Aggregate Hours and Local Projections with Long-Run Restrictions *

E. Charlie Nusbaum[†]

UC Santa Barbara

July 13, 2020

Abstract: I extend the local projections method to identify structural shocks using long-run restrictions. I show that the proposed estimator is substantially more robust than structural VARs (SVARs) to both the choice of lag length and the order of integration of the endogenous variables using data from a standard real business cycle (RBC) model. Benchmark simulations show that the proposed estimator can yield substantial reductions in both the bias and mean squared error of estimated impulse response functions relative to SVARs, particularly at short forecast horizons. In all cases, the proposed estimator correctly estimates the direction of the contemporaneous response and the shape of the full impulse response function, and can eliminate virtually all of the bias for some specification choices. Using my proposed estimator and data obtained from the Bureau of Labor Statistics, I then estimate the response of aggregate labor hours to productivity shocks. In contrast to much of the evidence based on SVARs, I find that labor hours rise in response to positive productivity shocks and subsequently follow a hump shaped profile. This result is robust to a number of specification choices and provides new evidence in support of the standard RBC model.

JEL Classification: E24, E32, E37, C32, C52

Keywords: Technology shocks; Local projections; Long-run restrictions; Real business cycle

*I thank Doug Steigerwald, Peter Rupert, Thomas Cooley, Javier Birchenall, and Kieran Walsh for thoughtful comments and discussions on earlier drafts of this paper. I would also like to thank participants of the UCSB Econometrics and Macroeconomics Research Groups for reading earlier drafts of this paper in detail and facilitating further discussions. Use was made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara.

[†]Department of Economics, UC Santa Barbara, 2127 North Hall, Santa Barbara, CA. enusbaum@umail.ucsb.edu.

1 Introduction

The estimation of impulse response functions is central to understanding the impact of shocks on the macroeconomy. Macroeconomists have largely relied on estimating vector auto-regressions (VARs) and imposing the minimum number of additional identifying restrictions to interpret them as structural. One set of identifying restrictions used with VARs are long-run restrictions, wherein practitioners restrict the long-run impact of shocks within the model. For example, the long-standing debate surrounding the response of hours to productivity shocks spurred by Galí (1999) has largely relied on evidence from structural VARs (SVARs) identified by assuming that demand shocks have no long-run effect on productivity growth.[1]

Long-run identification schemes such as these are not robust to two central specification choices: assumed order of integration of the endogenous variables and included lag length. In this paper, I propose an alternative estimator and investigate its properties. Using Monte-Carlo evidence, I show that it is significantly more robust to the choice of lag-length and order of integration of the endogenous variables than SVARs identified with long-run restrictions. I then provide new evidence that, consistent with Real Business Cycle (RBC) models and in contrast to much of the SVAR evidence, labor hours rise in response to productivity shocks.

Identification with long-run restrictions follow the work of Shapiro and Watson (1988), Blanchard and Quah (1989), and King et al. (1991). In this framework, a finite lag VAR is first estimated and the sums of the implied moving average coefficients are subsequently constrained to recover the structural parameters of the model. Cooley and Dwyer (1998) , Ravenna (2007), and Chari et al. (2008) have raised doubts regarding the validity of this approach. Together, they show that an inappropriate choice of lag length can severely bias impulse response functions estimated in this way. An inappropriate lag length not only biases the estimated AR coefficients due to omitted variables, but more importantly ignores terms in the moving average coefficients of the true data generating process.

My proposed estimator extends the local projections studied in Jordà (2005) to achieve structural identification. Local projections regress an endogenous variable on lags of itself and other endogenous variables independently for each forecast horizon. In this way, local projections estimate the moving average coefficients of the underlying data generating process directly rather than relying on recursive substitution as in VARs. As a result, the long-run impact of each shock may be constructed and constrained without excluding terms in the moving average representation of the underlying data generating process. Moreover, omitted variable bias is not passed between moving average coefficients through recursive substitution. Because of these features, structural

[1] Work by Christiano et al. (2003), Christiano et al. (2004), Francis and Ramey (2005), Galí and Rabanal (2005), Fernald (2007), Canova et al. (2010), and Saijo (2019) have all used variants of this approach to show that hours may fall on impact in response to positive productivity shocks.

identification using the proposed estimator should be expected to outperform SVARs of similar lag length regardless of the order of integration of the endogenous variables. To the best of my knowledge, I am the first to implement and study the properties of local projections with long-run restrictions.

To test my previous claim, I rely on a two-shock RBC model developed in Chari et al. (2007) and Chari et al. (2008) (hereafter CKM) as the data generating process. This model has several advantages. Its linearized form not only has a VAR(∞) representation for a wide range of parameter values, but its structural parameters can be identified with long-run restrictions. Moreover, the CKM model has been used to evaluate many empirical models, including SVARs, in McGrattan (2010), Kascha and Mertens (2009), and CKM. For each simulated data series, I estimate the impulse response function of labor hours to a productivity shock using the standard SVAR estimator and the proposed estimator. I find that the proposed estimator yields significant bias reductions relative to the SVAR both in estimating the full impulse response function and the contemporaneous response. For example, the proposed estimator reduces the bias of the contemporaneous response by 73% and 47% when labor hours are included in first differences and levels, respectively. My proposed estimator correctly estimates the direction of the contemporaneous response and the shape of the impulse response function in all cases, and can eliminate all of the biases for some specification choices.

I then isolate the small sample bias of the proposed estimator from other specification choices by increasing the length of each simulated data series. Similar to the findings of Erceg et al. (2005) in the context of SVARs, I find that the effect of small sample sizes on the structural parameter estimates depends on the empirical specification used.[2] Because the estimated AR suffer from the well known small sample bias first described in Hurwicz (1950), their sum and therefore the estimated structural parameters are also biased in small samples. The bias in the structural parameters, however, results from a non-linear transformation of that in the reduced form coefficients. As a result, the structural estimates may be biased upward even when the reduced form estimates are biased downward. This issue is particularly important at long forecast horizons, but quantitatively small at short forecast horizons. Despite these sensitivities, the proposed estimator outperforms the SVAR in terms of impact error and integrated bias across all forecast horizons considered.

