to accommodate high-dimensional controls and propose a post-double selection lasso procedure tailored to the time-varying setting, along with theoretical guarantees. Section 5 evaluates the finite-sample coverage properties of our confidence intervals through simulations, while Section 6 applies the method to estimate the time-varying effects of weather shocks on macroeconomic outcomes. All mathematical proofs are provided in an online appendix.

2. Model Framework

Our baseline specification for time-varying local projections is given by the multi-step-ahead direct forecast equation:

$$y_{t+h} = \beta_t^h \varepsilon_t + \gamma_{0,t}^h + \gamma_{1,t}^{h\top} x_t + \nu_t^h, \tag{1}$$

where y_{t+h} is the scalar endogenous variable of interest h periods ahead, ε_t is the (structural) shock of interest, β_t^h denotes the time-varying impulse response at horizon h, x_t is a vector of control variables (typically including lags of endogenous variables), and ν_t^h is the local projection residual.

To streamline notation, we define the regressor vector as $X_t = (\varepsilon_t, 1, x_t^\top)^\top$ and the associated time-varying parameter vector as $\gamma_t^h = (\beta_t^h, \gamma_{0,t}^h, \gamma_{1,t}^{h\top})^\top$, yielding the compact form:

$$y_{t+h} = \gamma_t^{h\top} X_t + \nu_t^h. \tag{2}$$

To be precise, X_t contains m regressors. We adopt a nonparametric, agnostic approach and do not impose a specific functional form on how γ_t^h evolves over time. Instead, we assume the coefficients vary smoothly as functions of standardized time:

$$\gamma_t^h = \gamma^h(t/T),$$

where $\gamma^h : [0,1] \to \mathbb{R}^m$ is a smooth (twice continuously differentiable) function. Standardizing time as t/T allows us to leverage infill asymptotics (Robinson, 1989), ensuring that the amount of local information increases with the sample size and consistency is preserved.

2.1. Time-Series Dependence

It is well known that the local projection residual ν_t^h is typically serially correlated. If the underlying data-generating process (DGP) is a well-specified and invertible structural VAR (SVAR), then ν_t^h will generally be a linear combination of structural shocks from $t+1$ to $t+h$. In this case, inference requires explicit consideration of the long-run variance. More generally, however, if the DGP is misspecified, ν_t^h may exhibit more complex and potentially unknown forms of serial correlation, further complicating inference.

Montiel Olea and Plagborg-Møller (2021) offer an alternative perspective by assuming that the structural shock ε_t satisfies a strengthened martingale difference sequence (MDS) condition in a stationary VAR setting. Specifically, that its conditional expectation is zero given both past and future innovations. Under this stronger assumption, one may bypass the use of heteroskedasticity and autocorrelation (HAC)-robust standard errors. This condition is satisfied, for example, by i.i.d. or ARCH-type shock processes. A standard MDS condition alone is generally insufficient to justify this simplification.

Crucially, however, there are important cases in which even MDS-type assumptions fail. For instance, if the DGP is an SVAR identified via a recursive ordering (see subsection below), then ε_t corresponds to a contemporaneous variable in the endogenous vector. By construction, this variable follows an autoregressive process and is therefore serially correlated.

For these reasons, we allow for very general forms of temporal dependence in both $\{X_t\}$, which includes the shock and lagged controls, and $\{\nu_t^h\}$ throughout our analysis. Specifically, we assume that $\{X_t\}$ and $\{\nu_t^h\}$ belong to the broad class of τ-mixing processes in the sense of Dedecker and Prieur (2004) (see Section 3 for technical details). This dependence structure is appealing for at least three reasons. First, it is well known that even linear autoregressive processes (e.g., stationary AR(1)) may fail to satisfy commonly used mixing conditions, such as strong mixing, unless additional assumptions are made on the distribution of the innovations (Andrews, 1984). Such processes are, however, accommodated under

τ-mixing. Second, in the high-dimensional setting of Section 4, we allow the number of controls (typically, lags of endogenous variables) to diverge with the sample size. This setup aligns with the recommendations of Plagborg-Møller and Wolf (2021) for obtaining local projection impulse responses that are comparable to those from VARs. Yet, it is known that linear processes with infinitely many lags may fail to be strong or β-mixing (see Section 1.5 of Dedecker et al., 2007), whereas they can still satisfy τ-mixing. Finally, our framework accommodates heavy-tailed time series with polynomial decay, which are common in impulse response analysis involving financial or macroeconomic variables. The analysis of empirical processes under τ-mixing and polynomial tails was not well-developed until recently. This has changed following the sharp Fuk-Nagaev-type concentration inequality established in Babii et al. (2024), which we leverage in our theoretical results.

2.2. Structural Interpretation

Before proceeding further, we clarify the interpretation of β_t^h as a structural impulse response function. Without loss of generality, we focus on a single structural shock, as the extension to multiple shocks is straightforward. In our framework, ε_t serves as a placeholder for the researcher's chosen identification strategy.

The most straightforward and frequently adopted approach in applied economics is to treat the exogenous shock as directly observable. Several prominent examples follow this logic. For instance, Tenreyro and Thwaites (2016) analyze U.S. monetary policy shocks using state-dependent local projections and construct ε_t using a nonlinear adaptation of the Romer and Romer (2004) approach. In their study of government spending multipliers, Ramey and Zubairy (2018) employ a military news-based spending measure, identified via the narrative approach of Ramey (2011), as the shock. This series captures changes in government spending driven by political or military developments, which are plausibly exogenous to the economy. Similarly, Känzig (2021) studies the effects of oil supply news shocks (i.e. shocks to the expectation on future oil supply) on macroeconomic outcomes. These shocks are identified from high-frequency surprises around OPEC[2] announcements.

Alternatively, local projection impulse responses can be given a structural interpretation by comparison with a SVAR. Plagborg-Møller and Wolf (2021) detail the conditions under which impulse responses from local projections are equivalent to those from an appropriately identified SVAR. A widely used example is recursive (Cholesky) ordering, which imposes a (lower) triangular structure on the contemporaneous

[2]The Organization of the Petroleum Exporting Countries.

relationships among endogenous variables. In their monetary policy example, the vector of endogenous variables Y_t is ordered with several real and price variables first (collected in the subvector r_t), followed by the federal funds rate ($rate_t$), and then several financial variables (in the subvector q_t). Under this ordering, the local projection specification in equation (1) corresponds to identifying the structural shock as $\varepsilon_t = rate_t$, while the control vector x_t includes r_t, lags of r_t, lags of $rate_t$, and lags of q_t. In addition, Plagborg-Møller and Wolf (2021) show that equation (1) can also accommodate structural interpretations based on long-run restrictions and sign restrictions. See their paper for further details.

2.3. Related Models in the Literature

To situate equation (1) within the literature and to highlight the broad applicability of our approach, we now describe several approaches for estimating time-varying impulse responses that are either special cases of, or closely related to, the proposed model.

2.3.1. Parametric Time-varying VAR

Kim et al. (2024) study the time-varying impact of severe weather shocks on the macroeconomy between 1963 and 2019 using the following time-varying VAR:

$$Y_t = (1 - t/T)\left(\alpha_1 + \sum_{i=1}^{P} A_{i,1} Y_{t-i} + \Sigma_1 e_t\right) + (t/T)\left(\alpha_2 + \sum_{i=1}^{P} A_{i,2} Y_{t-i} + \Sigma_2 e_t\right),$$

where Y_t is a vector of endogenous variables, and e_t is standard normal and i.i.d. distributed. In this specification, both the autoregressive coefficient matrices and the variance-covariance matrices of the errors evolve as convex combinations of two distinct regimes, with the weights determined by standardized time (t/T). The VAR is recursively identified by ordering the exogenous weather shock first, followed by the remaining macroeconomic variables, effectively constituting a proxy VAR. The weather shock is captured by the Actuaries Climate Index (ACI), a composite measure reflecting the frequency of severe weather events and changes in sea level.

For tractability, we focus on a VAR(1) formulation and the impulse response of the second variable to the first (i.e., the weather shock). It is straightforward to show that the impulse response is itself time-varying, and under this setup, it is given by:

$$\mathrm{IRF}_{2,1}^{(h)}(t) = [\Psi_{h,t} B_t]_{2,1},$$

where $[\cdot]_{2,1}$ denotes the $(2,1)$ element of the matrix, $\Sigma_t = B_t B_t^\top$ is the Cholesky decomposition of the time-varying error covariance matrix, and the moving average (MA) coefficients are defined recursively as $\Psi_{0,t} = I$ and $\Psi_{j,t} = A_t \Psi_{j-1,t}$, with $A_t = (1 - t/T)A_1 + (t/T)A_2$ and $\Sigma_t = (1 - t/T)\Sigma_1 + (t/T)\Sigma_2$.

Under the conditions and recommendations of Plagborg-Møller and Wolf (2021), it is plausible that the time-varying local projection model in equation (1) can recover the same impulse responses, provided that the identification strategy and control set are correctly specified. However, this equivalence hinges critically on the correct specification of the underlying dynamics. For instance, if the true evolution of the VAR coefficients deviates from the assumed linear-in-time form (e.g., in the presence of nonlinear trends) the estimated impulse responses from the nonparametric local projection may outperform those from the misspecified VAR. Conversely, if the linear-in-time specification is valid, the resulting impulse response function becomes a special case of the more general nonparametric estimator. We return to both a detailed discussion of the ACI shock and a comparison of these approaches in our empirical application.

