Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Random Subspace Local Projections

CAMA Working Paper 34/2023
July 2023

Viet Hoang Dinh
Monash University

Didier Nibbering
Monash University

Benjamin Wong
Monash University
Centre for Applied Macroeconomic Analysis, ANU

Abstract

We show how random subspace methods can be adapted to estimating local projections with many controls. Random subspace methods have their roots in the machine learning literature and are implemented by averaging over regressions estimated over different combinations of subsets of these controls. We document three key results: (i) Our approach can successfully recover the impulse response function in a Monte Carlo exercise where we simulate data from a real business cycle model with fiscal foresight. (ii) Our results suggest that random subspace methods are more accurate than factor models if the underlying large data set has a factor structure similar to typical macroeconomic data sets such as FRED-MD. (iii) Our approach leads to differences in the estimated impulse response functions relative to standard methods when applied to two widely-studied empirical applications.

Keywords

Local Projections, Random Subspace, Impulse Response Functions, Large Data Sets

JEL Classification

C22, E32

Address for correspondence:

(E) cama.admin@anu.edu.au

ISSN 2206-0332

The Centre for Applied Macroeconomic Analysis in the Crawford School of Public Policy has been established to build strong links between professional macroeconomists. It provides a forum for quality macroeconomic research and discussion of policy issues between academia, government and the private sector.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

Random Subspace Local Projections*†

Viet Hoang Dinh[1], Didier Nibbering[1], and Benjamin Wong[1,2]

[1] *Monash University, Australia*

[2] *CAMA, ANU*

July 28, 2023

Abstract

We show how random subspace methods can be adapted to estimating local projections with many controls. Random subspace methods have their roots in the machine learning literature and are implemented by averaging over regressions estimated over different combinations of subsets of these controls. We document three key results: (i) Our approach can successfully recover the impulse response function in a Monte Carlo exercise where we simulate data from a real business cycle model with fiscal foresight. (ii) Our results suggest that random subspace methods are more accurate than factor models if the underlying large data set has a factor structure similar to typical macroeconomic data sets such as FRED-MD. (iii) Our approach leads to differences in the estimated impulse response functions relative to standard methods when applied to two widely-studied empirical applications.

Keywords: Local Projections, Random Subspace, Impulse Response Functions, Large data sets

*Viet.Dinh@monash.edu, Didier.Nibbering@monash.edu, and Benjamin.Wong@monash.edu.

†We thank participants at a seminar at BI Norwegian Business School, the 2023 Econometric Society of Australasia Meeting, and the 2023 International Conference in Computing in Economics and Finance for helpful comments and suggestions. Benjamin Wong acknowledges funding support from the Australian Research Council (DP190100202 and DE200100693)

1 Introduction

Impulse response functions (IRFs) are a common tool in macroeconomics to study the dynamics of variables in response to shocks. A recent development is to estimate IRFs using local projections, developed by Jordà (2005) (e.g., see Tenreyro and Thwaites, 2016; Ramey and Zubairy, 2018; Swanson, 2020). Local projections are single-equation methods, and can be thought of as the direct forecast counterpart of traditional multiple equation systems, such as vector autoregression (VAR), which rely on iterating on a system of equations.

A parallel development in the macroeconometrics literature has been the use of large data sets. The cost of putting together a large data set for empirical analysis, at least for U.S. macroeconomic data, is now becoming exceedingly trivial, especially with data sets such as FRED-MD and FRED-QD (McCracken and Ng, 2020). Like any regression model, the choice of control variables to use in a local projection forms part of the model specification. With the availability of large data sets, the dimension of this control set has become potentially large.

The contribution of our paper is to introduce random subspace methods for the estimation of IRFs using local projections. The method is simple to implement by three steps: First, we include a random subset of controls in the local projection and estimate the IRF. Second, we repeat the first step many times. Finally, we take the average IRF. Through averaging over random subsets of controls, where the subsets are generated independently of the data, the variance of the IRF estimate is reduced while maintaining most of the signal of the controls.

A practical implication of our approach is that it frees up the researcher to focus on the IRFs, which is the key object of interest, and leaves the issue of appropriately including controls to the random subspace method. This is an appealing practical feature because the coefficients on the controls are almost never the object of interest, as long as the appropriate controls are included in order to obtain appropriate estimates of the IRFs.

We highlight that random subspace methods apply more naturally within the local projection setting despite the now-known result that local projection and VAR estimate the same IRF in population (see Plagborg-Møller and Wolf, 2021). In a VAR setting, an additional control variable implies an extra equation. In the local projection setting, extra variables imply just more controls within the single equation. With the potential control variables numbering into the tens or hundreds, appropriately dealing with extra, possible extraneous, controls is arguably more manageable than estimating tens or hundreds of extra, possibly irrelevant, equations.

Our key results are as follows. First, we show that random subspace methods can recover the true IRFs under plausible factor structures and usual identification techniques for the structural shock of interest. We base this conclusion of Monte Carlo experiments using a real business cycle model with fiscal foresight introduced by Leeper et al. (2013). In this setting, the proposed method is implemented with both SVAR identification (Plagborg-Møller and Wolf, 2021), and instrumental variable (IV) identification (Stock and Watson, 2018). Random subspace methods help accurately estimating responses to a simulated tax shock, using a large set of controls that exhibits different factor structures resembling empirical macroeconomic data sets.

Second, while we demonstrate that we are able to recover the true IRF in different factor structures, our Monte Carlo exercise shows that random subspace methods might be more appropriate than the factor-augmented VAR (FAVAR) when the factor structure resembles that of U.S. macroeconomic data sets such as FRED-MD. The intuition is as follows: It is known that when the large macroeconomic data set is characterized by moderate –but not extremely strong– comovement, random subspace methods can better process the signals of relevant comovement from the data than factor models (see Boot and Nibbering, 2019).

Third, we document in two widely studied empirical applications (i.e. the effect of the macroeconomy to an identified monetary policy and technology shock) that random subspace local projections produce meaningful differences in the estimated IRFs relative to common empirical strategies. Our results are encouraging as it suggests that the random subspace local projection is a simple procedure or robustness check for practitioners using local projections who may be concerned that they may have omitted relevant controls.

We link our work to two strands of the broader literature. First, although random subspace methods have their roots in the machine learning literature, researchers have recently applied them to improve forecast accuracy for economic indicators based on a large number of possible predictors (Koop et al., 2019; Kotchoni et al., 2019; Boot and Nibbering, 2020; Pick and Carpay, 2022). In general, they find strong performance of random subspace methods across different macroeconomic forecasting exercises. Since the data sets used in macroeconomic structural analysis are similar, exploring random subspace methods to estimating IRFs seems natural.

Second, Barnichon and Brownlees (2019); Ferreira et al. (2023); Ho et al. (2023) propose approaches that may improve the efficiency of the local projection estimator, which has a lower bias but higher variance than VARs (Li et al., 2022). One can view our approach in a similar vein. Given that proliferation of controls leads to inefficiency, researchers naturally economize on the number of controls. Random subspace methods are a form of regularization, in which averaging across subsets of controls targets the bias-variance trade-off between omitting potentially relevant information and including all available data as controls.

The remainder of the paper is organised as follows: Section 2 provides a detailed discussion of our framework. Section 3 uses a Monte Carlo exercise to understand how random subspace methods help in appropriately estimating IRFs. Section 4 applies our proposed approach to two widely studied empirical applications. Section 5 provides some concluding remarks.

2 Method

Consider the standard local projection regressions one specifies to estimate the IRF to the variable y from an exogenous one unit impulse on x_t:

$$y_{t+h} = \mu_h + \beta_h x_t + \Phi_h' W_t + \xi_{t+h}, \quad h = 0, 1, \ldots, H, \tag{1}$$

where μ_h, β_h, and Φ_h are projection coefficients, ξ_{t+h} is the projection error, and W_t is a vector of controls. We are interested in the response of y_{t+h} with respect to an exogenous one unit impulse to x_t, which equals β_h. The problem we investigate in this paper is how one deals with the control set W_t. While one is almost never interested in the coefficients Φ_h, the specification of the control set can matter to obtaining accurate estimates of β_h.

We briefly highlight two roles the controls play in the estimation of β_h. First, the set of control variables accounts for relevant determinants which are correlated with x_t. In this setting, failure to include the relevant controls results in omitted variable bias, and so biased estimates of the IRF. Second, even if x_t is strictly exogenous, or we possess a strictly exogenous instrument, including additional controls in the projection may reduce the variation in forecasting y_{t+h}, which may result in a more precise estimate for β_h. However, in this setting, the effect is less obvious. If the included set of controls is too large, the increase in parameter uncertainty reverses possible efficiency gains, and may then increase the variance of the estimate for β_h.

With large macroeconomic data sets like FRED-MD, empirical work in macroeconomics now has access to over a hundred variables that may serve as potential controls. This leads to the familiar trade-off for practitioners where omitting relevant controls results in bias, but including extraneous and irrelevant variables leads to an increase in variance. This is the setting that we propose solving through random subspace local projections (RSLP): we alleviate the issue of specifying W_t, mindful that a practitioner is often only interested in β_h.

2.1 Random Subspace Local Projections (RSLP)

Instead of a generic set of controls W_t, we make a distinction between the $p_V \times 1$ vector V_t of variables which are considered essential controls, and the $p_G \times 1$ vector G_t of possibly relevant controls.[1] Rewriting (1) in terms of these essential and possible controls:

$$y_{t+h} = \mu_h + \beta_h x_t + \Theta_h V_t + \Psi_h G_t + \xi_{t+h}, \tag{2}$$

where Θ_h and Ψ_h are the projection coefficients of V_t and G_t, respectively.

The number of possibly relevant controls in G_t is potentially large as previously motivated. Instead of estimating β_h conditional on all these controls, a form of dimension reduction is usually applied to G_t. Consider the linear projection from (2) with a $k \times p_G$ compression matrix $R^{(j)}$ indexed by j, where $k \leq p_G$:

$$y_{t+h} = \mu_h^{(j)} + \beta_h^{(j)} x_t + \Theta_h^{(j)} V_t + \Gamma_h^{(j)} R^{(j)} G_t + \xi_{t+h}^{(j)}, \tag{3}$$

where $\Gamma_h^{(j)}$ is a k-dimensional vector of projection coefficients, instead of the p_G-dimensional vector of projection coefficients Ψ_h in (2). The construction of the compression matrix can be data-driven. For instance, testing procedures (e.g. Chudik et al., 2018) or variable selection with information criteria use data to estimate a selection matrix for $R^{(j)}$. Factor-augmented models take $R^{(j)}$ as the matrix of the principal component loadings corresponding to the k largest eigenvalues from the sample covariance matrix of G_t. In these approaches, the selection of the controls in G_t, or the derived factors, may be subject to substantial uncertainty.