Having established the advantages of my proposed estimator, I show that it has first order implications for existing empirical discussions that rely on SVARs identified with long-run restrictions by revisiting the debate on the response of hours to productivity shocks.[3] The overwhelming

[2]Faust and Leeper (1997) make a related point when constructing confidence bands and hypothesis testing after imposing long-run restrictions on estimated VARs.

[3]Notable exceptions in this literature are Basu et al. (2006) and Sims (2011), who do not rely exclusively on SVARs identified with long-run restrictions. They, however, come to differing conclusions.

conclusion of this literature has been that labor hours fall in response to productivity shocks, a result that is at odds with RBC models à la Kydland and Prescott (1982) and King and Rebelo (1999). Using data from the Bureau of Labor Statistics (BLS) on non-farm private business labor productivity and labor hours spanning 1948Q1 to 2019Q2, my proposed estimator indicates that labor hours rise in response to a productivity shock and subsequently follow a hump shaped profile. In contrast to Christiano et al. (2003) and Christiano et al. (2004) who come to a similar conclusion, my results are robust to whether or not labor hours are first differenced and the inclusion of regime switches in labor productivity growth.[4] Moreover, my results provide new evidence in support of the standard RBC model.

My findings also contribute to the growing literature assessing the ability of local projections to improve estimates of impulse response functions. Kilian and Kim (2011), for example, use data simulated from a VAR(12) to compare the coverage rates of impulse response functions estimated using local projections and VARs. They argue that the local projections method provides no apparent advantages. Brugnolini (2018) shows that their use of the Akaike (1974) information criterion in choosing the lag length rather than the Schwarz (1978) information criterion drives their results.[5] Both papers apply structural shocks estimated using a VAR with short-run restrictions to the local projections method. Choi and Chudik (2019) instead tests the local projections against several alternative estimation procedures when the sequence of structural shocks is perfectly observed rather than estimated. Recent work by Plagborg-Møller and Wolf (2019) proves that impulse response functions estimated from local projections and VARs are asymptotically equivalent and suggest that any identification schemes relying on local projections succeed if and only if SVARs succeed. I show that structural identification through my proposed estimator and the previously described SVAR approach can differ wildly in empirically relevant sample sizes.

The rest of the paper is organized as follows. Section 2 reviews long-run restrictions in the context of VARs to illustrate the crucial source of bias and describes the proposed alternative. Section 3 details both the assumed data generating process and the Monte-Carlo approach used to test my claims of the proposed estimator. Section 4 describes the results of the Monte-Carlo exercise. Section 5 uses the proposed estimator to study the response of aggregate hours to a productivity shock. Section 6 summarizes my findings.

[4]Garín et al. (2019) finds that hours decline on impact away from the zero lower bound, but rise when the zero lower bound is binding.

[5]The AIC prefers over-parameterized models. Brugnolini (2018) argue that the exercise presented in Kilian and Kim (2011) effectively asks the local projections method to outperform the true data generating process, which is not possible and therefore provides an inappropriate comparison between local projections and VARs.

2 Long-Run Restrictions

In this section, I begin by presenting a simple structural VAR, highlighting the need to impose additional identifying restrictions. I then review how to recover the structural shocks by imposing long-run restrictions to an estimated VAR and describe why this method is so sensitive to the omission of relevant lags. Finally, I present my proposed alternative estimator that imposes long-run restrictions using local projections. I focus on the simplest case with only a single lag throughout to simplify notation and facilitate exposition. The discussion below can be generalized to the case of an arbitrary number of lags.

2.1 The Problem with SVARs

The goal of imposing sufficiently many identifying restrictions to an estimated VAR is to orthogonalize the forecast errors, thereby recovering the sequence of structural shocks. Given that the reduced form parameters have been estimated, recovering the contemporaneous correlation matrix of the endogenous variables is sufficient to accomplish this goal. Estimating this matrix through long-run restrictions requires imposing assumptions on the estimated long-run impact of the structural shocks and subsequently inverting the reduced form parameters.

Consider, for example an n-variable structural VAR(1) given by

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t \tag{1}$$

where B is the contemporaneous correlation matrix and subsumes that variance of ε_t. B captures both the indirect effect of $\varepsilon_{i,t}$ on $x_{i,t}$ through the other endogenous variables as well as the contemporaneous effect of $\varepsilon_{-i,t}$ on $x_{i,t}$. In this way, the contemporaneous correlation matrix, B, orthogonalizes the shocks that drive the stochastic process, ε_t, so that they may be interpreted as structural (causal) shocks with $\mathbb{E}(\varepsilon_t \varepsilon_t') = I$, where I is the identity matrix. Because each endogenous variable is potentially a function of not only lagged variables, but also the contemporaneous values of each other endogenous variable, rearranging Eq. 1 and simply applying OLS methods will yield biased coefficients.

Instead, one must first estimate the reduced form of the structural VAR by inverting B and impose additional identifying restrictions to recover the contemporaneous correlation matrix. The reduced form VAR implied by Eq. 1 is given by

$$x_t = \underbrace{B^{-1}\Gamma_0}_{A_0} + \underbrace{B^{-1}\Gamma_1}_{A_1} x_{t-1} + \underbrace{B^{-1}\varepsilon_t}_{e_t} \tag{2}$$

where A_0 and A_1 are transformations of the structural parameters and e_t are forecast errors such that

$\mathbb{E}(e_t e_t') = \Omega$ is no longer the identity matrix. Because $\mathbb{E}(x_{t-1}\varepsilon_t) = \mathbb{E}(x_{t-1}e_t) = 0$, the parameters of Eq. 2 may be consistently estimated using OLS. Including the covariance matrix of e_t, this yields only $n + n^2 + \frac{n^2+n}{2}$ parameter estimates while the structural VAR is characterized by $n + 2n^2$ parameters.[6] Thus, $\frac{n^2-n}{2}$ restrictions on the long run impact of the structural shocks must be imposed to recover estimates of the structural parameters.