2.3.2. Nonparametric Time-varying VAR

A potential DGP associated with our local projection framework is the nonparametric time-varying VAR studied by Gao et al. (2024):

$$Y_t = \sum_{i=1}^{P} A_i(t/T) Y_{t-i} + \Sigma(t/T) e_t, \tag{3}$$

where $A_i(\cdot)$ and $\Sigma(\cdot)$ are smooth matrix-valued functions of standardized time, and $\{e_t\}$ is a MDS. Under the assumption that $\{Y_t\}$ is locally stationary, Proposition 1 of Gao et al. (2024) shows that the VAR process can be approximated by a time-varying vector moving average (VMA) representation:

$$\tilde{Y}_t = \sum_{j=0}^{\infty} B_j(t/T) e_{t-j}, \tag{4}$$

where $B_j(t/T) = \Psi_j(t/T)\Sigma(t/T)$, and $\Psi_j(t/T)$ denotes the time-varying impulse response matrix at horizon j (see their paper for a precise definition).

The key implication is that the VAR-implied impulse responses are deterministic and nonparametric functions of time, which is fully aligned with the structure of our local projection specification in equation (1). Nevertheless, we deliberately choose to model impulse responses directly via local projections for two main reasons, as highlighted in the introduction. First, local projections are typically more robust to misspecification of the underlying VAR structure. Second, they offer a more tractable framework for conducting valid post-selection inference, particularly when estimation involves high-dimensional controls and penalization.

2.3.3. Parametric Smoothly-Varying Local Projections

Consider the following smooth transition local projection model from Tenreyro and Thwaites (2016):

$$y_{t+h} = F(z_t)(\beta_b^h \varepsilon_t + \gamma_{0,b}^h + \gamma_{1,b}^{h\top} x_t) + (1 - F(z_t))(\beta_r^h \varepsilon_t + \gamma_{0,r}^h + \gamma_{1,r}^{h\top} x_t) + \nu_t^h,$$

where parameters indexed by 'b' refer to an expansionary regime, and those indexed by 'r' refer to a recessionary regime. The transition function $F(z_t)$ is a logistic function defined as

$$F(z_t) = \frac{\exp\left(\theta \frac{z_t - c}{\sigma_z}\right)}{1 + \exp\left(\theta \frac{z_t - c}{\sigma_z}\right)},$$

where z_t is a state variable, θ and c are slope and threshold parameters, respectively, and σ_z is the standard deviation of the state variable. In the baseline specification, z_t is defined as a seven-quarter moving average of real GDP growth.

While this approach is closely related to equation (1), there are two key differences. First, the impulse responses in this model vary smoothly between two regimes, with weights determined by a random variable. In contrast, our framework models impulse responses as deterministic functions of time. Second, the logistic specification imposes a parametric functional form, whereas our method places minimal restrictions on the nature of time variation and instead learns it nonparametrically from the data.

Thus, the approach in equation (1) can be viewed as agnostic on two fronts: both in terms of the source of time variation and the functional form through which it operates. Regarding the former, as discussed in the introduction, impulse responses can be highly sensitive to the choice of state variable, making a flexible, data-driven alternative appealing. As such, our nonparametric model may serve as a useful complement or robustness check to more tightly parameterized models such as this one.

2.4. Nonparametric Estimation

To estimate the time-varying coefficients, we will use data local to (i.e. close to) the desired time point (t/T) in a weighted fashion. This is the key idea behind the kernel smoothing approach (Cai, 2007). Consider another (standardized) time point s/T close to t/T such that $|s-t|/T < b$, and refer to b as the bandwidth. Asymptotically, we require $b(T) \equiv b \to 0$ as $T \to \infty$. We employ local linear estimation, and motivate it by taking a Taylor-expansion of our coefficients around t/T up to the second order in (2). This yields:

$$y_{s+h} = \gamma^h(s/T)^\top X_s + \nu_s^h \approx \gamma^h(t/T)^\top X_s + \left(\frac{s-t}{T}\right)\gamma^{h\prime}(t/T)^\top X_s + \nu_s^h. \tag{5}$$

Subsequently, define $k_{s,t} = \frac{1}{b}K(\frac{s-t}{Tb})$ where the kernel $K : [-1,1] \to \mathbb{R}^+$ is a user-specified symmetric probability density function. Examples include the uniform (which can coincide with a rolling window estimation), Epanechnikov, and quartic kernels. Based on (5), we can obtain the local linear estimator at time t by solving:

$$\min_{\theta_t^h = (\alpha_{0,t}^{h\top}, \alpha_{1,t}^{h\top})^\top \in \mathbb{R}^{2m}} T^{-1} \sum_{s=1}^{T-h} \left[y_{s+h} - \alpha_{0,t}^{h\top} X_s - \alpha_{1,t}^{h\top}\left(\frac{s-t}{T}\right) X_s \right]^2 k_{s,t}, \tag{6}$$

where $\alpha_{j,t}^h$ is the $m \times 1$ coefficient horizon-h vector for $(\frac{s-t}{T})^j X_s$ with $j = 0, 1$. Our estimate $\gamma_t^h = \gamma^h(t/T)$ is therefore:

$$\hat{\gamma}_t^h = \mathbf{e}^\top \hat{\theta}_t^h$$

where $\mathbf{e}$ is a $2m \times 1$ vector with 1 in its first m entries and 0 everywhere else, and the local projection impulse response can be isolated as the first coordinate of $\hat{\gamma}_t^h$:

$$\hat{\beta}_t^h = \mathbf{e}_1^\top \hat{\gamma}_t^h,$$

where $\mathbf{e}_1$ is a $m \times 1$ vector with 1 in the first entry and 0 elsewhere. To obtain the impulse responses for other horizons, say h', repeat the same optimization procedure, and replace y_{t+h} with $y_{t+h'}$.

The optimization of this weighted least squares function is computationally simple and typically as fast as ordinary least squares. This can be seen by noting that the kernel weights $k_{s,t+h}$ are non-negative scalars, and thus we can include it into the parentheses by taking its square root. More specifically, we have:

$$\min_{\theta_t^h = (\alpha_{0,t}^{h\top}, \alpha_{1,t}^{h\top})^\top} T^{-1} \sum_{s=1}^{T-h} \left[y_{s+h} \sqrt{k_{s,t}} - \alpha_{0,t}^{h\top} \sqrt{k_{s,t}} X_s - \alpha_{1,t}^{h\top} \left(\frac{s-t}{T} \right) \sqrt{k_{s,t}} X_s \right]^2 \tag{7}$$

which reduces to the regular ordinary least squares estimation using variables transformed with the deterministic sequences of $\{k_{s,t}\}$ and $\{(s-t)/T\}$. If one is interested in multiple time points, the least squares estimation can be easily looped across all time. This procedure is computationally fast and may have advantages over more intensive typical time-varying estimation techniques such as Bayesian methods, along with less restrictive assumptions on the time-variation.

3. Asymptotic Properties

Before stating the statistical guarantees, we discipline our analysis with the following assumptions:

Assumption A.1 (Smoothness): $\gamma_t^h = \gamma^h(t/T)$ is a smooth function with continuous second order derivative for all h.

Assumption A.2 (Moments): For a random variable $x \in \mathbb{R}$, define the L_p norm ($p \geq 1$), $\|x\|_p$, to be $(E|x|^p)^{1/p}$. Then, we have:

(i) $\|\nu_t^h\|_p < \infty$ and $\max_i \|X_{ti}\|_q < \infty$ for $p > 2q/(q-2)$ and $q > 4$, for all t and h;

(ii) $M(t/T) \equiv E[X_t X_t^\top], \sigma_h^2(t/T) \equiv E[(\nu_t^h)^2]$, and $\Gamma_{j,h}(t/T) \equiv Cov(X_t \nu_t^h, X_{t+j} \nu_{t+j}^h)$ are Lipschitz continuous in $[0, 1]$ for all t,h and j, and $M(t/T)$ is non-singular;

(iii) $E[\nu_t^h | X_t] = 0$ for all t and h.

Assumption A.3 (Weak Dependence): Let $\{\xi_t\}_{t\in\mathbb{Z}}$ be a sequence of random variables in $\mathbb{R}^l$, and let $\mathcal{F}^t_{-\infty}$ to be the σ-field generated by $\{\xi_j : j = t, t-1, t-2, ...\}$. Define $\Delta_1(\mathbb{R})$ to be the set of 1-Lipschitz functions from $\mathbb{R}^l$ to $\mathbb{R}$, and the following weak dependence coefficient

$$\tau(\mathcal{F}^t_{-\infty}, \xi_t) = \left\| \sup_{f\in\Delta_1(\mathbb{R})} |E(f(\xi_t)|\mathcal{F}^t_{-\infty}) - E(f(\xi_t))| \right\|_1,$$

where $\|x\|_1 = E|x|$ is the L_1-norm. For a $k \geq 0$, the τ-mixing coefficient is defined as

$$\tau_k = \sup_{j\geq 1} \frac{1}{j} \sup_{t+k\leq t_1<...<t_j} \tau(\mathcal{F}^t_{-\infty}, (\xi_{t_1}, ..., \xi_{t_j})),$$

with the latter supremum taken with respect to t and $(t_1, ..., t_j)$. We say a process is τ-mixing if its coefficient $\tau_k \to 0$ as $k \to \infty$ and require:

(i) For every $i, j = 1, ..., m$ and h, $\{X_{ti}X_{tj}\}_{t\in\mathbb{Z}}$ and $\{\nu^h_t X_{ti}\}_{t\in\mathbb{Z}}$ are τ-mixing with coefficients satisfying $\tau_k = O(k^{-\varphi})$ and $\tau^*_k = O(k^{-\varphi^*})$ respectively, with $\varphi > (q-2)/(q-4)$, $\varphi^* > (R-1)/(R-2)$, and $R = pq/(p+q)$;

(ii) For every $i, j = 1, ..., m$, h and $r \geq 0$, $\{\nu^h_t \nu^h_{t+r} X_{t,i} X_{t+r,j}\}_{t,r\in\mathbb{Z}}$ is τ-mixing with coefficients satisfying $\tilde{\tau}_k = O(k^{-\tilde{\varphi}})$ with $\tilde{\varphi} > 1$.