The random subspace approach to dimension reduction is to generate the elements of $R^{(j)}$ from a probability distribution that is independent of the data. More precisely, let $R^{(j)}$ be a random subset selection matrix which randomly selects a subset of k predictors out of p_G available predictors. For instance, if there are 5 possible predictors and we wanted to choose 3, $p_G = 5$ and $k = 3$, and one possible draw j could be the following:

$$R^{(j)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

[1] While nothing precludes specifying V_t as an empty set, most macroeconomic applications would *a priori* treat some variables as essential. For example, a monetary policy application naturally suggests that one would need a real activity, inflation, and interest rate variable as being essential if taking guidance from a standard three equation New Keynesian model.

Conditional on a draw $R^{(j)}$, $\beta_h^{(j)}$ in (3) can be estimated using least squares. The random subspace estimate for $\beta_h^{(j)}$ is constructed by averaging over the least squares estimates $\hat{\beta}_h^{(j)}$ corresponding to different draws $R^{(j)}$, with $j = 1, \ldots, n_R$:

$$\hat{\beta}_h = \frac{1}{n_R} \sum_{j=1}^{n_R} \hat{\beta}_h^{(j)}. \tag{4}$$

Through averaging, random subspace methods reduce the variance of the local projection estimate, while it maintains most of the signal of the controls.

Random subspace methods have shown to improve forecast accuracy for economic indicators based on a large number of possible predictors (Koop et al., 2019; Kotchoni et al., 2019; Boot and Nibbering, 2019; Pick and Carpay, 2022). We argue that the random subspace approach also neatly fits into the local projection framework. The IRFs β_h can be seen as the differences in forecasts with and without an exogenous impulse to x_t. Where forecasting exercises focus on accurate estimates for y_{t+h}, our proposed method focuses on β_h. Both exercises are only interested in V_t and G_t to the extent that they matter for the estimation of y_{t+h} or β_h.

2.2 Structural identification within RSLP

The exposition so far discusses how one can use random subspace methods to estimate IRFs in a local projection setting. Since macroeconomists often estimate IRFs to study the effect of exogenously identified shocks on various macroeconomic variables, we connect our approach to structural models.

Define w_t as an $n \times 1$ vector of macroeconomic variables. Both the variable of interest y_t and the variable x_t to which we are introducing an exogenous impulse in (1) are included in w_t. The variables in w_t are driven by a $m \times 1$ vector of uncorrelated shocks ϵ_t:

$$w_t = \Theta(L)\epsilon_t, \qquad \epsilon_t \sim iid(0, I_m), \tag{5}$$

with lag polynomial $\Theta(L) \equiv \sum_{j=0}^{\infty} \Theta_j L^j$, and $n \times m$ coefficient matrices Θ_j.

Suppose, without loss of generality, that we are interested in the effect of the first shock ϵ_{1t}, and y_t and x_t are respectively the i^{th} and j^{th} variable in the vector w_t. Therefore, the IRF of y_{t+h} to a one unit increase in ϵ_{1t} is θ_h^{i1}, where θ_h^{kl} is the (k, l) element of Θ_h. Since a local projection is cast by normalizing the effect of the shock on a one unit exogenous increase in x_t, the IRF of y_{t+h} with respect to an impulse to x_t due to ϵ_{1t} is $\theta_h^{i1}/\theta_h^{j1}$. Connecting back to (1), $\theta_h^{i1}/\theta_h^{j1}$ is the object of interest that β_h in the local projection should identify.

If one possesses the series of ϵ_{1t}, one can consistently estimate β_h as a direct regression of the shock ϵ_{1t} on y_{t+h} without any controls.[2] In this setting, the inclusion of controls are to the extent that they may lead to efficiency in finite samples. Therefore, our proposed RSLP can only be used to exploit possible efficiency gains in this setting. Since possessing an observed shock series is less common, we consider the role of controls in settings encountered in most applied work: Local projections implementing "SVAR identification" or the use of an IV.

2.2.1 Implementing SVAR identification

Application of SVAR identification schemes (i.e. short-run restrictions, long-run restrictions, sign restrictions etc.) in local projections relies on satisfying SVAR invertibility. Invertibility implies one can map a VAR in w_t to the form in (5) by inverting the VAR lag polynomials. The literature has also long recognized the link between invertibility and the inclusion of all relevant information (see, e.g. Hansen and Sargent, 1991; Fernández-Villaverde et al., 2007; Stock and Watson, 2018). The equivalent SVAR representation of the local projection would be one where the control set of the local projection, W_t, includes all lags of w_t and hence the information from all past shocks. Therefore, under invertibility, the control set for the local projection satisfies:

$$\text{Proj}\left(y_{t+h} \mid x_t, W_t\right) = \text{Proj}\left(y_{t+h} \mid x_t, W_t, \{\epsilon_\tau\}_{-\infty\leq\tau\leq t}\right). \tag{6}$$

Suppose a researcher aims to estimate (2) using SVAR identification. To fulfil the condition implied by (6), short of magically correctly choosing all the lags of w_t as implied by the unknown DGP as a trivial special case, both V_t and G_t have to span the information implied by all past shocks. If not all relevant information is included, identification fails due to not satisfying invertibility. The application of random subspace methods to deal with the control set may therefore aid with satisfying invertibility. For context, factor models such as those by Bernanke et al. (2005) and Forni and Gambetti (2014), while in a VAR setting, were designed to include all the relevant information in order to fulfill the invertibility condition. While we propose to use random subspace methods within local projections, our motivation is the same.

[2] Note that one may also need to regress ϵ_{1t} on x_t to scale β_h if one did not treat x_t as ϵ_{1t}.

2.2.2 Implementing IV identification

Consider again ϵ_{1t} in (5) as the unobserved shock of interest, for which we aim to estimate the IRF. For simplicity, we normalize ϵ_{1t} to imply a unit increase in x_t.[3] Recall that x_t is an element of the vector w_t, and x_t can thus be written as a linear combination of shocks:

$$x_t = \epsilon_{1t} + f(\epsilon_{2:n_\epsilon,t}, \epsilon_{t-1}, \epsilon_{t-2}, ...), \tag{7}$$

where $\epsilon_{2:n_\epsilon,t} \equiv (\epsilon_{2t}, ..., \epsilon_{n_\epsilon t})'$ and $f(.)$ is a linear combination of its argument. Leading the i^{th} equation of (5), which describes the dynamics of y_t, by h periods, and subsequently rearranging (7) to substitute out for ϵ_{1t}, we obtain

$$y_{t+h} = \theta_h^{i1} x_t + f(\epsilon_{t+h}..., \epsilon_{t+1}, \epsilon_{2:n_\epsilon,t}, \epsilon_{t-1}, \epsilon_{t-2}, ...), \tag{8}$$

where the argument in $f(\cdot)$ now subsumes y_{t+h} being a linear function of past and future shocks and $f(\epsilon_{2:n_\epsilon,t}, \epsilon_{t-1}, \epsilon_{t-2}, ...)$ from (7).

If one were able to reliably estimate θ_h^{i1}, one would obtain the correct IRF since θ_h^{i1} in (8) would be analogous to β_h in (1). Due to endogeneity, (8) cannot be consistently estimated since both x_t and y_{t+h} are a function of all past and current shocks. More precisely, in our context, x_t is not only correlated with ϵ_{1t} but also with other shocks that affect y_{t+h}.

The IRF can be estimated using two-stage least squares and an instrument z_t that satisfies the following conditions (Stock and Watson, 2018):

1. $\mathbb{E}[\epsilon_{1t} z_t] \neq 0$ (relevance),
2. $\mathbb{E}[\epsilon_{2:n_\epsilon,t} z_t] = 0$ (contemporaneous exogeneity),
3. $\mathbb{E}[\epsilon_{t+j} z_t] = 0$, $j \neq 0$ (lead-lag exogeneity).

Most instruments used in a macro setting are constructed to isolate the shock in question contemporaneously.[4] We therefore see it as less empirically relevant how controls are used to fulfill the first two conditions and focus on lead-lag exogeneity. The third condition demands the instrument to be uncorrelated with past and future shocks, which can be seen from (8) as

[3] This normalization is without any loss of generality since shocks are unobserved, so the variance and sign of shocks are ultimately normalized. Our choice of normalization simplifies the exposition from needing to introduce a parameter to link ϵ_{1t} to x_t.

[4] For example, Kilian (2008) proposes an instrument for oil supply shocks based on foreign oil production shortfalls during events such as wars or civil disturbance. Gertler and Karadi (2015) propose an instrument for monetary policy shocks constructed from high frequency surprises around policy announcements.

both x_t and y_{t+h} are a function of all past and future shocks. This is a strong assumption but can be relaxed by including controls:

$$\mathbb{E}\left[\epsilon_{t+j}^{\perp} z_t^{\perp}\right] = 0,\ j \neq 0\,\text{(conditional lead-lag exogeneity)}, \tag{9}$$

where $u_t^{\perp} = u_t - Proj(u_t \mid W_t)$ for some variable u.

We elaborate on the role controls may play in the conditional lead-lag exogeneity. The instruments proposed by Romer and Romer (2004) and Gertler and Karadi (2015) have been shown to be possibly forecastable by other macro variables (see Ramey, 2016). By the logic of (5), this implies that the instruments are correlated with past shocks and lead-lag exogeneity is violated. However, the instrument may still be valid if the conditional lead-lag exogeneity is fulfilled if controls account for the information contained in all past shocks.

We thus modify random subspace methods to the two-stage least squares settings where the first and second stage regressions are

$$x_t = \alpha^{(j)} + \rho^{(j)} z_t + \Lambda^{(j)} V_t + \Upsilon^{(j)} R^{(j)} G_t + \eta_t^{(j)}, \tag{10}$$

$$y_{t+h} = \mu_h^{(j)} + \beta_h^{(j)} \hat{x}_t^{(j)} + \Theta_h^{(j)} V_t + \Psi_h^{(j)} R^{(j)} G_t + \xi_{t+h}^{(j)}, \tag{11}$$

where $\Lambda^{(j)}$ and $\Upsilon^{(j)}$ are the projection coefficients from the first stage regression, $\hat{x}_t$ is the fitted value from (10), and (11) is now analogous to our original RSLP from (4) in which we similarly average over $\beta_h^{(j)}$'s corresponding to random draws for selection matrices.

Depending on whether the instrument satisfies (conditional) lead-lag exogeneity, V_t and G_t can be empty sets. Recall that in addition to identification, controls may also help with reducing the variance in the estimated IRF. Hence, the practitioner facing a large set of controls who is unsure what to include, can also resort to RSLP with IV identification.

3 Monte Carlo Experiments

To understand our RSLP procedure, we consider Monte Carlo experiments based on a Real Business Cycle (RBC) model with fiscal foresight, as discussed in Leeper et al. (2013). Fiscal foresight is a setting in where economic agents know of a future tax shock, but this is not reflected in the information set used by the econometrician to estimate IRFs which then leads to erroneous estimation of the IRFs.