To obtain the long-run impact of a unit impulse of the structural shocks to the endogenous variables, it is instructive to re-write Eq. 2 in its moving average form. In particular, recursive substitution and the definition of e_t yields

$$x_t = B^{-1}\varepsilon_t + A_1 B^{-1}\varepsilon_{t-1} + A_1^2 B^{-1}\varepsilon_{t-2} + \dots \tag{3}$$

Thus, the time t impact of a structural shock i periods prior is $A_1^i B^{-1}\varepsilon_{t-i}$. Moreover, the long-run impact of shocks, D, is simply the sum of each of time t impact of those same shocks, i.e. $D = \sum_{i=0}^{\infty} A_1^i B^{-1}$. Further assuming that the eigenvalues of A_1 are less than 1 in modulus, the long-run impact matrix can then be written as

$$D = (I - A_1)^{-1} B^{-1} \tag{4}$$

where $(I - A_1)^{-1} = I + A_1 + A_1^2 + \dots$ are the reduced form moving average coefficients. As only A_1 in Eq. 4 can be estimated from Eq. 2, an additional equation to pin down D and recover B^{-1} is necessary. Given that $\mathbb{E}(\varepsilon_t \varepsilon_t') = I$, we have $\mathbb{E}(e_t e_t') = \Omega = B^{-1}B^{-1\prime}$ and so

$$DD' = (I - A_1)^{-1}\Omega(I - A_1')^{-1} \tag{5}$$

where A_1 and Ω can both be estimated from the reduced form VAR. Long-run restrictions can now be imposed on D such that Eq. 5 holds and A_1 and Ω satisfy Eq. 2. Given D, B^{-1} is then obtained by inverting the reduced form moving average coefficients, $(I - A_1)^{-1}$.

Taking a Cholesky decomposition of Eq. 4, for example, returns a triangular matrix for D and is equivalent to assuming that structural shocks have no long-run effect on the endogenous variables higher in the Cholesky ordering. By replacing A_1 and Ω with $\widehat{A}_1$ and $\widehat{\Omega}$, respectively, $\widehat{B}$ can be consistently estimated. Given estimates for the contemporaneous correlation matrix and reduced form moving average (lag) coefficients, it is straightforward to obtain the estimated s-step ahead impulse response to a structural shock of size d.

$$\widehat{IR}(s, d) = \widehat{A}_1^s \cdot \widehat{B^{-1}} \cdot d \tag{6}$$

[6] Generally, an estimated VAR(p) will provide $n + np + \frac{n^2+n}{2}$ parameter estimates compared to $n + n^2 p + n^2$ parameters in the structural VAR(p).

As is evident by Eq. 3-Eq. 5, consistently estimating the moving average coefficients of the underlying data generating process is critical to consistently estimating B and therefore $IR(s,d)$. Lag length mis-specification not only results in inconsistent estimates of the lag coefficients in Eq. 2, but more importantly in missing terms in all but the first two moving average coefficients. Suppose, for example, that the true data generating process took the form of a VAR(2) given by $x_t = A_0 + A_1 x_{t-1} + A_2 x_{t-2} + e_t$ rather than the VAR(1) described above. The first two coefficients of the moving average representation of both the true and assumed data generating process are equal to I and A_1. The third coefficient of the moving average representation on the other hand is A_1^2 for a VAR(1) and $A_1^2 + A_2$ for a VAR(2). A similar discrepancy exists between the two for each coefficient in their respective moving average representation after the third. This discrepancy poses first order consequences when imposing long-run restrictions on a VAR with incorrect lag length. CKM, for example, show that estimated contemporaneous responses may be more than double that of the true response, and may be of the incorrect sign even after controlling for small sample bias.[7]

2.2 Identification with Local Projections

The structural VAR methodology developed in the previous section is indeed useful for capturing the dynamics in a high-dimensional system. Estimated VARs, however, are a linear global approximation for the underlying system. That is to say that the dynamics of the system are determined recursively from one-step ahead forecasts. The local projections method provides an alternative.

The local projections method first presented in Jordà (2005) allows for a more direct and flexible estimation of the impulse response function. Rather than relying on recursive forecasts, he instead suggests estimating the dynamic relationships of a system at each forecast horizon independently with a collection of regressions. The local projections form of Eq. 3 is given by

$$x_{t+s-1} = A_0^{(s)} + A_1^{(s)} x_{t-1} + u_{t+s-1} \qquad s = 1, 2, \dots, s_{max} \tag{7}$$

where $A_0^{(s)}$ is an $n \times 1$ vector of constants, $A_1^{(s)}$ are matrices of coefficients for the lag dependent variable, s denotes the s-step ahead forecast, s_{max} is the maximum forecast horizon, and u_{t+s} is a mean zero forecast error. The forecast error contains information of all shocks from time t to time $t + s - 1$.[8] Jordá denotes the collection of equations given by Eq. 7 for $s = 1, 2, \dots, s_{max}$ as *local projections* due to the fact that each coeffecient is estimated equation by equation for each forecast horizon. As a result, these local projections are a set of local approximations to the true data

[7]This discussion puts aside omitted variable bias in the OLS step to better illustrate the central problem with the standard SVAR estimator. The reduced form coefficients of an estimated VAR(1) will of course also be biased due to the omitted second lag.

[8]It can be shown that the forecast errors given by u_{t+s-1} are a moving average of the reduced form residuals, $\{e_{t+i}\}_{i=0}^{s-1}$, when the true data generating process is a VAR.

generating process rather than a single global approximation as is the case with VARs. Jordà and Kozicki (2011) prove that, for a given forecast horizon, the coefficients in Eq. 7 can be consistently estimated by simple OLS.