Assumption A.4 (Kernel and Bandwidth): $K : [-1, 1] \to \mathbb{R}^+$ is a bounded and symmetric probability density function with $\int K^2(u)du < \infty$ for all $u \in [-1, 1]$. The kernel satisfies a Lipschitz condition and $K(u) \leq K(0) < \infty$. The bandwidth $b \to 0$ as $T \to \infty$ but $Tb \to \infty$.

Assumption A.1 is standard in local linear nonparametric estimation and is not restrictive given that it applies to all the examples in Section 2.3. Notable exceptions would include a non-smooth varying coefficient local projection model where the state variable is discrete (e.g. a binary variable indicating the state of the economy). Here, we note that the results in this section will still hold away from the break point. Assumption A.2(i) requires at least a little more than 4 finite moments and the local projection residuals having more than 2 finite moments. Given the tradeoff between p and q, the error term need only slightly more than 2 moments if the regressors have sufficiently many moments. A.2(ii) is a standard characterization of smooth time-varying moments (see for e.g. Chen and Maung, 2023). Importantly, we do not require stationarity as the moments are allowed to vary gradually with time. This flexibility is crucial when modeling macroeconomic and financial series for which stationarity can be questionable. Assumption A.2(iii) is the typical exogeneity condition and is reasonable in this setting, as ν^h_t typically depends on structural shocks occurring in periods $t+1, \ldots, t+h$ in SVARs, which should be exogenous with respect to the lagged regressors in X_t. The shock ε_t is assumed to be constructed

so that this condition is satisfied. Assumptions A.3(i) and A.3(ii) closely follow Assumptions 2.1 and 2.5 in Babii et al. (2024), with A.3(i) required for consistency and A.3(ii) necessary for establishing asymptotic normality. As emphasized in Section 2.1, the τ-mixing framework is particularly appealing for its flexibility and generality.[3] Notably, strong mixing processes are nested within the τ-mixing class as a special case. Assumption A.4 is standard within the nonparametric literature and implies the following: $\int K(u)du = 1$, $\int uK(u)du = 0$ and $\int u^2K(u)du < \infty$.

We now summarize the asymptotic properties of $\hat{\gamma}^h(\tau)$ where $\tau \in (0,1)$. This can be viewed as an extension of Theorem 2 of Cai (2007) and of Proposition 2 in Chen and Maung (2023) to the τ-mixing setup.

Proposition 1. *Under assumptions A.1-A.4, then for all $\tau \in (0,1)$, we have for every horizon h,*

$$\sqrt{Tb}\left\{\boldsymbol{B}[\hat{\theta}^h(\tau) - \theta^h(\tau)] - \begin{bmatrix} \frac{b^2}{2}\mu_2\gamma^{h''}(\tau) \\ 0 \end{bmatrix} + o_p(b^2)\right\} \to^d N(0, \tilde{M}^{-1}(\tau)\tilde{\Omega}_h(\tau)\tilde{M}^{-1}(\tau)), \tag{8}$$

where we have defined $\tilde{M}(t/T) = diag\{M(t/T), \mu_2 M(t/T)\}$, $\tilde{\Omega}_h(t/T) = diag\{\nu_0\Omega_h(t/T), \nu_2\Omega_h(t/T)\}$ for $\Omega_h(\tau) = \sum_{j=-\infty}^{\infty}\Gamma_{j,h}(\tau)$, $\boldsymbol{B} = diag\{I_{(m\times m)}, bI_{(m\times m)}\}$, $\mu_j = \int u^jK(u)du$, and $\nu_j = \int u^jK^2(u)du$. We have defined $M(\tau)$ and $\Gamma_{j,h}(\tau)$ in Assumption A.2. $\gamma^{h''}$ refers to the second derivative of γ^h. Specifically we have,

$$\sqrt{Tb}[\hat{\gamma}^h(\tau) - \gamma^h(\tau) - \mu_2\frac{b^2\gamma^{h''}(\tau)}{2} + o_p(b^2)] \to^d N(0, \nu_0 M^{-1}(\tau)\Omega_h(\tau)M^{-1}(\tau)). \tag{9}$$

This result yields the consistency and asymptotic normality of the time-varying horizon h impulse response. There is however a bias due to the local linear approximation which depends on the curvature of $\gamma^h(\tau)$. Even though this result is fairly standard in the literature on nonparametric time-varying estimation, we remark that our approach is atypical and uses the triangular array Central Limit Theorem (CLT) of Neumann (2013). This strategy is particularly attractive because it relies on minimal dependence assumptions and accommodates non-stationarity.

[3] In their analysis of time-varying VAR models, Gao et al. (2024) similarly avoid imposing strong or β-mixing conditions. Instead, they take a bottom-up approach by characterizing dependence via the VMA representation. In our setting, where we do not impose a specific underlying DGP, we instead use the general τ-mixing framework to discipline dependence while maintaining modeling flexibility.

3.1. Confidence Intervals and Variance Estimation

Confidence interval construction is particularly important for impulse response functions. Here, we discuss how to obtain sample estimates of the asymptotic variance in Proposition 1. From Lemma $A.2$, we have

$$\hat{M}(t/T) \equiv \frac{1}{T}\sum_{s=1}^{T-h} X_s X_s^\top k_{s,t} \xrightarrow{p} E[X_t X_t^\top] = M(t/T),$$

so it remains to obtain a consistent estimator of the long-run variance. From the proof of Lemma $A.3$, it follows that

$$Tb \cdot \mathrm{Var}(r_0^h(t/T)) \equiv Tb\,\mathbb{E}\left[\left(\frac{1}{T}\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t}\right)^2\right] = \frac{b}{T}\,\mathbb{E}\left[\left(\sum_{s=1}^{T-h} X_s \nu_s^h k_{s,t}\right)^2\right] \to \nu_0 \Omega_h(t/T), \quad \text{as } T \to \infty,$$

where the simplification follows from the exogeneity condition in Assumption A.2(iii). Hence, a feasible method of moments estimator for each horizon h can be constructed as

$$\hat{\Omega}_h(t/T) = \frac{b}{T\nu_0}\left(\sum_{s=1}^{T-h} X_s \hat{\nu}_s^h k_{s,t}\right)\left(\sum_{s=1}^{T-h} X_s \hat{\nu}_s^h k_{s,t}\right)^\top,$$

where $\hat{\nu}_s^h$ denotes the estimated local projection residual at horizon h:

$$\hat{\nu}_t^h = y_{t+h} - \hat{\gamma}^h(t/T)^\top X_t.$$

Here, ν_0 is a parameter that is determined by the kernel. For the Epanechnikov, uniform, and quartic kernels it can be computed as $3/5$, $1/2$ and $5/7$ respectively.

Our next result states the consistency of the variance estimator:

Proposition 2. *Under the conditions of Proposition* 1 *and Assumption V in Appendix A, we have for all* $\tau \in (0,1)$ *and* h, $\hat{\Omega}_h(\tau) \xrightarrow{p} \Omega_h(\tau)$.

A key implication is that confidence intervals for the estimated impulse response functions $\hat{\beta}_t^h$ at time t and horizon h can be computed using:

$$\sqrt{\left[\hat{M}(t/T)^{-1}\hat{\Omega}_h(t/T)\hat{M}(t/T)^{-1}\right]_{(1,1)}},$$

where the impulse response is normalized to be the first component of the coefficient vector.[4]

The kernel smoothing over time ensures that nearby observations contribute to the long-run variance estimate in a locally weighted fashion. This construction leads to a computationally straightforward estimator, avoiding the need for additional bandwidth or lag-length selection as in HAC estimators like the

[4]This normalization is maintained throughout the paper for notational clarity.

popular Newey–West approach. It therefore underscores one of the practical strengths of the nonparametric kernel-based approach in local projections.

3.2. Bandwidth Selection

We have assumed the bandwidth b to be known in our analysis thus far. In practice, this is a tuning parameter that must be selected prior to estimation, and its choice can influence the bias–variance tradeoff. We highlight two main approaches to selecting b in the time-varying setting.

First, in macroeconomic and financial applications, it is very common to select b based on ad hoc rules. For example, in their analysis of short-horizon return predictability using a similar nonparametric time-varying estimator, Farmer et al. (2023) tune the bandwidth such that Tb is equivalent to 2.5 years. The authors are primarily interested in 'real-time' forecasting and emulate this with a (one-sided) boundary kernel. Effectively, the local (two-sided) window used is 5 years which is a typical rolling-window choice in many financial applications. Macroeconomic applications typically have longer rolling window lengths as the frequency of data is lower. We remark that a traditional 'rolling window' analysis coincides with the nonparametric time-varying estimator when a local constant model is used with the uniform kernel. Nonetheless, the choice of bandwidth here is informed by user experience, intuition or field of study and is computationally simple to obtain.

Alternatively, one may wish to employ a data-driven algorithm to select the bandwidth via optimizing some metric relating to estimation error. Given that the bandwidth is the same for each time-point[5] t, $b(h) \equiv b$ can be selected to optimize the integrated mean squared error for each local projection:

$$\int_0^1 MSE_t^h(b) d(t/T)$$

where

$$MSE_t^h(b) = E[(y_{t+h} - \gamma_t^{h\top} X_t)^2].$$

The optimal solution is given by

$$b^{opt} = T^{-1/5} \left(\frac{\nu_0 \int_0^1 \mathrm{Tr}\left[M(\tau) V_\gamma^h(\tau)\right] d\tau}{\mu_2^2 \int_0^1 \gamma^{h\prime\prime}(\tau)^\top M(\tau) \gamma^{h\prime\prime}(\tau) d\tau} \right)^{1/5},$$

where $V_\gamma^h(\tau) = M^{-1}(\tau)\Omega_h(\tau)M^{-1}(\tau)$, and hence the bandwidth satisfies the optimal rate of $O(T^{-1/5})$. Although one may attempt a plug-in approach to estimate b^{opt} via the given formula, it is computationally

[5]It is possible to allow for adaptive bandwidths, b_t, however, this would introduce additional computational costs and could potentially yield noisy estimates.

challenging to do so. Instead, it can be shown, albeit tediously, that a bandwidth chosen via the following leave-one-out cross-validation approach is consistent for the optimal bandwidth:

$$b^{CV} = \underset{c_1 T^{-1/5} \leq b \leq c_2 T^{-1/5}}{\operatorname{argmin}} CV(b)$$

where $CV(b) = T^{-1} \sum_{s=1}^{T-h} (y_{s+1} - \hat{\gamma}_{-s}^{h\top} X_s)^2$ and $\hat{\gamma}_{-s}$ is the local linear estimator estimated from using data that span $1, ..., T-h$ excluding the s-th observation. A simple execution of the minimization is to conduct it over a finite grid of potential bandwidths. This procedure can further be repeated for each horizon to obtain $b^{opt}(h)$ at the cost of computational complexity.