3.1 Data generating process

The model includes income taxes, inelastic labor supply, and full capital depreciation. The log-linearized equilibrium condition for capital is given by[5]

$$k_t = \alpha k_{t-1} + u_{at} - (1-\theta)\frac{\tau}{1-\tau}\sum_{k=0}^{\infty}\theta^k E_t \hat{\tau}_{t+k+1}, \tag{12}$$

where k_t and $\hat{\tau}_t$ are the percentage deviations from the steady state capital and tax rate respectively, τ is the steady state value of the tax rate, θ and α are the discount factor and capital share in a Cobb-Douglas production function satisfying the inequalities $0 < \theta < 1$ and $0 < \alpha < 1$, and u_{at} is an independent identically distributed (i.i.d) technology shock. The tax rule is $\hat{\tau}_{t+h} = u_{\tau t}$, where $u_{\tau t}$ is an i.i.d tax shock. The phenomenon of fiscal foresight occurs when at time t agents know the tax rate they will face at time $t+h$.

Allowing $h = 2$ (a two-period foresight), (12) becomes

$$k_t = \alpha k_{t-1} + u_{at} - \kappa(\theta u_{\tau t} + u_{\tau t-1}), \tag{13}$$

where $\kappa = \tau(1-\theta)/(1-\tau)$. The structural moving average representation equals

$$\begin{bmatrix} \hat{\tau}_t \\ k_t \end{bmatrix} = \begin{bmatrix} L^2 & 0 \\ -\frac{\kappa(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L} \end{bmatrix} \begin{bmatrix} u_{\tau t} \\ u_{at} \end{bmatrix}. \tag{14}$$

To parametrize the above, we set $\theta = 0.2637$, $\tau = 0.25$, $\alpha = 0.36$, and $u_{\tau t}$ and u_{at} have independent standard Normal distributions.

Due to fiscal foresight, there is insufficient information to recover the IRFs by just observing both capital and the tax rate at time t. More precisely, the missing piece of information is the tax rate at $t+2$ (or the current tax shock). The Monte Carlo experiments assume that the econometrician observes a vector of 100 informational variables which mimic an empirically relevant setting where the econometrician has access to a data set such as FRED-MD or FRED-QD. The individual information series, y^*_{it}, are generated as:

$$y^*_{it} = b_i u_{\tau t} + (1-b_i)u_{at} + \xi_{it}, \quad \xi_{it} \sim N(0,\sigma_i^2), \quad i = 1,\ldots,100, \tag{15}$$

where we draw b_i as a Bernoulli random variable assuming value 1 with probability 0.1 and value 0 with probability 0.9.

To examine the role of the strength of the factor structure of the data in accurately recovering the IRFs, we consider two settings that differentiate the factor structure of the informational

[5] See Leeper et al. (2013), page 1118, Equation (4).

variables:

$$\text{Strong Case:} \quad \sigma_i \sim U(0,1), \quad \text{Weak Case:} \quad \sigma_i \sim U(0,4). \tag{16}$$

In expectation, lower values of σ_i are associated with less variation in the idiosyncratic component ξ_{it}, and so the informational variable provides a stronger signal of the relevant information. We have parameterized the strong and weak cases in order to provide some guidance in situations an applied macroeconomist may encounter. In the strong case, the two factors explain about 80% of the variation in expectation, while this falls to slightly over 20% in the weak case.

The strong case mimics a structure that one would probably encounter in a panel where these individual series have a high level of comovement, such as cross country asset prices and interest rates (e.g. Miranda-Agrippino and Rey, 2020), or commodity prices (e.g. West and Wong, 2014; Alquist et al., 2020). On the other hand, the weak case is akin to the sort one would expect to encounter when using U.S. macroeconomic data, such as the FRED-MD data set. Therefore, we stress that we view the weak factor not as an implausible case, but only to the extent that it is weaker than the strong case. Instead, the weak case is a relevant and plausible case, which we show in Appendix A: The first two factors of both the FRED-MD and FRED-QD data set explain a similar amount of variation as in the weak case.

3.2 Structural identification on the simulated data

While there is insufficient information to recover the true IRFs, the econometrician will use the informational variables as controls in order to estimate the IRFs. We revisit the two settings discussed in Section 2.2: implementing SVAR and IV identification.

Setting 1: Implementing SVAR identification

The missing piece of information is the tax rate two periods ahead, which is contained in $u_{\tau t}$. Therefore, restoration of the information contained in $u_{\tau t}$ will render the model informational sufficient. As seen from (14), the technology shock has no cumulative impact on the tax rate. Hence, conditional on being able to recover the reduced form forecast errors, one can implement SVAR identification.[6] This is the case considered by Forni and Gambetti (2014), except that we consider a local projection rather than an SVAR. If the random subspace methods can

[6] We implement SVAR identification by using the procedure presented by Plagborg-Møller and Wolf (2021). We leave the implementation details to Appendix B.

adequately control for the missing information, the SVAR identification can be implemented, and one should be able to recover the IRFs.

Setting 2: Implementing IV identification

In the second setting, the researcher possesses an instrument generated as follows:

$$Z_t = 0.7u_{\tau t} + u_{at-1} + u_{\tau t-1} + \epsilon_{Zt}, \quad \epsilon_{Zt} \sim N(0, 0.01), \tag{17}$$

which is meant to instrument for the tax shock $u_{\tau t}$, but the presence of lagged shocks u_{at-1} and $u_{\tau t-1}$ render the instrument invalid as it violates lead-lag exogeneity. If random subspace methods can adequately control for the information contained in the lagged shocks, the instrument is conditional exogenous, and one should be able to recover the IRF.

3.3 Simulation settings

Each Monte Carlo replication simulates an artificial data set of 200 observations for capital and the tax rate according to (14), together with 100 informational series in either the strong or weak case according to (15) and (16). For the IV setting, we also generate the instrument according to (17). The number of Monte Carlo replications is 2000.

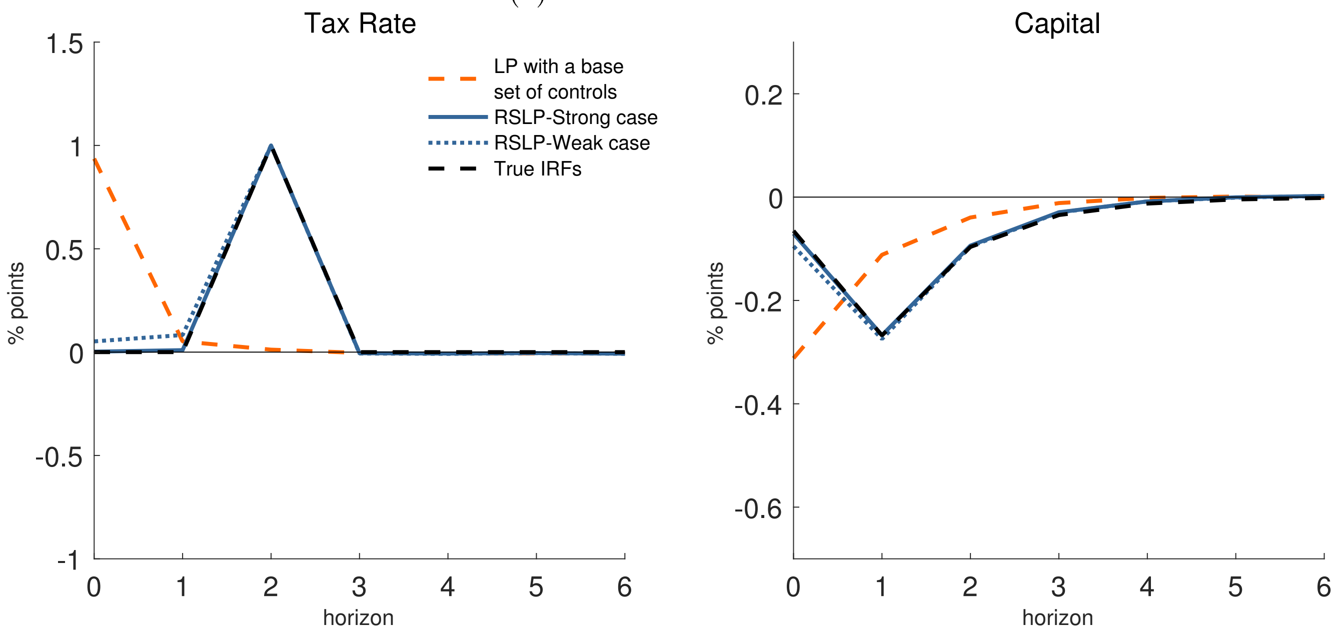
The subspace dimension for RSLP is set equal to 50. Appendix C elaborates on this choice of subspace dimension. For each artificial data set, we then use the procedure in Section 2.1 with 1000 draws of the selection matrix $R^{(j)}$ and we set G_t equal to the first lag of y_t^*. We also estimate local projections (LPs) without G_t, which we label LP with a base set of controls, to investigate whether RSLP is a valid approach to accounting for the omitted information. Both specifications include the two lags of tax rate and capital in V_t, together with two lags of the instrument when using IV identification. Appendix B elaborates on the specification of the local projections when using SVAR identification.

3.4 RSLP can recover the true impulse response functions

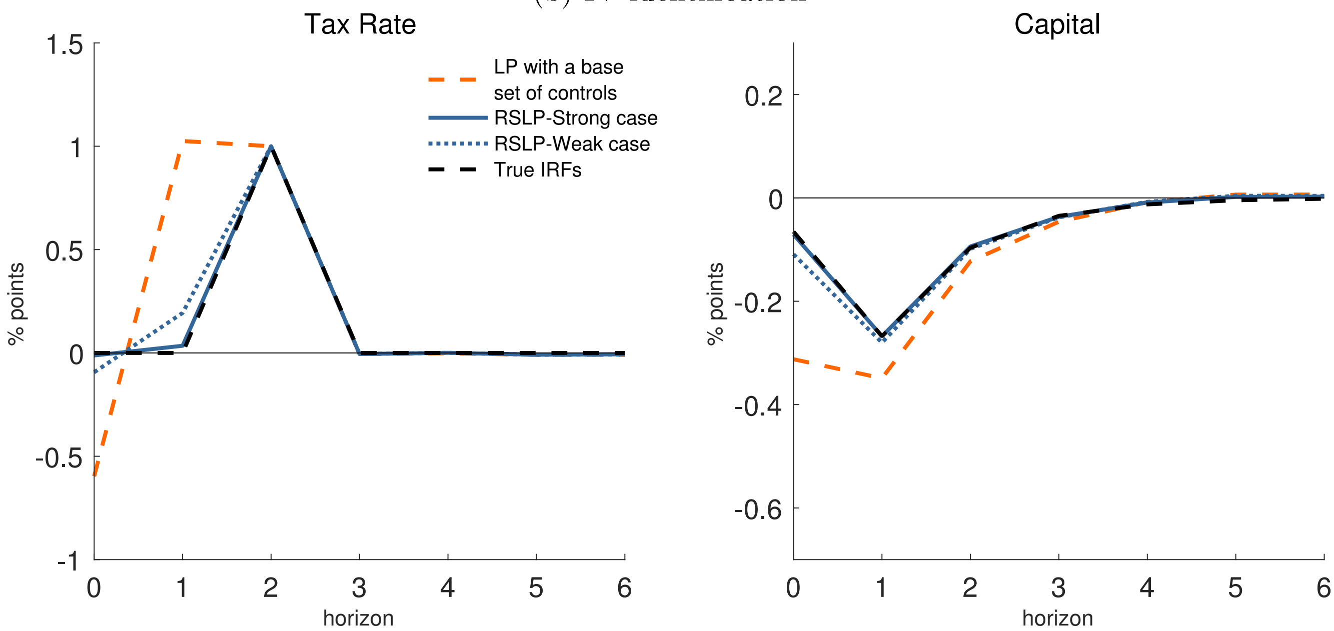
To understand if random subspace methods are able to recover the true IRFs, we study the expectation of the estimated IRFs. Figure 1 presents the IRFs of both capital and tax rate to a tax shock from both the SVAR and IV identification, under both the strong and weak case. To make the role of the informational variables explicit, we compare relative to IRFs estimated