Moreover, the local projections technique allows for a simple method to compute the impulse response functions. In particular, the set of reduced form impulse responses may be orthogonalized as in Eq. 6. The key distinction between the two is that rather than using recursive substitution to obtain $\widehat{A}_1^s$, A_1^s is estimated directly with $\widehat{A}_1^{(s)}$. A key question remains, however: how does one recover B? Current approaches to estimating structural impulse response functions using local projections simply apply a contemporaneous correlation matrix obtained elsewhere to the reduced form coefficients. For example, Jordà (2005) accomplishes this by recovering B using an estimated SVAR.[9] This approach, however, causes the shortcomings of SVARs to be passed on to the local projections through the inconsistent estimation of B described previously.[10]

Rather than obtaining B by imposing long-run restrictions on VARs, the local projections can instead be used directly. In particular, recall that a crucial source of bias in estimating B using VARs, is missing terms in each moving average coefficient as a result of lag misspecification. Because the local projections method does not rely s-steps of recursive substitution, each projection attempts to estimate its respective moving average coefficient directly. To see this, again consider the set of local projections given by Eq. 7 and substitute out x_{t-1} using the $s = 1$ local projection *only once*.

$$\begin{cases} x_t &= A_0^1 + A_1^{(1)} A_1^{(1)} x_{t-2} + A_1^{(1)} u_{t-1} + u_t \\ x_{t+1} &= A_0^2 + A_1^{(2)} A_1^{(1)} x_{t-2} + A_1^{(2)} u_{t-1} + u_{t+1} \\ x_{t+2} &= A_0^3 + A_1^{(3)} A_1^{(1)} x_{t-2} + A_1^{(3)} u_{t-1} + u_{t+2} \\ \vdots \end{cases} \tag{8}$$

Noting several features of Eq. 8. First, note that the structural shocks, ε_t, drive the stochastic process and that x_{t-2} summarizes all past shocks by assumption. Further note that u_{t+s-1} summarizes information on $\{e_{t+i}\}_{i=0}^{s-1}$. Then, the only information contained in u_{t-1} is that of the time $t-1$ structural shocks, ε_{t-1}. Thus, $A_1^{(1)}$ is the $t-1$ moving average coefficient, $A_1^{(2)}$ is the $t-2$ moving average coefficient, and so on. The time t moving average coefficient is trivially given by the identity matrix. Given this collection of local projections, the moving average coefficients used construct Eq. 5, A_1^i, can simply be replaced by the local projections coefficients, $A_1^{(s)}$. The tern

[9] This is similarly done in the critiques by Kilian and Kim (2011) and Brugnolini (2018).

[10] If the sequence of structural shocks is known (e.g. Hamilton, 1985; Romer and Romer, 2004; Ramey, 2011), then they may instead be included as regressors themselves.

$(I - A_1)^{-1}$ in Eq. 5 is then given by

$$(I - A_1)^{-1} = \begin{bmatrix} 1 + \sum_s a_{11}^{(s)} & \sum_s a_{12}^{(s)} & \dots & \sum_s a_{1n}^{(s)} \\ \sum_s a_{21}^{(s)} & 1 + \sum_s a_{22}^{(s)} & \dots & \sum_s a_{2n}^{(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_s a_{n1}^{(s)} & \sum_s a_{n2}^{(s)} & \dots & 1 + \sum_s a_{nn}^{(s)} \end{bmatrix} \tag{9}$$

The estimation of B then proceeds as before after substituting $\widehat{a}_{ij}$ for a_{ij} and $\widehat{\Omega}$ for Ω, where $\widehat{\Omega}$ obtained from an estimated VAR and the $s = 1$ local projection are equivalent.[11]

This alternative estimator has several distinct advantages. First, any omitted variable bias in the estimated lag coefficients from including too few lagged dependent variables is not compounded across forecast horizons as a result of s-steps of recursive substitution. Second, and most importantly, because the local projections method provides a collection of local approximations rather than a single global approximation as in VARs, summing across the estimated lag coefficients as described above does not result in missing terms in the implied moving average coefficients as is the case with improperly specified VARs. Instead, $A_1^{(s)}$ estimates each moving average coefficient directly with the obvious normalization that the first, i.e. $s = 0$, moving average coefficient is not estimated and is instead given by the identity matrix. This fact may not be obvious at first glance as the proposed local projections estimator still relies on recursive substitution. The key distinction is that the proposed local projections estimator relies only on one-step of recursive substitution. That is, the recursive substitution used in the proposed estimator relies only on the fact that the first moving average coefficient for a VAR of arbitrary lag length is the identity matrix. Thus, relying on one step of recursive substitution does not reintroduce the same issues present in the standard SVAR estimator.

It is immediately evident, that the proposed method relies on an appropriate choice of maximum forecast horizon, s_{max}, to construct the sums Eq. 9. It is also immediately evident that this estimator does not simply trade one problem for another. Increasing the accuracy of impulse response functions estimated using an n-variable VAR with long-run restrictions requires including more lags. Each additional lag results in a degrees of freedom reduction of n. Instead, improvements to local projections with long-run restrictions can be made by increasing the maximum forecast horizon rather than increasing the lag length of each regression. This results in a degrees of freedom reduction of only one. Thus, the proposed method for imposing long-run restrictions is

[11] Christiano et al. (2006) develop a non-parametric method to estimate Ω that may further improve my proposed method. As shown in Section 4, however, my proposed estimator yields significant bias reductions even without relying on the methods developed in Christiano et al. (2006).

less limited by the length of data series available.[12] I return to this issue in Section 4 where I show that empirically relevant sample sizes allow for a sufficiently long forecast horizon to improve the estimation of both contemporaneous responses and the subsequent path of the impulse response function using my proposed estimator.

3 Testing the Proposed Estimator

To test the properties of the proposed method against that of the standard SVAR approach, I rely on a two-shock version of a Real Business Cycle model developed in Chari et al. (2007) and CKM as the data generating processes.[13] This model has the distinct advantage that its linearized form satisfies the invertibility conditions described in Fernández-Villaverde et al. (2007) and Lippi and Reichlin (1994) that permit a VAR(∞) representation for a wide range of parameter values. Moreover, Chari et al. (2007) show that including stochastic wedges into a canonical RBC model such as this describes post-war aggregates well. This allows me to test the claim that local projections with long-run restrictions (LRLPs) are more robust than SVARs to specification choices. I begin by simulating the model for 285 quarters 1,000 times and estimating the impulse response of labor to a productivity shock using both an SVAR and LRLPs for each simulated series.[14] I adjust the lag length of both the estimated LRLPs and SVAR, and the maximum forecast horizon of the local projections method to investigate its advantages in empirically relevant samples. I then repeat this exercise using a sample length of 10,000 quarters to investigate the effects of small sample bias for the proposed estimator. Throughout, I focus on the response of hours to a productivity shock as this is the statistics most used in the literature to choose between business cycle models. Moreover, this is the statistic I estimate in Section 5 where I revisit the hours debate.