4. High-dimensional Controls

In this section, we consider the possibility of controlling for potentially many endogenous variables and their lags. This issue is exacerbated in our local linear setting for two reasons: (i) when estimating time-varying coefficients, the effective sample size is much smaller than the full sample, especially when the bandwidth b is small; (ii) local linear estimation effectively doubles the number of covariates compared to a local constant specification, as slope terms are introduced. When the number of regressors is large relative to the effective sample size, estimation may become computationally infeasible without some form of regularization.

A common solution in this context is to employ ℓ_1-penalized estimation (e.g., the lasso). Specifically, we can apply local linear lasso regression (see Chen and Maung, 2023, for a time-varying predictive regression context) to select relevant control variables in (1), and then perform a standard local linear regression of the target outcome on the shock and the selected controls. Specifically, in our nonparametric time-varying local projection context, the lasso solution is given by the penalized version of (6):

$$\tilde{\theta}_t^h = \underset{\theta_t^h = (\alpha_{0,t}^{h\top}, b\alpha_{1,t}^{h\top})^\top \in \mathbb{R}^{2m}}{\operatorname{argmin}} T^{-1} \sum_{s=1}^{T-h} \left[y_{s+h} - \alpha_{0,t}^{h\top} X_s - b\alpha_{1,t}^{h\top} \left(\frac{s-t}{Tb} \right) X_s \right]^2 k_{s,t} + \lambda_1 \|\alpha_{0,t}^{h(\neg)}\|_1 + \lambda_2 \|b\alpha_{1,t}^{h(\neg)}\|_1, \tag{10}$$

where $\alpha_{0,t}^{h(\neg)}$ indicates all the parameters in $\alpha_{0,t}^h$ excluding the first element which corresponds to the coefficient on the shock variable (likewise for $a_{1,t}^{h(\neg)}$), and $\|\cdot\|_1$ is the ℓ_1 norm. In other words, the shock variable is not subject to ℓ_1 penalization and would always be selected[6].

However, naive statistical inference on the estimated impulse response from (10) based on Proposition 1 may be problematic. The lasso is not generally variable selection consistent unless strong assumptions

[6]Additionally, we point out that for this section, the slope parameters have been appended with the bandwidth to simplify the theoretical discussion. As one implication, we can now take the orders of λ_1 and λ_2 to be proportional to each other.

are satisfied (such as the irrepresentability condition). Conducting standard inference after lasso-based selection is not advisable, as relevant variables correlated with the included regressors may be omitted, leading to omitted variable bias. As a result, the finite-sample distribution of the estimator can deviate substantially from the normal approximation (Leeb and Pötscher, 2008).

That said, in the special case where the observed shock is (i) exogenous (in the sense of a zero conditional mean) with respect to the other variables in the system, and (ii) serially uncorrelated, and where the control variables x_t consist solely of lags of the endogenous variables, the shock would be exogenous to the controls. In this case, variable selection mistakes would not impact the consistency or inference of the impulse response. However, such conditions are likely to be the exception rather than the rule. In particular, this framework breaks down under recursive identification schemes, such as those proposed by Plagborg-Møller and Wolf (2021), where the shock (i.e., a contemporaneous endogenous variable) would generally be correlated with the controls if its lags are included. Furthermore, it is not uncommon to include lags of the shock variable as controls, which might be problematic if the shock exhibits serial correlation.

For a robust and uniform approach to addressing this issue, we consider the post-double selection procedure of Belloni et al. (2014). This method has been adapted to the VAR context by Hecq et al. (2023) for conducting uniform inference (i.e., inference that is robust to variable selection errors) in Granger causality testing. Their procedure involves joint hypothesis testing of a subset of coefficients in a VAR. However, because impulse response functions are typically nonlinear functions of the full set of VAR coefficients, it is not straightforward to extend the post-double selection framework to inference on impulse responses in that setting.

In contrast, in local projections, the impulse responses are typically obtained as specific coefficients from a regression, and they are the primary objects of interest while the coefficients on other controls can be treated as nuisance parameters. This structure makes the local projection framework particularly well suited for applying the post-double lasso procedure[7].

[7]Certainly, alternative approaches to robust model selection can be considered. In the local projection context, Cha (2024) proposes estimating high-dimensional local projections using a flexible Orthogonal Greedy Algorithm together with a High-dimensional Akaike information criterion. Here, we prefer the lasso to preserve parsimony due to additional technical complications from the nonparametric estimation.

4.1. Post-double Lasso Framework

We formalize our framework here. First, given the high-dimensional set up, we now allow the number of regressors to deviate to infinity. Specifically, $m \equiv m_T \to \infty$ as $T \to \infty$. Similar to Hecq et al. (2023), we adopt exact sparsity for parsimony and allow only a small number of regressors in our framework to have non-zero coefficients[8]. Write the $(m-1) \times 1$ vector $z_s = (1, x_s^\top)^\top$, and consider the following system of equations:

$$y_{t+h} = \beta_t^h \varepsilon_t + \widetilde{\vartheta}_t^{h\top} z_t + \nu_t^h, \tag{11}$$

$$\varepsilon_t = \vartheta_t^{(1)\top} z_t + e_{1t}, \tag{12}$$

where $\vartheta_t^{(1)}$ is a coefficient vector of conformable dimensions. By substituting the reduced form equation for ε_s into the local projection we get

$$y_{t+h} = \vartheta_t^{h(2)\top} z_t + e_{2t}^h \tag{13}$$

where $\vartheta_t^{h(2)} = \beta_t^h \vartheta_t^{(1)} + \widetilde{\vartheta}_t^h$ and $e_{2t}^h = \beta_t^h e_{1t} + \nu_t^h$.

We assume that only $s_{(1),T}$ coefficients out of the $(m_T - 1)$ parameters in $\vartheta_t^{(1)}$ are non-zero, while the rest are zeroes. Similarly, only $s_{(2),T}$ out of $(m_T - 1)$ coefficients in $\vartheta_t^{h(2)}$ are non-zero. We treat the sparsity level collectively as $s_T = s_{(1),T} + s_{(2),T}$ where s_T is allowed to diverge to infinity albeit at reasonable rates to be defined below.

The algorithm for post-double selection in the time-varying local projection is given as follows. For a given horizon h:

1. First solve the kernel-weighted program,

$$\min_{a_t^{(1)} = (a_{0,t}^{(1)\top}, a_{1,t}^{(1)\top})^\top} T^{-1} \sum_{s=1}^{T-h} \left(\varepsilon_s - a_{0,t}^{(1)\top} z_s - b a_{1,t}^{(1)\top} \left(\frac{s-t}{Tb} \right) z_s \right)^2 k_{s,t} + \lambda_1^{(1)} \| a_{0,t}^{(1)} \|_1 + \lambda_2^{(1)} \| b a_{1,t}^{(1)} \|_1. \tag{14}$$

   Define the index set of non-zero coefficients for the 'level' terms (i.e. z_s) to be $I_1 = \{ j : a_{0,t,j}^{(1)} \neq 0 \}$.

2. Next solve,

$$\min_{a_t^{h(2)\top} = (a_{0,t}^{h(2)\top}, a_{1,t}^{h(2)\top})} T^{-1} \sum_{s=1}^{T-h} \left(y_{s+h} - a_{0,t}^{h(2)\top} z_s - b a_{1,t}^{h(2)\top} \left(\frac{s-t}{Tb} \right) z_s \right)^2 k_{s,t} + \lambda_1^{(2)} \| a_{0,t}^{h(2)} \|_1 + \lambda_2^{(2)} \| b a_{1,t}^{h(2)} \|_1, \tag{15}$$

   and likewise, keep the index set $I_2 = \{ j : a_{0,t,j}^{h(2)} \neq 0 \}$.

[8] An extension to a framework of approximate sparsity is possible at the cost of more rigorous exposition.

3. Next, define the sub-vector $z_s^{\mathcal{S}}$ which refers to all elements of z_s whose index belongs in $I_1 \cup I_2$. Finally, we solve the local linear problem:

$$\hat{\theta}_t^h = \underset{(\theta_{0,t}^h, \theta_{1,t}^h, \tilde{\theta}_{0,t}^{h\top}, \tilde{\theta}_{1,t}^{h\top})^\top}{\operatorname{argmin}} T^{-1} \sum_{s=1}^{T-h} \left(y_{s+h} - \theta_{0,t}^h \varepsilon_s - \theta_{1,t}^h \left(\frac{s-t}{T} \right) \varepsilon_s - \tilde{\theta}_{0,t}^{h\top} z_s^{\mathcal{S}} - \tilde{\theta}_{1,t}^{h\top} \left(\frac{s-t}{T} \right) z_s^{\mathcal{S}} \right)^2 k_{s,t}, \tag{16}$$

and $\check{\beta}_t^h \equiv e_1 \hat{\theta}_t^h$, where e_1 is a row vector of zeroes and one in its first position, is the post-double lasso impulse response function for horizon h at time t.