Figure 1: Impulse response functions in the Monte Carlo experiments

(a) SVAR identification



(b) IV identification



Note: This figure shows impulse response functions estimated by LP with a base set of controls (orange, dot-dashed), RSLP using controls with a strong factor structure (blue, solid), RSLP using controls with a weak factor structure (blue, dotted), and the true responses (black, dashed). The upper and lower panels use SVAR and IV identification respectively. The shock is normalized to a 100 basis point increase in the tax rate. The results are the average impulse response functions across all 2000 Monte Carlo replications.

via local projection without the additional controls from the generated informational series. We also plot the true IRFs. The plotted estimated IRFs are taken as the average across all the Monte Carlo simulations, so deviations from the true IRF represent bias.

Since the tax shock occurs in period 2, the true IRF of the tax rate sees a one off spike in period $t+2$. Capital falls for two periods, then adjusts back in a monotonic fashion. Suppose we estimated the IRF with only information on the time series of the tax rate and capital. The dot-dashed orange line shows that we will miss the timing of the tax rate, and as a result, also miss the response of capital. This is an illustration of a fiscal foresight issue, where information on only the tax rate and capital is insufficient information to recover the true IRF which has been previously documented (e.g. Leeper et al., 2013; Forni and Gambetti, 2014). With IV identification, LP with a base set of controls is also biased given the instrument is invalid due to not being exogenous to the lagged shocks.

Our proposed approach is able to largely recover the true IRF. In particular, RSLP is able to estimate the tax rate spiking in period 2, and recover the dynamics of the response to capital. While our proposed method is unbiased in the strong case, the bias is still minimal under the weak case with either SVAR or IV identification. Appendix C explores the size of the subspace dimension in our Monte Carlo experiments, and finds an optimal subspace dimension within the 40-60 range in weak cases, and often even smaller in the strong cases.

3.5 Understanding differences of RSLP relative to FAVAR

FAVAR models are a common alternative to control for a large information set within the applied macroeconomics literature (e.g., see Bernanke et al., 2005; Stock and Watson, 2005; Forni et al., 2009; Forni and Gambetti, 2014). To understand how our approach fares relative to FAVAR, we compare these methods in the Monte Carlo experiments. The FAVAR augments a bivariate VAR system containing taxes and capital with the first two principal components of the informational variables. The SVAR identification case identifies the shock by assuming there is no cumulative effect of the technology shock on the tax shock. The IV identification case uses the instrument as per the proxy-VAR approach by Mertens and Ravn (2013) and Gertler and Karadi (2015).

Before we present our results, we first caveat that one should practice some caution in comparing RSLPs to FAVARs. IRFs from FAVARs are based on iterated forecasts, whereras local projections produce direct forecasts. We already know that iterative forecasts are more efficient, albeit more biased than direct forecast under misspecification.

Table 1 reports the root mean squared error (RMSE) of the FAVAR relative to our RSLP

Table 1: RMSE and absolute bias of FAVAR relative to RSLP in the Monte Carlo experiments

	SVAR identification				IV identification			
	Strong		Weak		Strong		Weak	
	Tax Rate	Capital	Tax Rate	Capital	Tax Rate	Capital	Tax Rate	Capital
RMSE	0.896	0.989	2.350	1.256	0.868	0.780	2.496	1.454
Absolute bias	0.872	0.375	1.795	3.485	0.709	0.873	1.256	2.301

Note: This table reports the root mean squared error (RMSE) and the absolute bias of FAVAR relative to RSLP for the SVAR and IV identification under both the strong and weak factor structure settings. Appendix D shows the average impulse response functions across all 2000 Monte Carlo replications for both RSLP and FAVAR.

approach.[7] That is, we compare the FAVAR relative to the RSLP in an identical setting where they both use the same identification strategy and observe the same informational variables. Values above one favour RSLP. In the strong factor case, the FAVAR models have a lower RMSE relative to our RSLP approach. However, the conclusions flip when we consider the weak case, where the RSLP approach outperforms the FAVAR, at times by substantial margins.

In order to contextualize why RSLP performs better with a weakened factor structure, Table 1 also reports the absolute bias of FAVAR relative to RSLP.[8] When the factor structure is strong, FAVAR and RSLP have a similar bias.[9] Therefore, it is no surprise that FAVAR does better from an RMSE perspective when the factor structure is strong: Both methods are close to being unbiased and the relative efficiency gains from the iterated forecast dominates. However, when we weaken the factor structure, the bias in FAVARs increases substantially to the extent that the less biased RSLP dominates.

We conclude that in situations where the factor structure is strong, such as cross-country data on asset prices and interest rates, factor models such as FAVARs do well. With a weak factor structure, resembling that of a typical macroeconomic data set such as U.S. macroeconomic data like FRED-MD, RSLP can be a valuable tool in the econometrician's toolkit.

[7] Note that there are seven horizons ($h = 0, \dots, 6$) in the IRFs, across which we average to calculate the RMSE: $\sqrt{\frac{1}{7}\sum_{h=0}^{6}\frac{1}{2000}\sum_{i=1}^{2000}(\hat{\beta}_{hi} - \beta_{h,\text{true}})^2}$, where $\hat{\beta}_{hi}$ is the estimated IRF at horizon h in the i-th replication and $\beta_{h,\text{true}}$ is the true IRF at horizon h.

[8] We calculate the absolute bias across seven horizons: $\frac{1}{7}\sum_{h=0}^{6}|\bar{\hat{\beta}}_h - \beta_{h,\text{true}}|$, where $\bar{\hat{\beta}}_h = \frac{1}{2000}\sum_{i=1}^{2000}\hat{\beta}_{hi}$ is the average IRF at horizon h across 2000 replications.

[9] Note that Figure 1 shows that the RSLP approach is close to being unbiased in all the considered cases. Hence, a relative bias of 10% to even 30% is not a meaningful difference in a practical sense.

3.6 Take-aways

We summarize two key take-aways from our Monte Carlo exercise. First, the results indicate that RSLP is capable of recovering the true IRF and provides at least *prima facie* evidence that it is a viable approach for applied work. In particular, RSLP is close to being unbiased both under weak or strong factor structures. Second, relative to a common alternative, namely the FAVAR model, RSLP may be an appropriate addition to the standard toolkit given its smaller bias can overcome its relative inefficiency, especially when the factor structure mimics that of U.S. macroeconomic data.

4 Empirical applications

We use RSLP to estimate the dynamic responses to technology and monetary policy shocks. Both applications can be traced to a broad empirical literature, and some of this literature suggests that baseline specifications with minimal controls are insufficient. Hence, these applications are natural settings to understand whether random subspace methods are a useful method to incorporate information from a large set of controls.

4.1 Technology shock application

Since Gali (1999), a strand of the SVAR literature has investigated the impact of technology shocks on a variety of macroeconomic variables, with a focus on labor market variables (e.g., Francis and Ramey, 2005; Forni and Gambetti, 2014; Barnichon, 2010). Keeping in the spirit of Gali (1999), the technology shock is identified as being the only shock that has a long-run impact on labor productivity. Since we work with local projections instead of SVARs, like the aforementioned papers, we implement the long-run identification restrictions as suggested by Plagborg-Møller and Wolf (2021).[10]

In the spirit of SVAR work which estimates the effect of technology shocks, our baseline local projection for this application will only consider the lags of two variables; labor productivity and unemployment. Owing to the model being specified at the quarterly frequency, we include four lags of both the growth rate of labor productivity and unemployment in V_t.

The set of possible controls in G_t includes 127 macroeconomic time series specified at the

[10]Implementation requires one to nominate a horizon at which all short-run effects of the shock are expected to dissipate. We set this horizon to 3 years. Appendix B discusses further implementation details.

quarterly frequency, which includes 117 series from the FRED-QD database (see McCracken and Ng, 2020), 6 total factor productivity series from Fernald's website[11], and 4 consumer confidence indicators from the Michigan Survey. We consider the first lags for these variables, which gives us a set of 127 possible control variables. Our sample is from 1960Q1 to 2019Q4.

4.2 Monetary policy shock application

This application features a baseline set of controls that echoes the Proxy VAR of Gertler and Karadi (2015). As the model is monthly, our baseline set of controls includes 12 lags of the log difference of CPI (or approximately quarter-on-quarter inflation), log difference of industrial production (IP), the Excess Bond Premium (see Gilchrist and Zakrajšek, 2012), and the 1-year government bond rate in V_t in (11).

Our possible control set G_t includes 111 FRED-MD series, and we consider the first lag of these variables. The sample spans January 1990 to June 2012 to match up with the time-span of the instrument. We use the three-month ahead funds rate surprise by Gertler and Karadi (2015) as an instrument for monetary policy shocks.

4.3 Results

Figure 2 presents the estimated IRFs from our two applications.[12] The RSLPs are based on 1000 draws of the selection matrix, a subspace dimension equal to 50, and accompanied by a one standard deviation interval constructed as in Appendix E. Appendix C investigates the robustness of the results with respect to the subspace dimension. In order to appreciate the effect of the additional controls in RSLP, we compare with a local projection with only the base set of controls (i.e. omitting G_t in (2)). We also compare relative to the FAVAR since this is a common alternative within the broader literature.