3.1 Two-Shock RBC Model

A unit mass of infinitely lived households maximize expected utility by making a consumption-savings decision and a labor supply decision in frictionless markets, and discount the future at rate β. Households are subject to stochastic labor taxes that are rebated lump-sum in each period. Furthermore, I assume that household preferences are additively separable within and across

[12]It is well known that properly specified VARs are more efficient than their local projection counterparts and so long-run restrictions imposed with VARs dominate asymptotically.

[13]I present only the key equations. Refer to Chari et al. (2007) and CKM for more details and proofs.

[14]The data series used in Section 5 is approximately 285 quarters in length.

periods, and are of the CES form. The problem of the representative household is then given by,

$$\max \mathbb{E}_{t_0} \sum_{t=t_0}^{\infty} [\beta(1+\gamma)]^t [\log c_t + \phi \log(1-l_t)] \quad s.t. \quad c_t + (1+\gamma)k_{t+1} - (1-\delta)k_t = (1-\tau_{lt})w_t l_t + r_t k_t + T_t \tag{10}$$

where all variables are in per-capita terms. Here, δ is depreciation, γ is population growth, k_t is per-capita capital stock, w_t is the wage rate, r_t is the return on capital, T_t are lump sum rebates of the stochastic labor wedge, τ_{lt}, and c_t and l_t are per-capita consumption and labor, respectively.

Firms produce a numeraire consumption good using a Cobb-Douglas production function with labor augmenting technology, Z_t. Firm's maximize their profit function, $(k_t)^{\alpha} (Z_t l_t)^{1-\alpha} - r_t k_t - w_t l_t$, by choosing labor and capital input each period.

The stochastic variables, $\log Z_t$ and τ_{lt}, are subject to independent mean zero shocks each period given by $\log z_t$ and ε_{lt}, respectively. The covariance matrix of these shocks is therefore given by $\begin{bmatrix} \sigma_z^2 & 0 \\ 0 & \sigma_l^2 \end{bmatrix}$. I further assume that Z_t is log-normally distributed and follows a random walk with drift, μ_z, in logs. τ_{ls} is instead normally distributed and is defined by a stationary AR(1) process in levels. The evolution of the stochastic variables is then characterized by

$$\log Z_t = \mu_z + \log Z_{t-1} + \log z_t \tag{11}$$

$$\tau_{lt} = (1-\rho_l)\mu_l + \rho_l \tau_{lt-1} + \varepsilon_{lt} \tag{12}$$

where ρ_l is the persistence of the stochastic labor wedge and $\log z_t$ and τ_{lt} are independent normally distributed shocks. With the assumption that $0 < \rho_l < 1$, innovations in the labor wedge are temporary. To close the model, the aggregate resource constraint is given by $c_t + (1+\gamma)k_{t+1} = y_t + (1-\delta)k_t$.

Log-linearizing about the steady state using standard methods, one may define a state-space system of the model. I assume that labor productivity growth, $\log\left(\frac{y_t}{l_t}\right)$, and the log of labor hours, $\log(l_t)$ are observable. The system is defined this way for a several reasons. If more variables are included in the observer equation than shocks driving the process–in this case two– then a subset of the included variables will be a linear combination of the others, precluding a valid VAR(∞) representation of the model (Ingram et al., 1994; Ireland, 2004; Fernández-Villaverde et al., 2007). Moreover, defining the observer equation in this way allows for a VAR(∞) representation of the model that is consistent with the large literature estimating the aggregate response of hours worked mentioned previously–shocks to the stochastic labor wedge have no long-run effect on labor-productivity growth.

To make the model quantitative, I set $\phi = 1.6$, $\alpha = 0.33$, depreciation to be 6%, the rate of time preference to be 2%, population growth to be 1%, and the technology growth rate to be 2%. All parameters are set to match these annualized rates such that a period in the model is equivalent to one quarter. The parameters governing the stochastic variables are set so that $(\mu_z, \mu_l) = (0.005, 0.4)$, the persistent of the stochastic labor wedge is $\rho_l = 0.95$, and the standard deviation of each shock as $(\sigma_z, \sigma_l) = (0.0114, 0.00725)$.[15]

3.2 Econometric Specification

To investigate the properties of the proposed method, I estimate the model analogue of Eq. 7 given by

$$\begin{cases} \Delta \log\left(\frac{y_{t+s}}{l_{t+s}}\right) & = a^{(s)}_{0,\frac{y}{l}} + a^{(s)}_{1,\frac{y}{l}} \Delta \log\left(\frac{y_{t-1}}{l_{t-1}}\right) + a^{(s)}_{1,l} \log l_{t-1} + \ldots + a^{(s)}_{p,\frac{y}{l}} \Delta \log\left(\frac{y_{t-p}}{l_{t-p}}\right) + a^{(s)}_{p,l} \log l_{t-p} + u^{(s)}_{\frac{y}{l},t+s} \\ \log l_{t+s} & = b^{(s)}_{0,l} + b^{(s)}_{1,\frac{y}{l}} \Delta \log\left(\frac{y_{t-1}}{l_{t-1}}\right) + b^{(s)}_{1,l} \log l_{t-1} + \ldots + b^{(s)}_{p,\frac{y}{l}} \Delta \log\left(\frac{y_{t-p}}{l_{t-p}}\right) + b^{(s)}_{p,l} \log l_{t-p} + u^{(s)}_{l,t+s} \end{cases} \tag{13}$$

for $s = 0, 1, 2, \ldots, s_{max}$. As described in Section 2, the two main advantages of imposing long-run restrictions using local projections are that each regression is a local approximation and so bias at any one forecast horizon is not compounded due to recursive substitution, the reduced form coefficients estimate the MA coefficients of the underlying data generating process directly, and that additional terms may be included in Eq. 9 by simply increasing the s_{max}. While I use a lag length of 4 and a maximum forecast horizon of 25 quarters as a benchmark, I adjust these two specification choices to illustrate the robustness of this method to the choice of lag length and investigate the sensitivity of LRLPs to the choice of s_{max}.