To discuss the validity of this double-selection procedure, we list the following conditions in addition to A.1-A.4:

Assumption H.1 (Model size and Penalty): We require:

(i) $s_T b^2 \propto (Tb)^{-1/2}$;

(ii) Let $\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_1^{(2)}, \lambda_2^{(2)} \propto \lambda$. For some $\delta \in (0,1)$, we require $\lambda_T(\delta) \equiv \lambda \to 0$ as $T \to \infty$, and satisfies

$$H_T(\delta) \equiv \left\{ \left(\frac{m_T}{\delta (Tb)^{\kappa-1}} \right)^{1/\kappa} \vee \sqrt{\frac{\log(24 m_T/\delta)}{(Tb)}} \vee \frac{m_T}{\delta (Tb)^{1/2}} \right\} \leq \lambda, \tag{17}$$

where $\kappa = ((\varphi^* + 1)R - 1)/(\varphi^* + R - 1)$, φ^*, and R are defined in Assumption A.3.

Assumption H.2 (Eigenvalues): The following conditions hold:

(i) Let

$$\tilde{M}(\tau) = \begin{bmatrix} M(\tau) & 0 \\ 0 & \mu_2 M(\tau) \end{bmatrix},$$

then we have $\lambda_{min}(\tilde{M}(\tau)) > \rho_1$ uniformly in $\tau \in [0,1]$ where $\lambda_{min}(\cdot)$ refers to the smallest eigenvalue, $M(\tau)$ and μ_2 have been defined in assumption A.2., and ρ_1 is a positive constant.

(ii) Define the set

$$C^\circ = \Bigg\{ v = (v_{11}, \ldots, v_{1(m_T-1)}, v_{21}, \ldots, v_{2(m_T-1)})^\top \in \mathbb{R}^{2(m_T-1)} : \sum_{j=1}^{m_T-1} (|v_{1j}| + |v_{2j}|) \leq 2(1+\eta) \sum_{j=1}^{s_T} (|v_{1j}| + |v_{2j}|) \Bigg\},$$

for some $\eta > 0$. Then, for all t, there exists positive constant κ such that with probability $1 - Q_T$ where $Q_T \to 0$ as $T \to \infty$, we have for $v \in C^\circ$ and $\kappa > 0$,

$$\kappa \|v\|_2^2 \leq \frac{1}{Tb} \|Z_t v\|_{K_t}^2 \tag{18}$$

where $\|e\|_{K_t} = \sqrt{e^\top K_t e}$ and $\|\cdot\|_2$ refers to the Euclidean norm. Z_t is the matrix of $(z_s^\top, z_s^\top(s-t)/(Tb))^\top$ stacked over $s = 1, ..., T-h$ and K_t is a diagonal matrix with diagonal elements given by $\{K((s-t)/Tb)\}_{s=1}^{T-h}$.

Assumption H.3 (Moments and dependence) We require the following analogous conditions for our post-double lasso framework: (i) $\|e_{it}\|_p < \infty$ where e_{it} is either e_{1t} from (12) or e_{2t}^h from (13); (ii) For every $i = 1, ..., m_T$ and h, $\{X_{ti}e_{1t}\}_{t\in\mathbb{Z}}$ and $\{X_{ti}e_{2t}^h\}_{t\in\mathbb{Z}}$ are τ-mixing with coefficients τ_k^*; (iii)$\{e_{1t}e_{1t+r}\nu_t^h\nu_{t+r}^h\}_{t,r\in\mathbb{Z}}$ is τ-mixing with coefficients given by $\tilde{\tau}_k$. Note that p, τ_k^* and $\tilde{\tau}_k$ are defined in assumptions A.2 and A.3.

Following Hecq et al. (2023), we assume a framework of exact sparsity to ensure that the resulting regression in step 3 of the algorithm is always computationally feasible, but the relaxation to approximate sparsity can be done. Noticeably, we do not require any beta-min conditions because we are not interested in the variable selection consistency of the selection procedure. The rate assumptions in H.1(i) are similar to Chen and Maung (2023) and Li et al. (2015) and are used to discipline the slope variables in the local linear expansion. H.1(ii) requires more explanation. Recall that we are working with τ-mixing variables with only finite moments to accommodate potentially heavy-tail processes. Hence our set-up deviates significantly from the i.i.d. sub-Gaussian scenario that is prevalent in the literature on lasso estimation. The first two rates are similar to rates in Babii et al. (2024) and the third rate stems from the local linear approximation. These rates help to bind the maximal probability bound of the first derivative or score vector which we show in the supplementary appendix in Lemma $A.6$. Nonetheless, the present rates are derived from a Fuk-Nagaev-type inequality processes with only a small number of finite moments and is sharp for polynomial τ-mixing time series processes[9] (as in assumption A.3). Assumption H.2(i) ensures that the data is well-behaved at the population level while H.2(ii) is commonly referred to as a restricted eigenvalue and the constraint set is the "cone constraint" for local linear lasso regression. We show additionally in Lemma $A.7$ that these assumptions are sufficient to provide for the estimation consistency of the lasso estimator of the time-varying coefficients, which is assumption 2(e) in Hecq et al. (2023) in the static VAR case.

The following theorem provides the asymptotic normality of the post-double lasso time-varying impulse response estimate $\check{\beta}_t^h$:

[9]The theory becomes much simpler if we assume exponential mixing instead. However, some processes are not necessarily exponentially mixing. For example, as shown in Dedecker and Prieur (2004), functional autoregressive process are only polynomially mixing. Given the reduced-form set-up of the proposed framework, we prefer to accommodate as large a class of processes as possible and therefore opt for polynomial mixing.

Theorem 1. *Under A.1-4, H.1-3, and suppose that the growth conditions hold:* $s_T\lambda^2 = o(\sqrt{Tb})$, $\varsigma_T^2 s_T = o(\sqrt{Tb})$, $\varsigma_T s_T \lambda = o(1)$ *where* $\varsigma_T = c\sqrt{\log(8m_T)/\delta_1}$ *and* δ_1 *is from Lemma A.6, then we have with high probability that for* $\tau \in (0,1)$*:*

$$\sqrt{Tb}\left(\check{\beta}^h(\tau) - \beta^h(\tau) - G(h)\right) \to^d N\left(0, \nu_0 m_e(\tau)^{-2}\omega^h_{\nu,e}(\tau)\right) \tag{19}$$

where $G(h) = O(h^2)$, $m_e(t/T) = E(e_{1t}^2)$ *and* $\omega^h_{\nu,u}(t/T) = \sum_{k=-\infty}^{\infty} Cov(\nu_t^h e_{1t}, \nu_{t+k}^h e_{1,t+k})$.

Theorem 1 can be seen as a non-parametric time-varying extension of Belloni et al. (2014), and the construction of an estimate of the asymptotic variance is analogous to the low-dimensional case as in the previous section.

5. Monte Carlo Simulations

In this section, we assess the finite sample properties of the nonparametric time-varying impulse response estimators. Specifically, we look at the mean squared error (MSE) of the nonparametric local projection estimator and the pointwise coverage rate of the 90% confidence interval. Our data generating process is a smoothed (and deterministic) version of the DGP in Gonçalves et al. (2024). Specifically, the general model is given by

$$\begin{aligned}
y_t &= \beta_t x_t + \alpha_t x_{t-1} + \sum_{i=1}^{p} \gamma_{t,i} y_{t-i} + \varepsilon_{2t} \\
x_t &= \rho x_{t-1} + \varepsilon_{1t} \\
\varepsilon_{2t} &= u_t + 0.75 u_{t-1},\ u_t \sim \mathcal{N}(0, \sigma_t^2),
\end{aligned}$$

where the time-varying parameters evolve as

$$\begin{aligned}
\beta(\tau) &= (2 + 0.2\sin(\tau))(0.5 + 0.05\cos(\tau)), \\
\gamma(\tau) &= \begin{cases} 0.1(-\tau^2) + 0.3 & \text{for } i = 1, \\ 0.044(0.1(-\tau^3) + 0.8)^{i-2} & \text{for } i > 1, \end{cases} \\
\sigma^2(\tau) &= 0.5(\tau - 0.5)^2 + 0.75.
\end{aligned}$$

We consider two designs. The first design (Design 1) is simpler and the 'shock' as exemplified by x_t is i.i.d. with $\rho = 0$ and $\alpha_t = 0$. Here, x_t corresponds to a directly observed shock that is commonly employed in applied work. In the second design (Design 2), we consider a serially correlated shock with $\rho = 0.4$ and $\alpha(\tau) = 0.3(0.1\tau - 1)^2$. Subsequently, for each design we consider two lag lengths with $p = 1$ (Low) and

$p = 12$ (High). Note that the coefficients on the higher-order lags decay rapidly, so although we do not have exact sparsity, the model can be made to be approximately sparse. Nonetheless, for all cases, our error term is serially correlated with time-varying variance.

We estimate our impulse response for horizons $h = 0, 2, 8, 12$ and at three distinct timepoints where $\tau = 0.25, 0.5$ and 0.75. These timepoints represent the end of the first quarter of the sample period, the midpoint, and the end of the third quarter of the sample, respectively. We consider sample sizes typically encountered in macroeconomic analyses with monthly data[10] with sample periods of $T \in \{400, 600, 800\}$.

5.1. Computation

Although our sample sizes may appear large, we emphasize that the effective sample size used in the local nonparametric estimation can be significantly smaller. To mimic common practice in macroeconomics, we employ $Tb \in \{48, 60, 72\}$ respectively which results in "windows" of approximately 8, 10 and 12 years[11]. Correspondingly, the bandwidths are $0.12, 0.10$ and 0.09. We note here that the chosen framework is indeed consistent with the required rates of $Tb \to \infty$ and $b \to 0$.

For the low-dimensional case, we estimate the impulse response as the first coefficient of the "level" terms (i.e. $\alpha_{0,t}^h$) in (6). For the high-dimensional set-up, the impulse response is estimated with the time-varying local projection via the double lasso procedure. In both lasso regressions, the tuning parameter is determined by 10-fold cross-validation. As mentioned in the discussion of (7), this problem is akin to doing 3 least squares regressions with transformed variables, 2 of which are with the lasso penalty (i.e. the steps in (14) and (15)). These 2 steps can easily be done with the 'glmnet' package in 'R' for example, in a matter of seconds even without requiring a high-powered device. To further reduce computational cost, we execute the tuning variable selection and conduct lasso selection for only the first forecast horizon. Subsequently, we use the same selected set of controls for all horizons. Extending this to the case with different selection sets for different horizons (for e.g. having different lag orders for different horizons) is straightforward and comes at a cost of more computation.