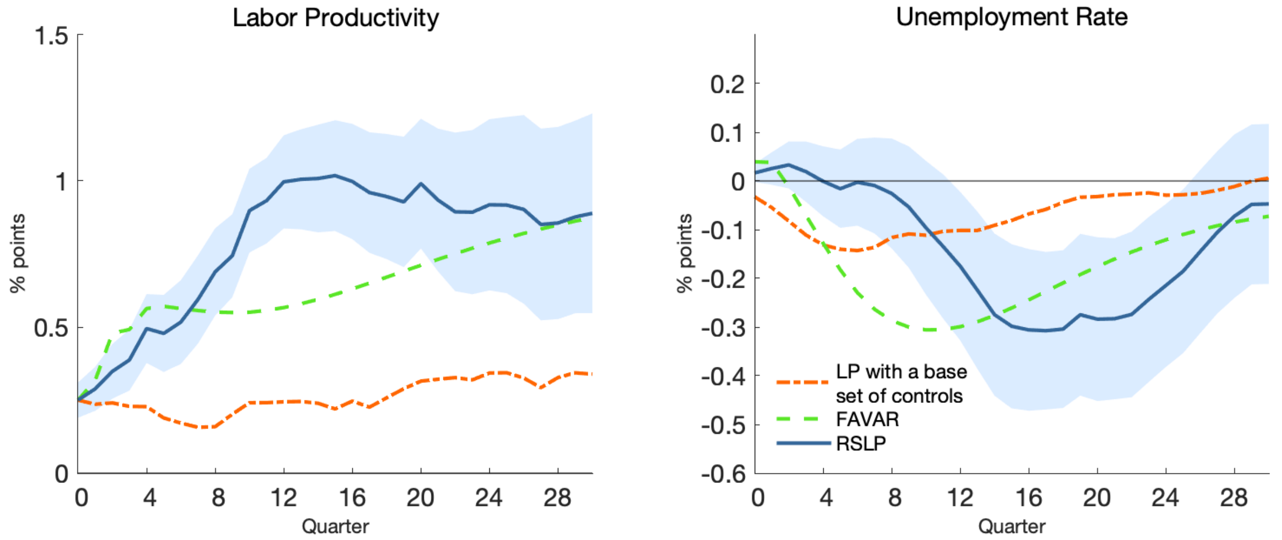
For the technology shock application, we present the response of labor productivity and the unemployment rate to a technology shock which raises labour productivity by 0.25%, which is equivalent to a one standard deviation shock in the FAVAR model. Using RSLP, labor productivity increases and the unemployment rate decreases in response to an expansionary

[11] John Fernald's website: https://www.johnfernald.net/TFP

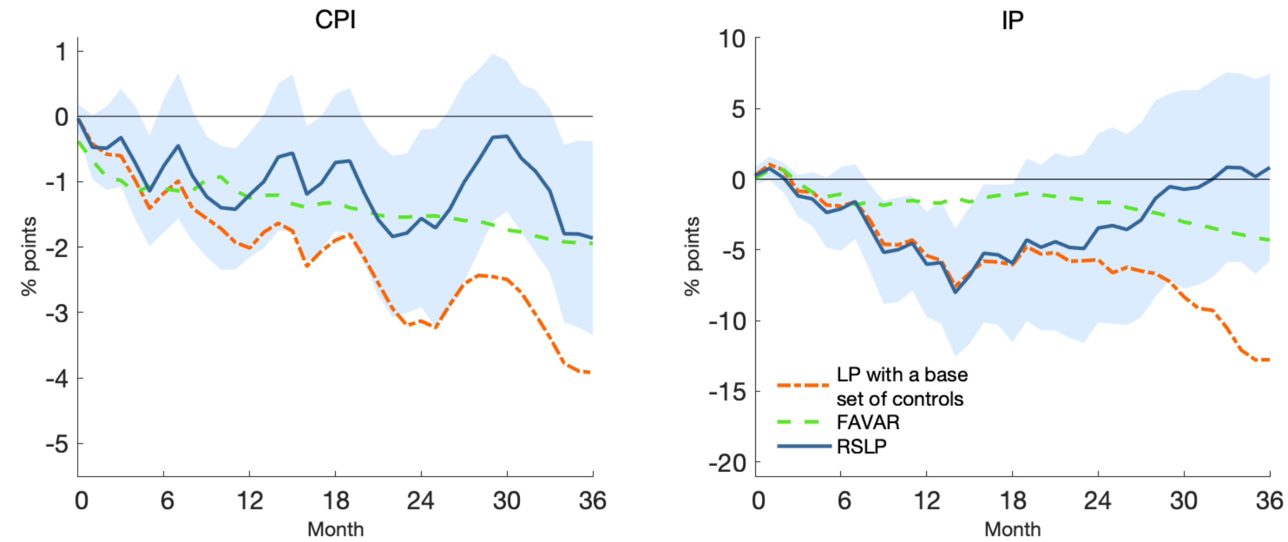
[12] Note that the IRFs are on the level of labor productivity, CPI and industrial production index. Hence, we follow the usual practice to specify the left-hand side variable in the local projection as $y_{t+h} - y_{t-1}$ (see Stock and Watson, 2018, Section 1.5).

Figure 2: Impulse response functions in the empirical applications

(a) Response to a technology shock



(b) Response to a monetary policy shock



Note: This figure shows the estimated local projections with only a base set of controls (orange, dot-dashed), FAVAR (green, dashed), and RSLP (blue, solid). The upper panel shows the impulse responses of labor productivity and unemployment rate to a technology shock that increases labor productivity by 0.25%. The lower panel shows the impulse responses of CPI and IP to a monetary policy shock that increases GS1 by a 100 basis points. The blue-shaded areas indicate the one standard deviation intervals of RSLP.

technology shock. These results are consistent with the predictions of the real business cycle model and those shown by, for example, Christiano et al. (2003).

For the monetary policy shock application, we present the response of CPI and industrial production to a contractionary monetary policy shock which increases the 1-year government bond rate by a 100 basis points. Using RSLP, both CPI and industrial production fall in response to a contractionary monetary policy shock, consistent with what one expects from a standard monetary policy shock (Bernanke et al., 2005; Gertler and Karadi, 2015).

Estimating local projections with just a baseline set of controls and the FAVAR produce IRFs that imply the same qualitative impact as RSLP, but there are large quantitative differences. For example, allowing for a set of additional controls using RSLP implies an unemployment response whose trough is twice that of not allowing for the additional controls. At the same time, CPI inflation and industrial production fall by more in response to a monetary policy shock if only a base set of controls is included. There are also marked differences relative to FAVAR, especially in the labor productivity responses to a technology shock, and in the dynamics of a monetary policy shock to industrial production.

The one standard deviation intervals suggest that at least some of the differences between RSLP and LP with a only a base set of controls are meaningful, implying that the controls *do* matter. Using our Monte Carlo experiments in Section 3 as a guide to interpreting our results, the empirical findings suggest that failure to allow for a larger set of controls can result in estimated IRFs that are statistically different, and potentially misleading. Some of the differences relative to the FAVAR may also be statistically significant. This is in line with the discussion in Section 3.5, which suggests that these differences may stem from how random subspace and factor methods control for missing information, with the possibility that random subspace methods are more suitable given the type of macroeconomic data.

In sum, we list two key take-aways from our empirical applications. First, RSLP can be a useful approach to estimating IRFs, given we are able to obtain reasonable estimates from both empirical exercises. Second, the issue of controls in local projections may require more scrutiny, as the baseline local projections that we estimate are quite representative of what a typical macroeconomist would use as a base specification. If the additional controls in the RSLP do not contain any useful information, one would not expect any differences between RSLP and local projections with a baseline set of controls. The fact we do find these differences, a finding consistent with our Monte Carlo exercise, suggests that the baseline local projections in our empirical exercise omit relevant information.

5 Conclusion

We show how one can apply a dimension reduction technique traditionally used in machine learning to local projections in order to estimate IRFs with many controls. This random subspace method is simple to implement and it basically contains 3 steps: Step 1: take a random draw; Step 2: do it many times; and Step 3: take averaging. We have shown that our approach is a plausible addition to the toolkit as it can recover the true IRFs in settings encountered in macroeconomics. In addition, we present suggestive evidence that in factor structures of U.S. macroeconomics data, such as FRED-MD, our method is perhaps more appropriate than the commonly used FAVAR.

It is worth stressing that while our proposed approach is a plausible empirical strategy, it does not compete to supplant any of the recent innovations in the broader development of local projection estimation. To highlight two important developments in the local projection literature, Barnichon and Brownlees (2019) consider smoothing local projection, and Ferreira et al. (2023) combine local projections with additional prior information. Both approaches can be applied to random subspace local projections instead of standard local projections with a base set of controls, so providing a more appropriate starting point for their procedures.

We also note that while we show that our method can outperform FAVAR in some empirical plausible settings, we do not argue that one has to choose between either factor or random subspace methods. For instance, there is nothing to stop one from using factors in conjunction with subspace methods. One possibility is the application of subspace methods to a set of estimated factors instead of the original set of controls. Therefore, one should not necessarily view our work as a replacement for existing methods. Instead, we are keen to stress the potential for future work to combine our insights with existing developments in the local projection literature in attempts to further improve the properties of these local projection estimators.

References

Alquist, R., Bhattarai, S., and Coibion, O. (2020). Commodity-Price Comovement and Global Economic Activity. *Journal of Monetary Economics*, 112:41–56.

Barnichon, R. (2010). Productivity and Unemployment Over the Business Cycle. *Journal of Monetary Economics*, 57(8):1013–1025.

Barnichon, R. and Brownlees, C. (2019). Impulse Response Estimation by Smooth Local Projections. *The Review of Economics and Statistics*, 101(3):522–530.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *The Quarterly journal of economics*, 120(1):387–422.

Blanchard, O. J. and Quah, D. (1988). The Dynamic Effects of Aggregate Demand and Supply Disturbances. Technical report, National Bureau of Economic Research.

Boot, T. and Nibbering, D. (2019). Forecasting Using Random Subspace Methods. *Journal of Econometrics*, 209(2):391–406.

Boot, T. and Nibbering, D. (2020). Subspace Methods. In *Macroeconomic Forecasting in the Era of Big Data*, pages 267–291. Springer.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). Model Selection: An Integral Part of Inference. *Biometrics*, pages 603–618.

Christiano, L. J., Eichenbaum, M., and Vigfusson, R. (2003). What Happens After a Technology Shock? Technical report, National Bureau of Economic Research.

Chudik, A., Kapetanios, G., and Pesaran, M. H. (2018). A One Covariate at a Time, Multiple Testing Approach to Variable Selection in High-Dimensional Linear Regression Models. *Econometrica*, 86(4):1479–1512.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J., and Watson, M. W. (2007). ABCs (and Ds) of Understanding VARs. *American economic review*, 97(3):1021–1026.

Ferreira, L. N., Miranda-Agrippino, S., and Ricco, G. (2023). Bayesian Local Projections. *The Review of Economics and Statistics*, pages 1–45.

Forni, M. and Gambetti, L. (2014). Sufficient Information in Structural VARs. *Journal of Monetary Economics*, 66:124–136.

Forni, M., Giannone, D., Lippi, M., and Reichlin, L. (2009). Opening the Black Box: Structural Factor Models with Large Cross Sections. *Econometric Theory*, 25(5):1319–1347.