Moreover, I estimate a VAR(p) with labor-productivity growth ordered first to illustrate the advantages of local projections with long-run restrictions. The estimated model is given by

$$\begin{bmatrix} \Delta \log\left(\frac{y_t}{l_t}\right) \\ \log l_t \end{bmatrix} = \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} \Delta \log \frac{y_{t-1}}{l_{t-1}} \\ \log l_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \tag{14}$$

where $A_{ij}(L)$ are lag polynomials of order p, and e_{1t} and e_{2t} are reduced form forecast errors. The length of the lag polynomial included in Eq. 14 is always consistent with that of Eq. 13.

While labor hours in the model are stationary by construction, Dickey-Fuller tests performed on the data used in Section 5 do not reject the null hypothesis of a unit root in hours. Moreover, Francis and Ramey (2005), Galí and Rabanal (2005), and CKM all argue that including labor in levels is inferior to an econometric specification including first differenced labor hours.[16] In light of these considerations, I also estimate the local projections and VAR using a differenced specification by

[15] CKM set the standard deviation of productivity shocks by dividing the standard deviation of TFP estimated in Prescott (1986) by labor share.

[16] Recent work by Saijo (2019) also uses a first differenced specification, but does not provide a clear justification of this choice.

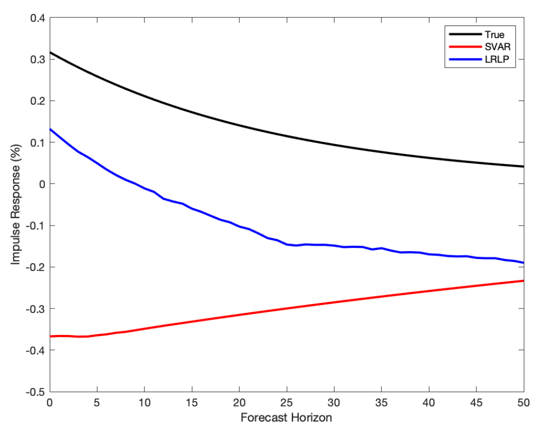
replacing $\log l_t$ with $\Delta \log l_t = \log l_t - 0.99 \cdot \log l_{t-1}$ in Eq. 13 and Eq. 14.[17]

4 Monte-Carlo Results

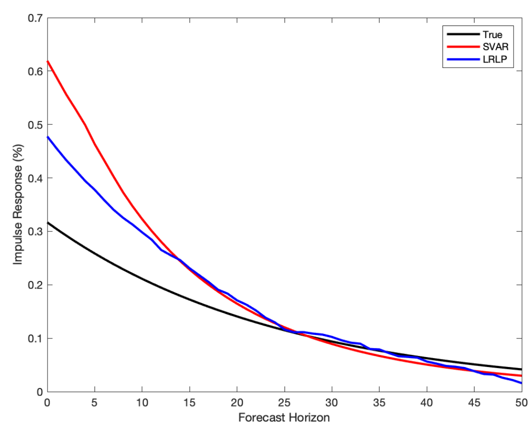
4.1 Small Sample Results

The benchmark results wherein 1,000 time series of 285 quarters are simulated and each estimator applied to all series are displayed below.[18] Figure 1 compares the mean estimated response of hours to a one standard deviation technology shock to the true response of the model. The results of the econometric specifications incorporating log-labor in differences are shown in Figure 1a and the results when including log-labor in levels are shown in Figure 1b. Throughout this section, the results from the benchmark local projections method and SVAR are shown in dark blue and red, respectively. The model implied response is always shown in black. Moreover, I present the estimated impulse response function to a forecast horizon of 50 quarters in order to illustrate both the full shape of the estimated impulse response function and how it compares across specification choices. In cases where $s_{max} < 50$, I estimate the local projections to a forecast horizon of 50 but use only the first s_{max} coefficients in the construction of the contemporaneous correlation matrix. I use the remaining $50 - s_{max}$ coefficients only to extend the estimated impulse response function to a forecast horizon of 50 for graphical purposes.

Figure 1: Comparison to True Response in Small Samples



(a) Differences Specification



(b) Levels Specification

On impact, the model implies a 32 basis point rise in hours worked. Relative to the mean response estimated from the SVAR, the bias of the LRLPs is substantially reduced. In fact, the

[17]This avoids over differencing but is quantitatively equivalent to a first differenced specification asymptotically.
[18]The length of these time series are chosen to match those in Section 5.

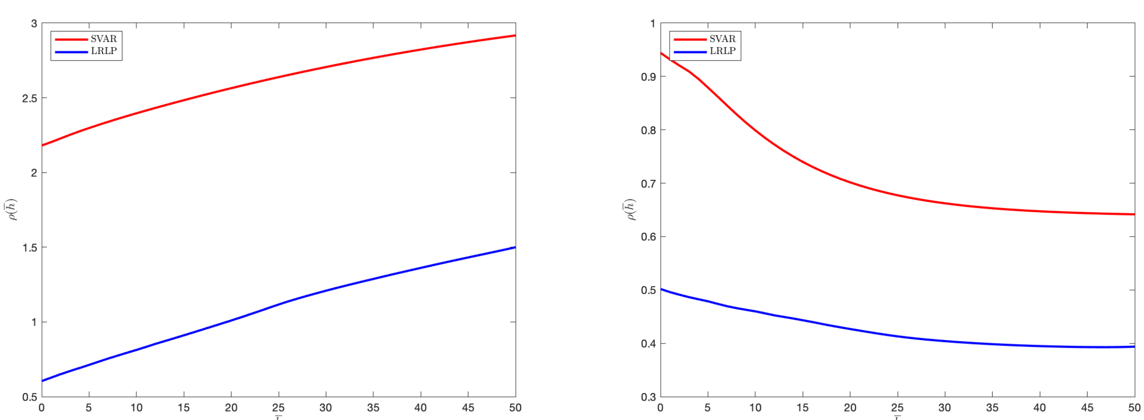
impact response of hours estimated using SVARs is 69 basis points lower and 30 basis points higher than the true response for the differences and levels specification, respectively. The mean impact response of the proposed estimator with labor included in differences and in levels is instead 19 basis points lower and 16 basis points higher than the true response, respectively. This translates into bias reductions in the response on impact for differences and levels specification of 72% and 47%, respectively. Moreover, the proposed estimator accurately represents the shape of the true response in both specifications whereas the shape of the impulse response function estimated using the SVAR changes drastically with the choice of hours used.