---

[10] For example, in our empirical application we have a sample size of 713 months.

[11] Given this, we regard the scenario where $p = 12$ as the 'high-dimensional' setup because when say $Tb = 48$, the effective sample size is 97. Due to the local linear estimation, the number of parameters to be estimated is doubled, which results in 28 coefficients. This yields a relatively high ratio of estimated coefficients to effective sample size.

5.2. Results

We report the MSE and the pointwise coverage rates for the 90% confidence interval for Design 1 in Table 1 and in Table 2 for Design 2. Our experiment consider 500 iterations. The standard errors are constructed according to the sample moments as implied by the asymptotic variances of Proposition 1 and Theorem 1. In all simulations, the Epanechnikov kernel was used, and therefore $\nu_0 = 3/5$.

[Table 1 about here.]

From Table 1, we can see that for both the low and high-dimensional setup in Design 1, the estimated confidence intervals attains coverage close to their nominal level. Although it has to be pointed out that in the majority of the cases where the coverage probabilities deviate, we do observe a situation of under coverage. Furthermore, as the sample size increases, we note that the MSE is monotonically decreasing for all horizons and time-points. Taken together this provides some evidence for the theoretical results established in Proposition 1 and Theorem 1 for the local linear estimator in the low-dimensional case and the post-double lasso local linear estimator respectively.

[Table 2 about here.]

We observe broadly similar results in Table 2 for the second design; however, there is a slightly higher degree of undercoverage when the sample size is small. These issues, however, become largely negligible as the sample size increases even for the high-dimensional setup.

6. Weather Shocks on the Macroeconomy

In a recent paper, Kim et al. (2024) study the impact of severe weather shocks on several key macroeconomic variables. They identify a novel shock series using the monthly indicator of the Actuaries Climate Index (ACI)[12], developed by the actuary associations in the U.S. and Canada. Essentially, the ACI index is a composite measure of the frequency of severe weather and changes to sea level. Specifically, the ACI consists of 6 components: (i) High temperatures, which looks at the changes in frequency of temperatures above the 90th percentile obtained from an early reference period; (ii) Low temperatures, which is the same as the above, but for below 10th percentile; (iii) Heavy precipitation, for the maximum amount of rainfall in a 5-day period in a month; (iv) Drought, which is similar to the previous indicator but for less than 1mm of precipitation; (v) High winds, which is the change in frequency of wind power above the

[12]This data is available at https://actuariesclimateindex.org/home/.

reference period's 90th percentile; (vi) Sea level, is the change in sea level as estimated by tide gauges at coastal stations (see Kim et al., 2024, for a detailed breakdown).

The measurements are taken by weather stations for the first 5 variables and to standardize them across all variables, they are first transformed into a z-score using the historical mean and standard deviation of the reference period (1961 - 1990). The sea level variable is obtained from coastal stations at fixed locations. The measurements are aggregated to the national level by taking the average of all stations, and a linear combination of all the variables is applied to arrive at the final index. When the index goes up, it suggests that anomalous weather activities are becoming more prevalent. As in the original paper, we will consider this to be a severe weather shock.

The authors estimate the impulse responses of this shock on several macroeconomic variables: (i) national industrial production growth, (ii) unemployment rate, (iii) CPI inflation, (iv) core CPI inflation, (v) short-term interest rate. This is done through a non-linear VAR that accommodates time-variation:

$$y_t = (1 - t/T)(\alpha_1 + \sum_{i=1}^{P} A_{i,1} y_{t-i} + \Sigma_1 e_t) + (t/T)(\alpha_2 + \sum_{i=1}^{P} A_{i,2} y_{t-i} + \Sigma_2 e_t), \tag{20}$$

where e_t is standard normal and i.i.d. distributed, and the intercept, coefficient matrices and error variance-covariance matrices can take on one of the two possible regime-options. The final coefficient matrices and variance-covariance matrix will be a convex combination of the two regimes depending on the deterministic time-varying weights. This model is estimated with Bayesian methods.

The key identification strategy of the VAR above is that severe weather shocks are not contemporaneously (short-run) related to the macroeconomic variables and the other variables are identified with a Cholesky ordering. We adopt the same approach here but with our local projection framework (see Plagborg-Møller and Wolf, 2021, for a detailed account on how to execute a recursive ordering in the local projection set-up).

It is important to point out that this non-linear VAR is a specific parameterization of the time-varying VAR in (3), and as such the impulse responses would also be a parameterized version of our nonparametric time-varying impulse responses. We rely on our nonparametric local projection framework and crucially, we do not impose a parameterization of the time-variation a priori. Instead, we adopt an agnostic and data-driven approach to uncovering the time-variation if any.

Our dataset is constructed to be exactly the same as Kim et al. (2024). This involves seasonally adjusting the monthly ACI index, and using the same data source (FRED) for all of the corresponding (seasonally adjusted) macroeconomic variables. Likewise, for the short-term interest rate, we replace the nominal interest rates during the zero lower bound periods with the shadow rate of Wu and Xia (2016).

We mirror our sample period to begin in 1961 January and to end in 2019 May.

The authors of the original paper looked at the impulse responses of the weather shocks at the very beginning of their sample, and at the very end of the sample. This is possible because of the Bayesian approach and the fact that VAR impulse responses are iterative. Here, local projections require us to know what the outcome value is at '$t+h$' to be used in the direct regression, and thus we would not be able to situate our impulse response analysis right at the end of our sample. Furthermore, the variance of our kernel-based estimator will suffer from boundary problems (i.e. lack of data on either side of the kernel) if we estimate at a time point that is too early or too late in the sample[13]. Hence, we consider points that are within the interior of the sample period but not too far off from 1961 and 2019. Specifically, we situate our first set of impulse responses in January 1970 and the latter set in December 2011.

We consider the same lag structure with 12 lags and 6 variables, to yield 72 control variables. An intercept is included in our estimation, and we use the post-double selection procedure of Section 4. Figure 1 reports the impulse responses of the macroeconomic variables (in order: IP growth, CPI inflation, core CPI inflation, unemployment rate, and the short-term interest rate) to a one-unit shock in the ACI-index for 1970 January and 2011 December.

[Figure 1 about here.]

Our results for 1970 are intuitive and align with expectations. Industrial production (IP) growth responds negatively to an ACI shock, with the effect becoming statistically significant roughly two years after the shock. This decline in output growth is consistent with panel-data evidence on weather shocks for U.S. states reported in Colacito et al. (2019), which, using data from 1957 to 2012, shows that increases in temperature negatively affect the growth rate of gross state product.

More interestingly, the estimated impulse responses for 2011 exhibit some surprising findings. Although an ACI shock marginally depresses IP growth on impact, it subsequently raises it in a statistically significant manner after roughly 20 months. The trajectory of IP, both the distinct phases of the response and the significant increase later in the horizon, is consistent with the findings of Tran et al. (2020), who document stages of recovery and a "build back better" effect. A similar mechanism is identified in Hornbeck and Keniston (2017), who show that widespread reconstruction following the Great Fire of 1872 led to improvements in Boston's capital stock. Although both studies focus on regional impacts,

[13]The latter problem can be circumvented by using a boundary or one-sided kernel as in Farmer et al. (2023) or by using data reflection as in Chen and Maung (2023).

greater economic interconnectedness, through longer, more geographically dispersed supply chains and the propagation of reconstruction-driven local demand shocks, makes it plausible that a severe weather event in one area can raise industrial production at the national level, especially in an environment where such events are becoming more frequent (as reflected in the upward trend in the ACI index).

Moreover, an adverse weather shock appears to exert inflationary pressures on both headline and core inflation. In 1970, these impacts are large, pronounced, and persistent. Such inflationary pressures may partly reflect disruptions to supply chains caused by adverse weather events. For example, Yim and Dall'erba (2025) find that severe weather events adversely impact the U.S. agrifood supply chain by reducing agricultural yields, altering domestic trade flows, and affecting food manufacturing. This suggests that supply chain disruptions represent a key channel through which adverse weather shocks can contribute to higher inflation. The observed reaction of core inflation indicates that the impact of the shock goes beyond agricultural supply chains, capturing wider adverse effects throughout the economy. Accordingly, the subsequent rise in the short-term interest rate is intuitive, as policymakers respond to these inflationary pressures.

These patterns are similarly reflected in the impulse responses for 2011, with one key difference: the magnitude of the responses is less than half of that observed in 1970. If inflation primarily reflects supply-side disruptions, the weaker impact in the later period may reflect improved resilience of the U.S. supply chain to weather-related shocks. Supporting this notion, Lim-Camacho et al. (2017) provide simulation evidence that complex supply chains with a large number of nodes and links are more resilient to climate-related disruptions. Consistent with this view, Cevik and Gwon (2024) find that, using data from 1997 to 2021, weather shocks do not have a statistically significant effect on U.S. supply chain pressures. Their sample period largely coincides with the latter portion of ours, lending further support to the argument that supply chain resilience has increased over time.

The impulse response of the unemployment rate to an ACI shock in 1970 is broadly consistent with that of IP growth, with the exception of a small dip around the five-month mark. Consistent with the longer-term decline in IP growth, unemployment begins to rise approximately 20 months after the shock, reflecting the delayed impact of output contractions on labor markets. In 2011, the direction of the responses remains the same, but, similar to our findings for inflation, the magnitude of the impulse response has fallen by more than half. This attenuation may reflect improvements in labor market flexibility, greater automation, and increased resilience of firms and supply chains to weather-related disruptions, which could dampen the transmission of local adverse shocks to national employment.