Francis, N. and Ramey, V. A. (2005). Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited. *Journal of Monetary Economics*, 52(8):1379–1399.

Gali, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American economic review*, 89(1):249–271.

Gertler, M. and Karadi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Gilchrist, S. and Zakrajšek, E. (2012). Credit Spreads and Business Cycle Fluctuations. *American economic review*, 102(4):1692–1720.

Hansen, L. P. and Sargent, T. J. (1991). Two Difficulties in Interpreting Vector Autoregressions. In *Rational expectations econometrics*, pages 77–119. CRC Press.

Ho, P., Lubik, T., and Matthes, C. (2023). Averaging Impulse Responses Using Prediction Pools. Working paper, FRB Richmond Working Paper.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American economic review*, 95(1):161–182.

Kilian, L. (2008). Exogenous Oil Supply Shocks: How Big are They and How Much Do They Matter for The US Economy? *The Review of Economics and Statistics*, 90(2):216–240.

Koop, G., Korobilis, D., and Pettenuzzo, D. (2019). Bayesian Compressed Vector Autoregressions. *Journal of Econometrics*, 210(1):135–154.

Kotchoni, R., Leroux, M., and Stevanovic, D. (2019). Macroeconomic Forecast Accuracy in a Data-Rich Environment. *Journal of Applied Econometrics*, 34(7):1050–1072.

Leeper, E. M., Walker, T. B., and Yang, S.-C. S. (2013). Fiscal Foresight and Information Flows. *Econometrica*, 81(3):1115–1145.

Li, D., Plagborg-Møller, M., and Wolf, C. K. (2022). Local Projections vs. VARs: Lessons from Thousands of DGPs. Technical report, National Bureau of Economic Research.

McCracken, M. and Ng, S. (2020). FRED-QD: A Quarterly Database for Macroeconomic Research. Technical report, National Bureau of Economic Research.

Mertens, K. and Ravn, M. O. (2013). The Dynamic Effects of Personal and Corporate Income Tax Changes in The United States. *American economic review*, 103(4):1212–47.

Miranda-Agrippino, S. and Rey, H. (2020). US Monetary Policy and the Global Financial Cycle. *The Review of Economic Studies*, 87(6):2754–2776.

Pick, A. and Carpay, M. (2022). Multi-Step Forecasting with Large Vector Autoregressions. In *Essays in Honor of M. Hashem Pesaran: Prediction and Macro Modeling*, volume 43, pages 73–98. Emerald Publishing Limited.

Plagborg-Møller, M. and Wolf, C. K. (2021). Local Projections and VARs Estimate the Same Impulse Responses. *Econometrica*, 89(2):955–980.

Ramey, V. A. (2016). Macroeconomic Shocks and Their Propagation. In *Handbook of macroeconomics*, volume 2, pages 71–162. Elsevier.

Ramey, V. A. and Zubairy, S. (2018). Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. *Journal of Political Economy*, 126(2):850–901.

Romer, C. D. and Romer, D. H. (2004). A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review*, 94(4):1055–1084.

Stock, J. H. and Watson, M. W. (2005). Implications of Dynamic Factor Models for VAR Analysis.

Stock, J. H. and Watson, M. W. (2018). Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal*, 128(610):917–948.

Swanson, E. T. (2020). Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets. *Journal of Monetary Economics*.

Tenreyro, S. and Thwaites, G. (2016). Pushing on a String: US Monetary Policy is Less Powerful in Recessions. *American Economic Journal: Macroeconomics*, 8(4):43–74.

West, K. D. and Wong, K.-F. (2014). A Factor Model for Co-movements of Commodity Prices. *Journal of International Money and Finance*, 42:289–309.

ONLINE APPENDIX

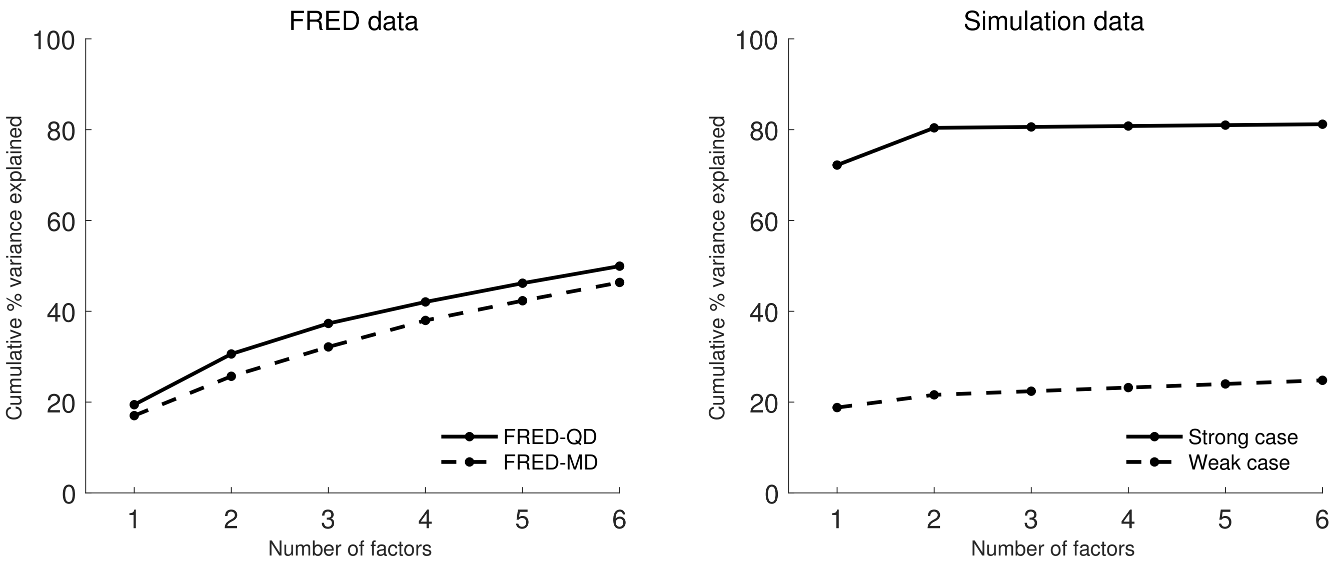
A FRED data structure

Figure A1 shows the factor structure of the FRED-QD and FRED-MD data sets, together with the factor structure of the simulated data in the weak case in Section 3. More precisely, we present the cumulative variance being accounted for as we sequentially add an additional principal component. This is a standard metric for understanding the strength of the factor structure. The figure shows that the first two factors of FRED-QD and FRED-MD account for 25% and 20% of the variation, respectively. Similarly, our simulation data under the weak case also requires two factors to account for 20% of the variation in the controls.

Note that in Figure A1, we calculate the factor structure of the simulated data in population. Let $n_{y^*} = 100$ be the number of informational series, which have mean zero and unit variance. Given the setting in Section 3.1, the expected number of informational series containing the tax shock equals 10, and the expected number of informational series containing the capital shock equals 90. The population covariance matrix of the informational series is therefore equal to $\Sigma_{y^*} = H + \sigma^2 I_{n_{y^*}}$, where H is a 100×100 blockmatrix with a 10×10 and 90×90 blockmatrix with all elements equal to 1 on its diagonal, $\sigma^2 = 0.25$ in the strong case and $\sigma^2 = 4$ in the weak case, and $I_{n_{y^*}}$ is a 100×100 identity matrix.

Then the factor structure of the data set can be calculated by solving the eigenvalue-

Figure A1: Factor structure



Note: This figure shows the factor structure of FRED-QD, FRED-MD, and our simulated data in the weak case.

eigenvector decomposition of the correlation matrix ρ_{y^*} of the covariance matrix Σ_{y^*}: $\rho_{y^*}\nu = \Lambda\nu$, where ν is a matrix whose columns are the eigenvectors and Λ is a diagonal matrix with the eigenvalues λ_j on its diagonal. The cumulative variance explained by the first i principal components can be calculated as $\frac{1}{100}\sum_{j=1}^{i}\lambda_j$, where λ_j is the j^{th} largest eigenvalue of λ.

B Implementing SVAR identification in RSLP

B.1 Monte Carlo experiments

Equation (14) shows that all the variation in the tax rate is from the tax shock (and not the technology shock). Hence, the tax shock can be identified by assuming that the cumulative impact of the tax rate on a technology shock is zero. In practice, this approach identifies the shocks using the long-run restrictions by Blanchard and Quah (1988).

However, there is an important difference between the SVAR identification we implement here and how long-run restrictions are usually implemented. In long-run restrictions, one has an I(1) variable in which all the variation is driven by the permanent shock. Since the I(1) variable enters the VAR or the local projection in differenced form, this implies a restriction where the cumulative impact on the *changes* from the transitory shocks to the I(1) variable is equivalent to zero, and thus has zero long-run impact on the level. In contrast, there is no long-run in our setting since neither the tax rate nor capital is I(1), but we can still identify the technology shock since the long-run cumulative impact of this shock on the tax rate (whether differenced or in levels) is zero.

The estimation of the cumulative identification in a local projection setting is shown by Plagborg-Møller and Wolf (2021). This method requires an I(1) tax rate variable which can be constructed by accumulating our I(0) tax rate variable in Equation (14):

$$\hat{\tau}_1^{I(1)} = \hat{\tau}_1, \qquad \hat{\tau}_t^{I(1)} = \hat{\tau}_t + \hat{\tau}_{t-1}^{I(1)}, \quad \text{for } t = 2, 3, ...T. \tag{A1}$$

Define $\upsilon_t = \hat{\tau}_{t+2}^{I(1)} - \hat{\tau}_{t-1}^{I(1)}$. The impulse response functions can now be estimated by two-stage least squares in the first and second stage regressions

$$\upsilon_t = \alpha^{(j)} + \Lambda^{(j)}V_t^1 + \Upsilon^{(j)}R^{(j)}G_t^1 + \eta_t^{(j)}, \tag{A2}$$

$$y_{t+h} = \mu_h^{(j)} + \beta_h^{(j)}\hat{\upsilon}_t^{(j)} + \Theta_h^{(j)}V_t^2 + \Psi_h^{(j)}R^{(j)}G_t^2 + \xi_{t+h}^{(j)}, \tag{A3}$$

where V_t^1 contains the contemporaneous values of the tax rate and capital, G_t^1 contains contemporaneous values of the informational series, V_t^2 consists of the two lags of tax rate and capital, and G_t^2 consists of the first lag of the informational series. The first stage uncovers the linear combinations of the data that explain the two-period ahead movement of the tax rate. Therefore the fitted values of the first stage recover the tax shock. Once the tax shock is recovered, one can estimate the impulse responses from the second stage.