To quantitatively characterize the performance of each method in estimating the full impulse response function, I rely on a normalized integrated root mean squared error defined by

$$\rho\left(\widehat{IR}, IR, \overline{h}\right) = \left(\frac{\int_0^{\overline{h}} \left(\widehat{IR}(h) - IR(h)\right)^2 dh}{\int_0^{\overline{h}} IR(h)^2 dh} \right)^{\frac{1}{2}} \tag{15}$$

where $\widehat{IR}(h)$ is the estimated impulse response at horizon h, $IR(h)$ is the true impulse response at time h, and $\overline{h}$ is the maximum forecast horizon of interest. This metric is both robust to scale and punishes deviations symmetrically. Figure 2 and Figure 3 shows this metric as a function of $\overline{h}$ for the benchmark estimates of the differences and levels specification. The former shows $\rho\left(\mathbb{E}\widehat{IR}, IR, \overline{h}\right)$ and so identifies squarely the integrated bias of each estimator. The latter instead displays $\mathbb{E}\rho\left(\widehat{IR}, IR, \overline{h}\right)$ and so captures the bias-variance tradeoff of each estimator.

Figure 2: Integrated Bias

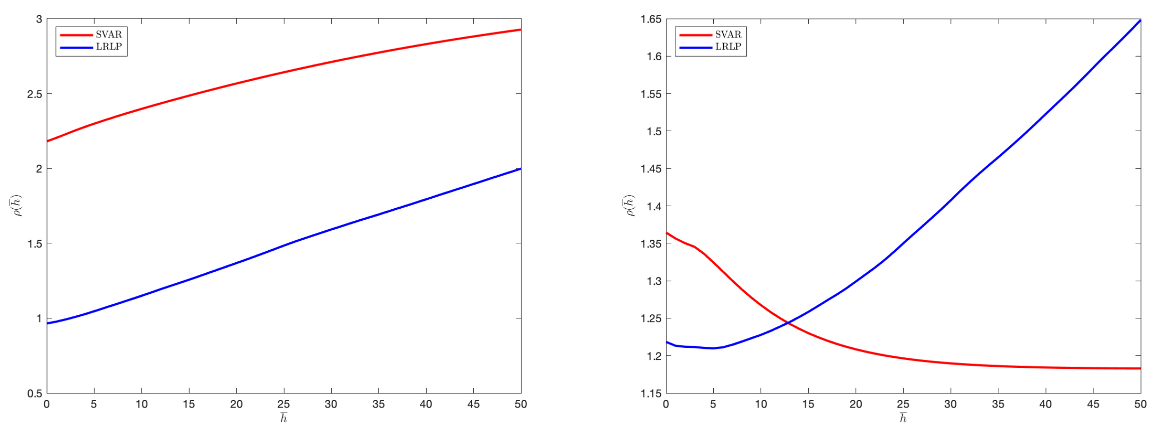


(a) Labor in Differences: $\rho\left(\mathbb{E}\widehat{IR}, IR, \bar{h}\right)$

(b) Labor in Levels: $\rho\left(\mathbb{E}\widehat{IR}, IR, \bar{h}\right)$

In both the levels and differences specification, the LRLPs outperform the SVAR method in terms of bias. Indeed, Figure 2a and Figure 2b show that integrated bias of the LRLPs is 50% lower than the SVAR method in the differences specification and 39% lower in the levels specification even after 48 quarters (12 years). Reductions in the integrated bias are even more substantial at shorter forecast horizons. Figure 3a further shows that the proposed estimator outperforms the SVAR method in terms of mean squared error at all horizons considered in the differences specification. In the levels specification, the LRLPs outperform the SVAR in terms of mean squared error at forecast horizons less than 3 years, after which the SVAR performs better. Clearly, however, this is a result of the reduced efficiency of local projections at long forecast horizons rather than the SVAR overtaking LRLPs in terms of bias. Moreover, the reductions in efficiency are relatively small as $\mathbb{E}\rho$ is only 8% lower for the SVAR than the LRLPs even after 24 quarters (i.e. 6 years). Also notice that the mean squared error of the SVAR estimator asymptotes whereas the same is not necessarily true for the local projections estimator. This is a direct result of the fact that the SVAR estimator relies on recursive substitution and the true response decays to 0 in the model. At horizons greater than the included lag-length, the effect of the shock decays due to multiplication of the estimated AR coefficients in the reduced form VAR. Thus, the estimated effect on hours eventually returns to 0 regardless of bias in $\widehat{B}$ as in the model. Because this is not a feature of the local projections estimator, the estimated impulse response of a shock does not necessarily decay to 0 at long forecast horizons.

Figure 3: Integrated Mean Squared Error



(a) Labor in Differences: $\mathbb{E}\rho\left(\widehat{IR}, IR, \bar{h}\right)$

(b) Labor in Levels: $\mathbb{E}\rho\left(\widehat{IR}, IR, \bar{h}\right)$

Finally, to further illustrate the relative performance of the local projections estimator and the standard SVAR estimator, Table 1 and Table 2 presents the bias and mean squared error, rather than their integrated counterparts, at each forecast horizon. I normalize each by the value of the true impulse response function and the square of this value, respectively, to induce scale invariance. In addition, I highlight in grey rows in which the LRLP estimator yields the correct qualitative result but the SVAR does not. That is, I highlight cases when the LPBQ estimator has the correct sign but the SVAR estimator does not. Cases in which the LRLPs do not yield improvements are indicated with a dash. Evidently, the LRLP yields large bias reductions at virtually every forecast horizons regardless of the specification used. Moreover, the LRLP estimator yields mean squared error reductions at all forecast horizons considered for the first differences specification. For the levels specification, the LRLPs yield mean-squared error improvements for throughout the first year and a half and subsequently do not outperform the SVAR estimator. As explained above, however, this is a result of increasing variance at long forecast horizons rather than SVARs outperforming LRLPs in terms of bias. Still, the LRLPs yield reductions in mean squared error at forecast horizons typically of most interest for both specification choices.