In summary, our findings differ markedly from those of Kim et al. (2024). Specifically, for the earlier

part of our sample, the estimated impulse responses are consistent with supply chain disruptions and cost-push inflation, whereas Kim et al. (2024) observe no statistically significant changes in any of the macroeconomic variables, except for persistent and statistically significant deflationary pressures on headline inflation. In contrast, the key difference in the latter-period impulse responses lies in the behavior of IP growth: we observe statistically significant positive responses over the longer term, whereas Kim et al. (2024) report a short-run negative change. We emphasize that these differences in empirical findings could stem from (i) differences in functional form assumptions; (ii) differences in indirect or iterative forecasting of impulse responses, as implied by the VAR approach, versus direct forecasting via local projections; or (iii) differences in the timepoints considered.

To further study the variation of impulse responses over time, we plot the point estimates of the responses for all horizons starting each month from January 1970 to December 2011 in Figure 2.

[Figure 2 about here.]

As discussed above, the responses of inflation, core inflation, and the short-term interest rate appear to have attenuated over time, as evidenced by an almost monotonic decline in their magnitudes. On the other hand, the variation in impulse responses for IP growth and unemployment are non-monotonic. Based on the argument that positive responses of IP growth may be driven by the "build back better" phenomenon, disaster-relief transfers are likely an important contributing factor. However, as documented by Zhou (2017), such transfers are highly variable over time, which could help explain the observed fluctuations in IP growth and unemployment responses across different periods. Periods with larger or more timely transfers may amplify recovery and boost production, while periods with smaller or delayed transfers may weaken the effect, leading to the non-monotonic pattern observed in the data.

7. Concluding Remarks

In this paper, we propose a nonparametric estimator for time-varying local projection impulse responses within a local linear kernel smoothing framework. Flexibility is a central theme of our approach, which we demonstrate along several dimensions. First, in terms of modeling, we remain agnostic about both the nature of time variation and the underlying data-generating process (DGP) for the impulse responses. Section 2.3 provides a detailed discussion of several existing models to illustrate the broad applicability of our method. Second, we explicitly allow for heavy-tailed processes and flexible forms of temporal dependence by working under τ-mixing conditions. We also propose a computationally straightforward estimator of the long-run variance to facilitate inference on the impulse response functions, thereby offering

an alternative to standard HAC estimators such as Newey-West. Finally, we show that local projections offer a natural and effective framework for analyzing impulse responses in high-dimensional settings. In such cases, we recommend a post-double selection lasso procedure to enhance finite-sample performance and enable uniformly valid inference, that is, inference that remains valid in the presence of model selection mistakes. Our simulation results, based on a variety of DGPs, corroborate the theoretical insights and demonstrate that the proposed confidence intervals achieve close-to-nominal coverage. We further apply our local linear estimator to evaluate the macroeconomic effects of severe weather shocks and uncover several notable differences compared to results from a parameterized nonlinear VAR. These differences potentially reflect the greater degree of data-driven flexibility afforded by the nonparametric strategy adopted in this paper.

References

Andrews, D.W., 1984. Non-strong mixing autoregressive processes. Journal of Applied Probability 21, 930–934.

Atak, A., Yang, T.T., Zhang, Y., Zhou, Q., 2025. Specification tests for time-varying coefficient panel data models. Econometric Theory 41, 123–170.

Babii, A., Ghysels, E., Striaukas, J., 2024. High-dimensional granger causality tests with an application to vix and news. Journal of Financial Econometrics 22, 605–635.

Belloni, A., Chernozhukov, V., Hansen, C., 2014. Inference on treatment effects after selection among high-dimensional controls. Review of Economic Studies 81, 608–650.

Cai, Z., 2007. Trending time-varying coefficient time series models with serially correlated errors. Journal of Econometrics 136, 163–188.

Cai, Z., Fang, Y., Xu, Q., 2022. Testing capital asset pricing models using functional-coefficient panel data models with cross-sectional dependence. Journal of Econometrics 227, 114–133.

Cevik, S., Gwon, G., 2024. This is going to hurt: Weather anomalies, supply chain pressures and inflation. International Economics 180, 100560.

Cha, J., 2024. Local projections inference with high-dimensional covariates without sparsity. arXiv preprint arXiv:2402.07743 .

Chen, B., Maung, K., 2023. Time-varying forecast combination for high-dimensional data. Journal of Econometrics 237, 105418.

Chen, B., Maung, K., 2025. High-dimensional time-varying forecast combination under general loss functions. Available at SSRN .

Colacito, R., Hoffmann, B., Phan, T., 2019. Temperature and growth: A panel analysis of the united states. Journal of Money, Credit and Banking 51, 313–368.

Dedecker, J., Doukhan, P., 2003. A new covariance inequality and applications. Stochastic processes and their applications 106, 63–80.

Dedecker, J., Doukhan, P., Lang, G., Rafael, L.R.J., Louhichi, S., Prieur, C., 2007. Weak dependence: With examples and applications. Springer.

Dedecker, J., Prieur, C., 2004. Coupling for τ-dependent sequences and applications. Journal of Theoretical Probability 17, 861–885.

Farmer, L.E., Schmidt, L., Timmermann, A., 2023. Pockets of predictability. The Journal of Finance 78, 1279–1341.

Gao, J., Peng, B., Yan, Y., 2024. Estimation, inference, and empirical analysis for time-varying var models. Journal of Business & Economic Statistics 42, 310–321.

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

Hecq, A., Margaritella, L., Smeekes, S., 2023. Granger causality testing in high-dimensional vars: A post-double-selection procedure. Journal of Financial Econometrics 21, 915–958.

Hornbeck, R., Keniston, D., 2017. Creative destruction: Barriers to urban growth and the great boston fire of 1872. American Economic Review 107, 1365–1398.

Inoue, A., Rossi, B., Wang, Y., 2024. Local projections in unstable environments. Journal of Econometrics , 105726.

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

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

Kim, H.S., Matthes, C., Phan, T., 2024. Severe weather and the macroeconomy. American Economic Journal: Macroeconomics , Forthcoming.

Kock, A.B., Callot, L., 2015. Oracle inequalities for high dimensional vector autoregressions. Journal of Econometrics 186, 325–344.

Leeb, H., Pötscher, B.M., 2008. Can one estimate the unconditional distribution of post-model-selection estimators? Econometric Theory 24, 338–376.

Li, D., Ke, Y., Zhang, W., 2015. Model selection and structure specification in ultra-high dimensional generalised semi-varying coefficient models. The Annals of Statistics 43, 2676–2705.

Lim-Camacho, L., Plagányi, É.E., Crimp, S., Hodgkinson, J.H., Hobday, A.J., Howden, S.M., Loechel, B., 2017. Complex resource supply chains display higher resilience to simulated climate shocks. Global Environmental Change 46, 126–138.

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

Montiel Olea, J.L., Plagborg-Møller, M., Qian, E., Wolf, C.K., 2024. Double Robustness of Local Projections and Some Unpleasant VARithmetic. Technical Report. National Bureau of Economic Research.

Müller, U.K., Petalas, P.E., 2010. Efficient estimation of the parameter path in unstable time series models. The Review of Economic Studies 77, 1508–1539.

Neumann, M.H., 2013. A central limit theorem for triangular arrays of weakly dependent random variables, with applications in statistics. ESAIM: Probability and Statistics 17, 120–134.

Owyang, M.T., Ramey, V.A., Zubairy, S., 2013. Are government spending multipliers greater during periods of slack? evidence from twentieth-century historical data. American Economic Review 103, 129–134.

Piger, J., Stockwell, T., 2025. Are the effects of monetary policy larger in recessions? a reconciliation of the evidence. The Journal of Economic Asymmetries 31, e00394.

Plagborg-Møller, M., Wolf, C.K., 2021. Local projections and vars estimate the same impulse responses. Econometrica 89, 955–980.

Pötscher, B.M., 2009. Confidence sets based on sparse estimators are necessarily large. Sankhyā: The Indian Journal of Statistics, Series A (2008-) , 1–18.

Ramey, V.A., 2011. Identifying government spending shocks: It's all in the timing. The Quarterly Journal of Economics 126, 1–50.

Ramey, V.A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from us historical data. Journal of political economy 126, 850–901.

Robinson, P.M., 1989. Nonparametric estimation of time-varying parameters. Springer.

Romer, C.D., Romer, D.H., 2004. A new measure of monetary shocks: Derivation and implications. American economic review 94, 1055–1084.

Ruisi, G., 2019. Time-varying local projections. Technical Report. Working Paper.

Tenreyro, S., Thwaites, G., 2016. Pushing on a string: Us monetary policy is less powerful in recessions. American Economic Journal: Macroeconomics 8, 43–74.

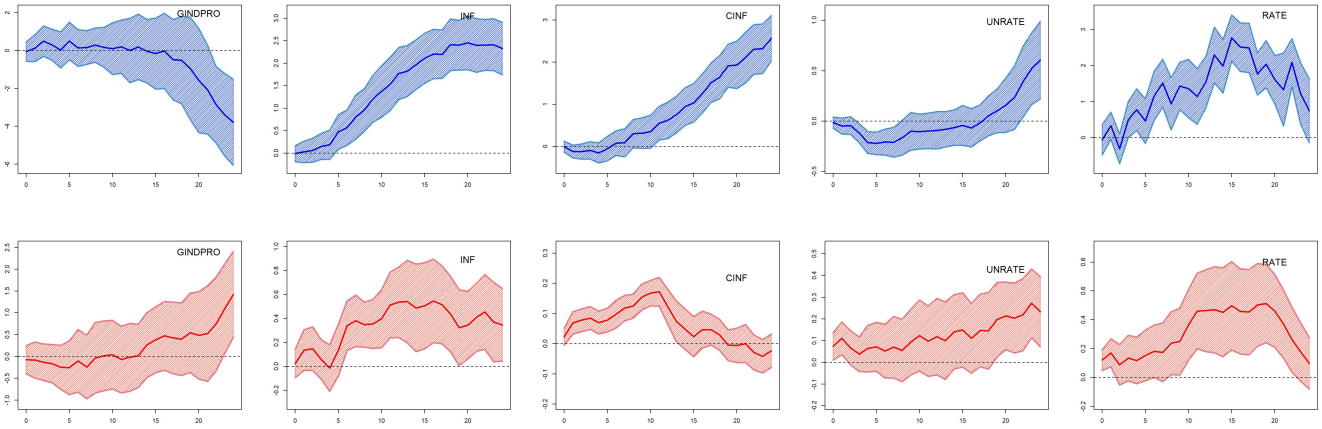
Tran, B.R., Wilson, D.J., et al., 2020. The local economic impact of natural disasters, Federal Reserve Bank of San Francisco.