B.2 Empirical application with a technology shock

The procedure outlined above also applies to our empirical application to a technology shock, with minor adjustments. First, υ_t now equals the long-run movement of labor productivity in level: $\upsilon_t = \text{labor}_{t+12}^{\text{level}} - \text{labor}_{t-1}^{\text{level}}$. Second, the controls in both stages are now as follows: The first stage controls in V_t^1 are the first-differences of contemporaneous values of labor productivity and the unemployment rate, and G_t^1 contains contemporaneous values of 127 FRED-QD series. The second stage controls in V_t^2 are the first four lags of the first-differences of labor productivity and the unemployment rate, and G_t^2 consists of the first lag of the 127 FRED-QD series.

C Subspace dimension

C.1 Subspace dimension in the Monte Carlo experiments

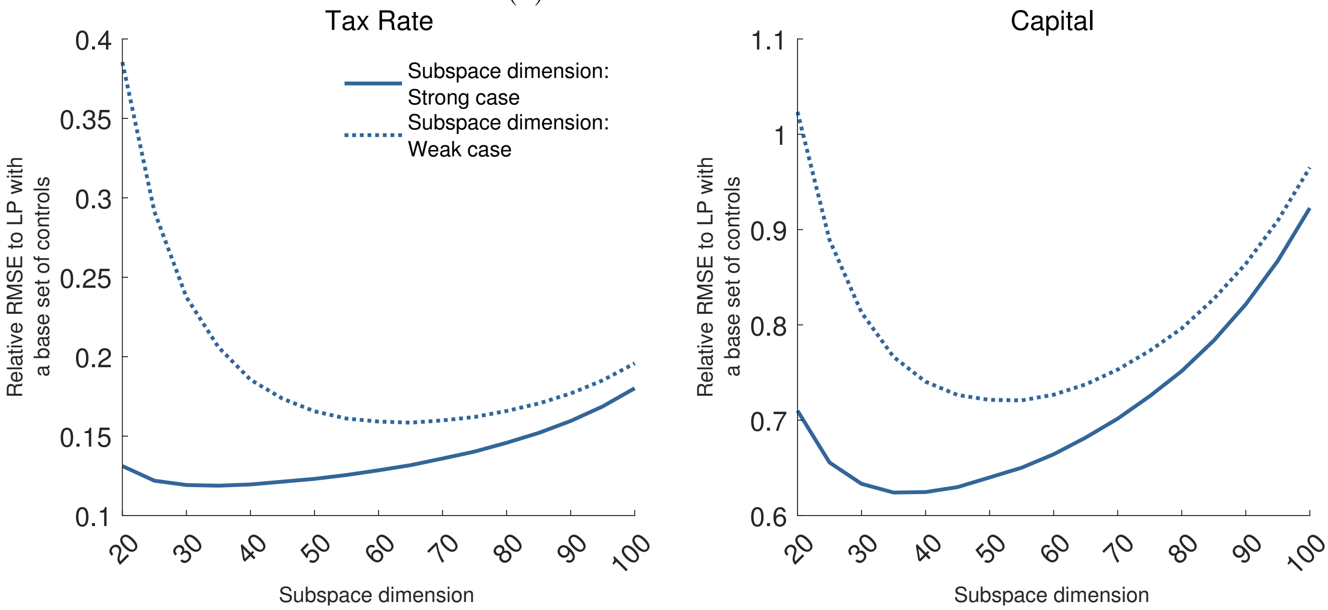
We explore the size of the subspace dimension in our Monte Carlo experiments. For both variables and both identification methods, we calculate the root mean squared error (RMSE) of the RSLP across horizons 0-6 in all generated data sets for each subspace dimension. Figure A2 reports these RMSEs relative to the RMSE of the local projection with the base set of controls.

While we do not find a unique subspace dimension that minimizes the RMSEs in each setting under consideration, we highlight two findings from this exercise. First, varying the subspace dimension reveals a "U-shaped" pattern, suggesting that when there are a few controls in each subspace regression, additional controls improve the estimates. This is, however, a trade-off where for some subspace dimension the variance in the estimates exceeds the reduction in bias. Second, we find that the minimum is often achieved with a subspace dimension within the 40-60 range, and the subspace dimension that minimizes the RMSEs is smaller in the strong cases. Intuitively, one needs fewer variables in the subspace regressions to filter out the relevant signal with stronger signals from the informational variables.

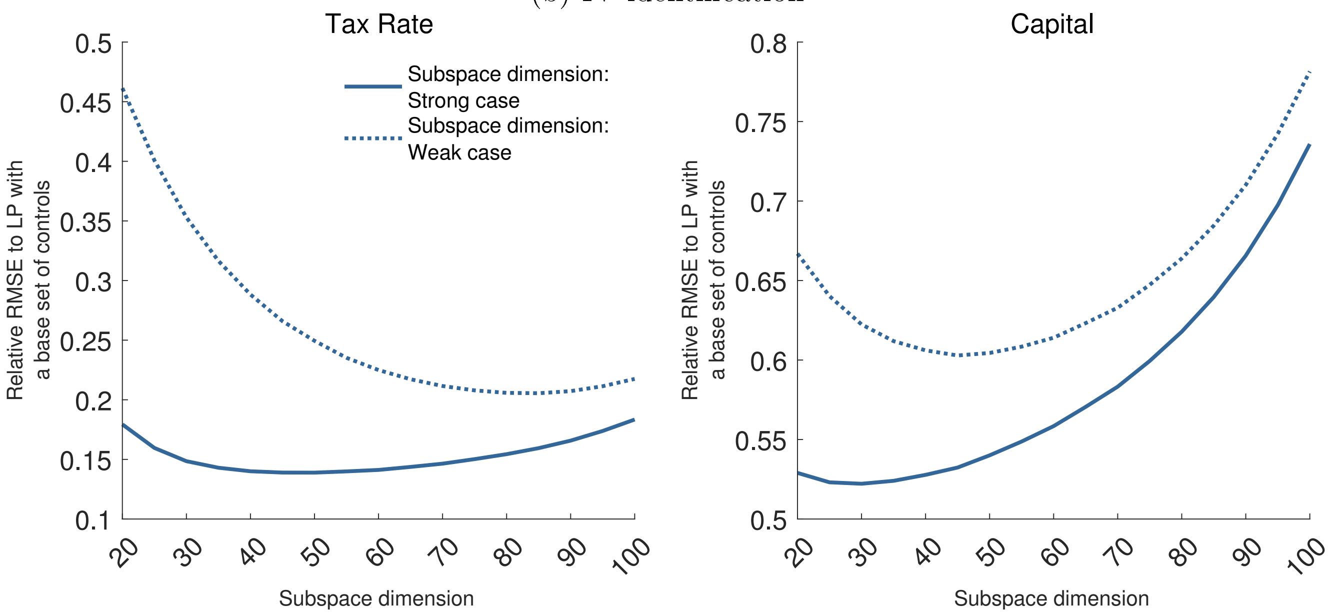
Although this is a simulated data example, we stress that the Monte Carlo experiments are

Figure A2: Subspace dimension RSLP in the Monte Carlo experiments

(a) SVAR identification



(b) IV identification



Note: This figure shows the root mean squared error of RSLP relative to LP with a base set of controls for various subspace dimensions, in the strong factor structure case (blue, solid) and the weak factor structure case (blue, dotted). The upper and lower panels correspond to the SVAR and IV identification, respectively.

useful in designing empirical strategies on how to set the subspace dimension for applied work. First, the suggestion from the Monte Carlo is that for a data set with a factor structure like the FRED-MD, a subspace dimension in the range of 40 to 60 is probably a good starting point. Moreover, Boot and Nibbering (2020) also find an optimal subspace dimension of 40-60 with the FRED-MD, albeit in a forecasting context. Second, given data sets are different, exploring some form of robustness of the subspace dimension is probably warranted for applied work.

Hence, we proceed with an empirical strategy that uses a subspace dimension in the 40-60 range, and subsequently check for the robustness of the results.

C.2 Subspace dimension in the empirical applications

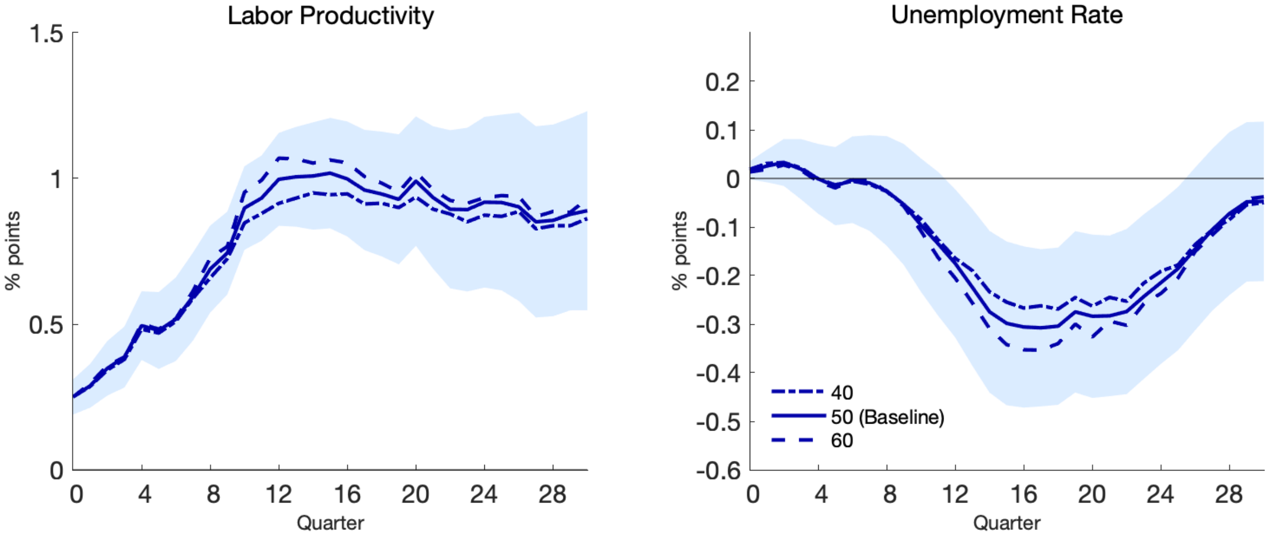
In the two empirical applications, we set the subspace dimension equal to 50 as follows from the discussion above. However, it remains an open question whether this choice of subspace dimension is appropriate in our applications. To investigate this choice, Figure A3 presents the estimated impulse response functions by RSLP for our applications, but now with varying subspace dimensions. The blue solid-line corresponds to a subspace dimension of 50, which is displayed alongside the estimates corresponding to the subspace dimensions of 40 and 60. We find that these estimated impulse response functions are very similar to that considered in our baseline. Moreover, they are within the one standard deviation interval of the baseline choice of 50. This indicates that our choice of subspace dimension is appropriate.