Wu, J.C., Xia, F.D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. Journal of Money, Credit and Banking 48, 253–291.

Yim, H., Dall'erba, S., 2025. Impact of extreme weather events on the us domestic supply chain of food manufacturing. Proceedings of the National Academy of Sciences 122.

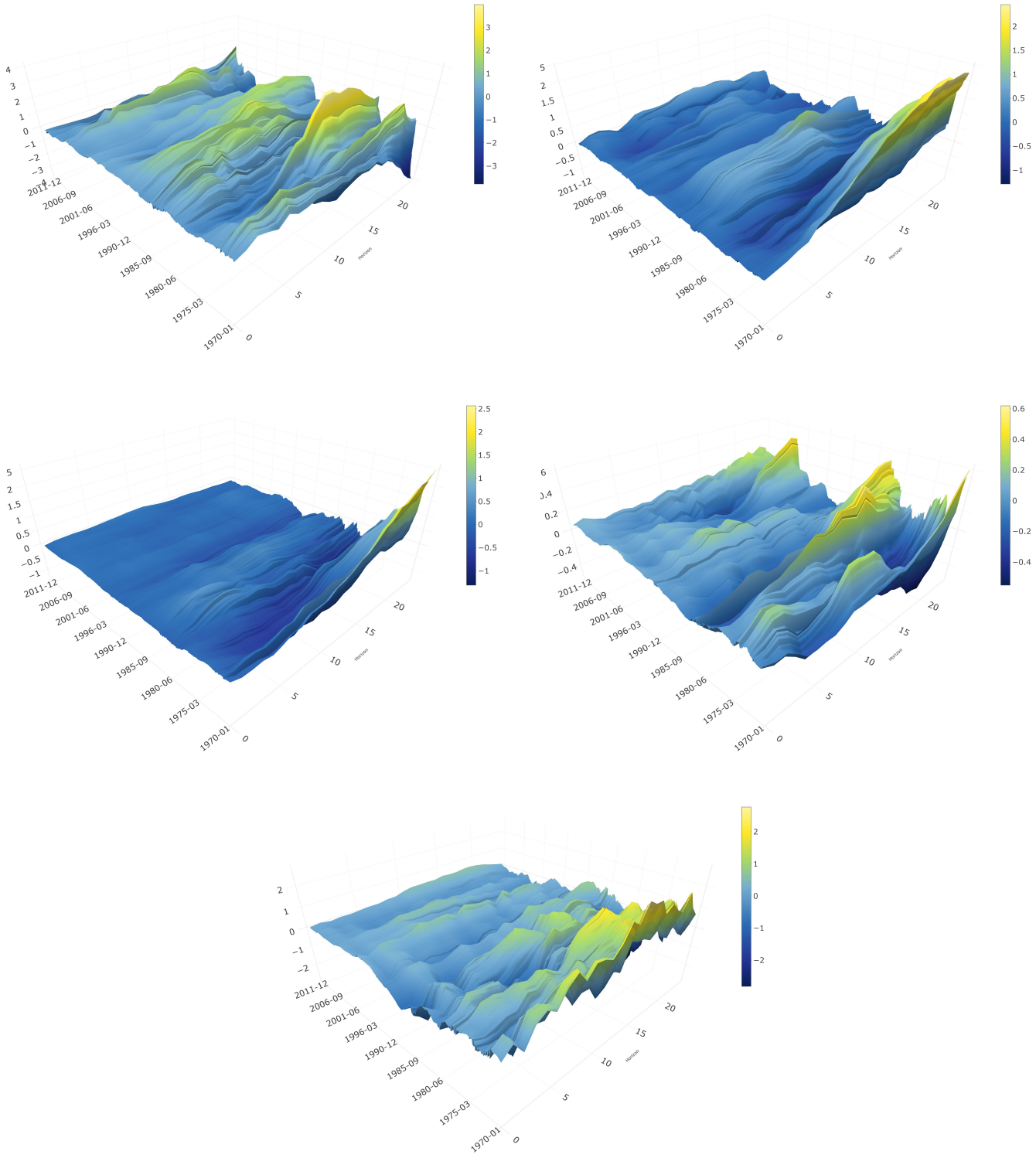
Zhou, X., 2017. Multiplier effects of federal disaster-relief spending: Evidence from us states and households. Available at SSRN 2928934 .

Figure 1: Impulse responses



Notes: Impulse responses of the macroeconomic variables (in order: IP growth, CPI inflation, core CPI inflation, unemployment rate, and the short-term interest rate) to a one-unit shock in the ACI index. The top row reports results for January 1970 and the bottom row is for December 2011. Shaded areas are 95% point-wise confidence interval.

Figure 2: Impulse responses over time



Notes: Impulse responses starting each month from January 1970 to December 2011 for a horizon of up to 24 months ahead. Top row: industrial production and inflation. Middle row: core inflation and unemployment. Bottom row: short-term interest rate.

Table 1: Simulation results for Design 1.

Low							High					
	MSE			90% Cov			MSE			90% Cov		
	tau											
h	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75
T=400												
0	0.008	0.008	0.010	0.88	0.86	0.85	0.010	0.009	0.010	0.82	0.84	0.83
2	0.031	0.034	0.032	0.89	0.83	0.85	0.031	0.035	0.032	0.87	0.84	0.86
8	0.029	0.033	0.030	0.88	0.86	0.87	0.030	0.033	0.030	0.86	0.83	0.88
12	0.033	0.031	0.029	0.87	0.85	0.88	0.032	0.027	0.028	0.84	0.89	0.88
T=600												
0	0.007	0.006	0.007	0.85	0.86	0.87	0.008	0.007	0.007	0.82	0.87	0.87
2	0.024	0.023	0.021	0.88	0.87	0.90	0.027	0.023	0.025	0.87	0.88	0.83
8	0.023	0.022	0.024	0.88	0.88	0.88	0.027	0.023	0.025	0.87	0.88	0.87
12	0.028	0.021	0.025	0.86	0.90	0.87	0.026	0.023	0.024	0.86	0.86	0.89
T=800												
0	0.006	0.005	0.006	0.87	0.88	0.86	0.006	0.005	0.006	0.87	0.86	0.86
2	0.020	0.019	0.020	0.88	0.88	0.87	0.022	0.017	0.018	0.87	0.88	0.87
8	0.023	0.020	0.020	0.87	0.89	0.88	0.020	0.019	0.020	0.88	0.87	0.87
12	0.018	0.019	0.020	0.91	0.89	0.88	0.021	0.019	0.020	0.87	0.88	0.89

Notes: MSE and 90% coverage probabilities for low ($p = 1$) and high ($p = 12$) dimensional cases. The first six columns correspond to the low-dimensional setup while the subsequent six columns are for the high-dimensional case. The first three columns of either half reports the MSE and the next three columns report the coverage probabilities. Each of the three columns correspond to estimating the impulse response at $\tau = 0.25, 0.5, 0.75$. Results are reported for $h = 0, 2, 8, 16$ and for sample sizes $T = 400, 600, 800$.

Table 2: Simulation results for Design 2.

Low							High					
	MSE			90% Cov			MSE			90% Cov		
	tau											
h	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75
T=400												
0	0.009	0.009	0.010	0.87	0.86	0.85	0.010	0.008	0.009	0.84	0.84	0.85
2	0.054	0.053	0.052	0.84	0.81	0.81	0.055	0.062	0.054	0.81	0.78	0.79
8	0.048	0.052	0.044	0.87	0.85	0.87	0.050	0.051	0.042	0.84	0.83	0.86
12	0.051	0.047	0.043	0.85	0.85	0.86	0.053	0.043	0.040	0.83	0.88	0.88
T=600												
0	0.008	0.007	0.007	0.86	0.85	0.88	0.008	0.006	0.007	0.84	0.87	0.87
2	0.039	0.038	0.036	0.85	0.84	0.85	0.047	0.037	0.041	0.78	0.83	0.82
8	0.038	0.038	0.037	0.87	0.86	0.88	0.044	0.039	0.035	0.85	0.85	0.86
12	0.042	0.034	0.040	0.85	0.90	0.86	0.043	0.038	0.035	0.87	0.88	0.87
T=800												
0	0.006	0.005	0.006	0.88	0.87	0.86	0.006	0.005	0.006	0.87	0.86	0.87
2	0.030	0.030	0.031	0.87	0.86	0.83	0.035	0.027	0.027	0.83	0.88	0.87
8	0.038	0.030	0.030	0.86	0.86	0.86	0.033	0.029	0.030	0.88	0.88	0.88
12	0.029	0.032	0.031	0.89	0.89	0.88	0.036	0.029	0.027	0.86	0.87	0.88

Notes: MSE and 90% coverage probabilities for low ($p = 1$) and high ($p = 12$) dimensional cases. The first six columns correspond to the low-dimensional setup while the subsequent six columns are for the high-dimensional case. The first three columns of either half reports the MSE and the next three columns report the coverage probabilities. Each of the three columns correspond to estimating the impulse response at $\tau = 0.25, 0.5, 0.75$. Results are reported for $h = 0, 2, 8, 16$ and for sample sizes $T = 400, 600, 800$.