D FAVAR in the Monte Carlo experiments

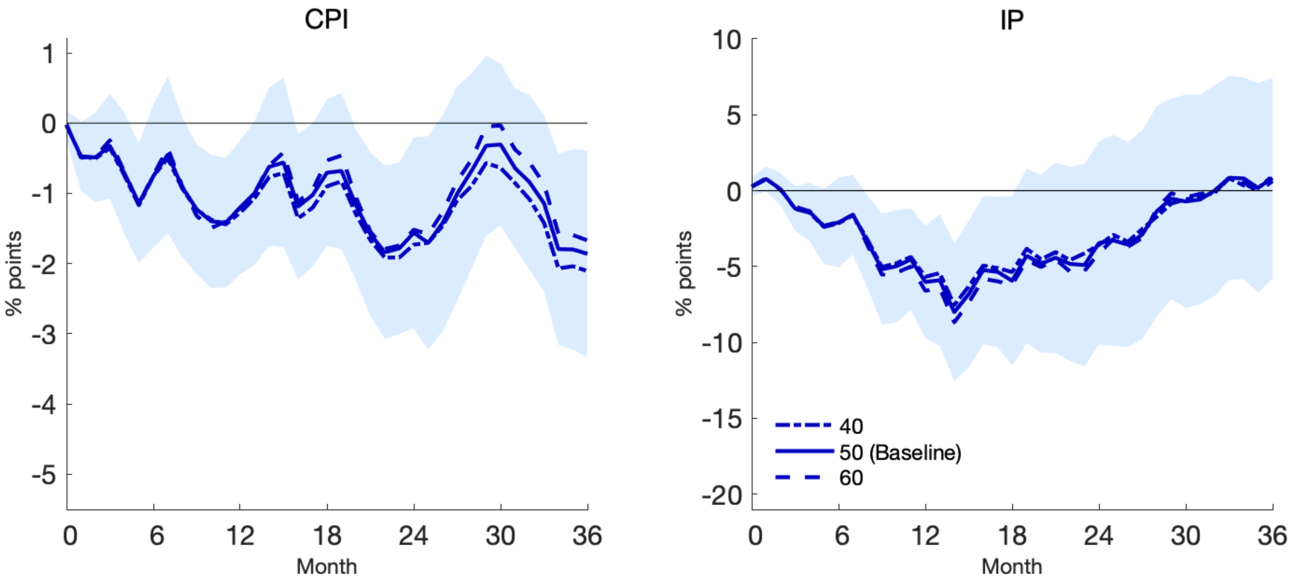
Figure A4 shows the average impulse response functions across 2000 replications for both SVAR identification (upper panel) and IV identification (lower panel) in the weak case. The black-dashed line represents the true impulse responses, and the green-dashed and blue-solid lines FAVAR and RSLP, respectively. When the factor structure is similar to the FRED data, the bias of FAVAR rises substantially. In contrast, RSLP is still tracking the true impulse response functions. The smaller bias in RSLP even dominates the efficiency in FAVAR, leading to a finding that RSLP outperforms FAVAR in terms of RMSE as shown in Table 1.

Figure A3: Subspace dimension RSLP in the empirical applications

(a) Response to a technology shock



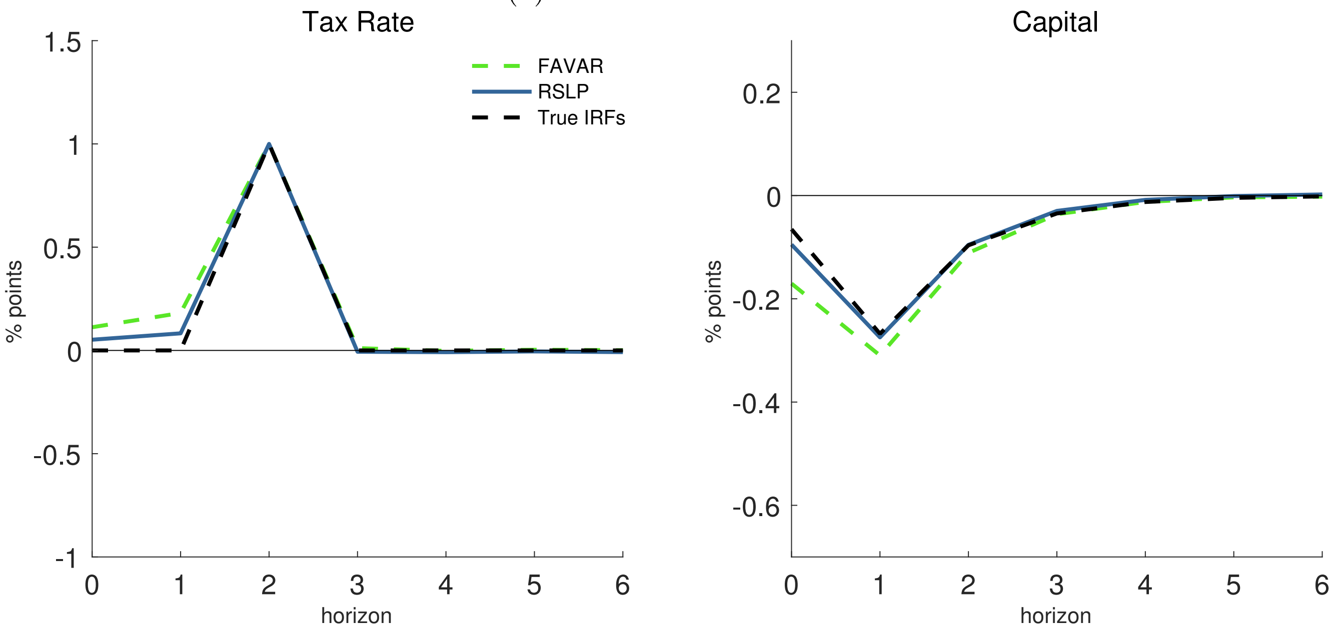
(b) Response to a monetary policy shock



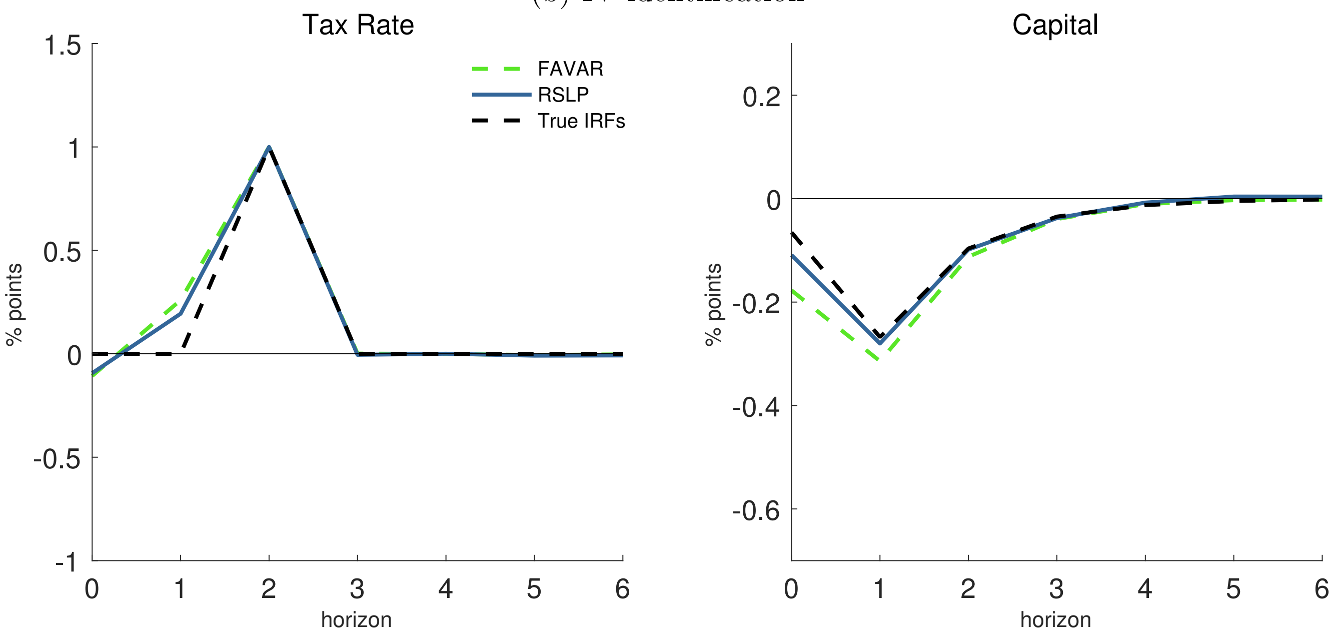
Note: This figure shows the estimated impulse response functions of labor productivity and unemployment rate to a technology shock (upper panel), and of CPI and IP to a monetary policy shock (lower panel). The lines correspond to different subspace dimensions, and the blue-shaded areas indicate the one standard deviation intervals of RSLP with a subspace dimension of 50.

Figure A4: FAVAR impulse responses functions in the Monte Carlo experiments

(a) SVAR identification



(b) IV identification



Note: This figure shows impulse response functions estimated by FAVAR (green, dashed), RSLP (blue, solid), and the true responses (black, dashed). FAVAR and RSLP use controls with a weak factor structure. The upper and lower panels use SVAR and IV identification, respectively. The shock is normalized to a 100 basis point increase in the tax rate. The results are the average impulse response functions across all 2000 Monte Carlo replications.

E Standard deviation of impulse response functions

Recall that our impulse response estimates are constructed as

$$\hat{\beta}_h = \frac{1}{n_R} \sum_{j=1}^{n_R} \hat{\beta}_h^{(j)}. \tag{A4}$$

Buckland et al. (1997) derive an expression for the variance of $\hat{\beta}_h$ under two assumptions. First, assume that the expectation across all possible draws $R^{(j)}$ of the misspecification bias in estimating $\beta_h^{(j)}$ in the model corresponding to $R^{(j)}$ is zero. That is, $\mathrm{E}[\beta_h^{(j)}] = \beta_h$. This holds per definition in case the shock x_t is exogenous. Second, assume that $\beta_h^{(j)}$ and $\beta_h^{(l)}$ are perfectly correlated, for all $l \neq j$. Although this is a strong assumption, each $\beta_h^{(j)}$ is estimated on the same underlying data set and these correlations are therefore indeed expected to be high.

It now follows that

$$\mathrm{SD}[\hat{\beta}_h] = \frac{1}{n_R} \sum_{j=1}^{n_R} \sqrt{\mathrm{var}(\hat{\beta}_h^{(j)} | \text{model } (j) \text{ is correct}) + (\beta_h^{(j)} - \bar{\beta}_h)^2}, \tag{A5}$$

which may be estimated by replacing $(\beta_h^{(j)} - \bar{\beta}_h)$ by $(\hat{\beta}_h^{(j)} - \hat{\beta}_h)$, and $\mathrm{var}(\hat{\beta}_h^{(j)} | \text{model } (j) \text{ is correct})$ by the squared Newey-West standard error for the ordinary least squares estimate $\hat{\beta}_h^{(j)}$ in model (j) to account for the serial correlation in the error terms (see Jordà, 2005).