Table 1: Bias and MSE in First Differences

	Bias			Mean-Squared Error		
Horizon	*LRLP*	*SVAR*	*% Reduction*	*LRLP*	*SVAR*	*% Reduction*
0	-0.185	-0.684	73.0	1.53	5.02	69.6
1	-0.191	-0.670	71.5	1.61	5.28	69.5
2	-0.197	-0.658	70.0	1.66	5.58	70.3
3	-0.203	-0.648	68.7	1.81	5.92	69.4
4	-0.205	-0.637	67.8	1.91	6.22	69.2
5	-0.209	-0.623	66.5	2.11	6.46	67.3
6	-0.214	-0.610	65.0	2.27	6.73	66.2
7	-0.217	-0.597	63.6	2.48	7.00	64.6
8	-0.219	-0.585	62.5	2.65	7.31	63.8
9	-0.219	-0.572	61.7	2.80	7.61	63.1
10	-0.222	-0.560	60.3	2.98	7.92	62.4
15	-0.232	-0.504	54.0	4.50	9.76	53.9
20	-0.244	-0.456	46.6	6.94	12.1	42.9
25	-0.261	-0.415	37.1	11.7	15.3	23.1

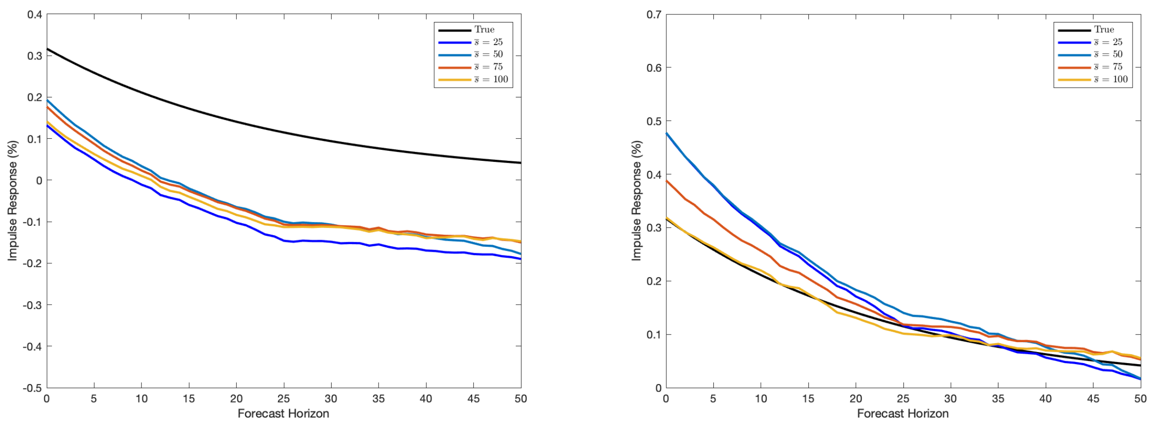
The improvements of the local projections method over the SVAR of course depend on the lag length used and maximum forecast horizon, s_{max}, included in the estimation of the contemporaneous response. I therefore re-estimate the impulse response function for each time series after varying the maximum forecast horizon, s_{max}, and lag length, p, included. Figure 4a and Figure 4b show the results of adjusting s_{max} in the differences and levels specifications, respectively. The results from adjusting the included lag length are shown in Figure 5.

Several properties of the proposed estimator are immediately evident. First, the proposed estimator is sensitive to the maximum forecast horizon, s_{max}. This results from the fact that s_{max} determines where the summations in Eq. 9 are truncated and therefore the number of estimated moving average coefficients included in the construction of $\widehat{B}$. This sensitivity is greater in the

Table 2: Bias and MSE in Levels

	Bias			Mean-Squared Error		
Horizon	*LRLP*	*SVAR*	*% Reduction*	*LRLP*	*SVAR*	*% Reduction*
0	0.161	0.302	46.8	2.24	2.61	14.3
1	0.151	0.284	46.8	2.23	2.59	13.7
2	0.141	0.264	46.5	2.20	2.53	13.1
3	0.133	0.248	46.1	2.23	2.55	12.4
4	0.125	0.230	45.6	2.19	2.48	11.9
5	0.119	0.205	41.7	2.22	2.32	4.23
6	0.110	0.185	40.3	2.21	2.21	0.00
7	0.102	0.164	37.8	2.26	2.08	–
8	0.096	0.144	33.0	2.34	1.96	–
9	0.093	0.127	26.9	2.44	1.86	–
10	0.087	0.112	22.2	2.51	1.77	–
15	0.059	0.056	–	3.11	1.48	–
20	0.030	0.024	–	4.07	1.34	–
25	0.001	0.006	73.7	6.09	1.29	–

Figure 4: Sensitivity to Choice of s_{max}



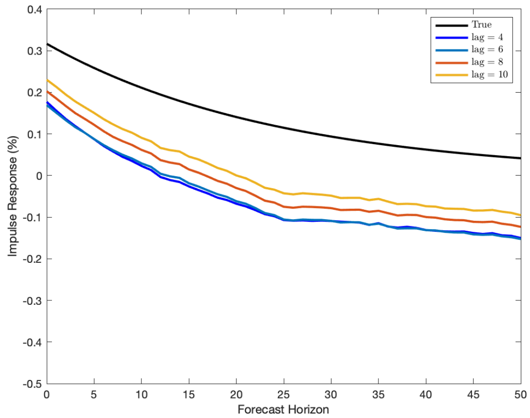
(a) Labor in First Differences

(b) Labor in Levels

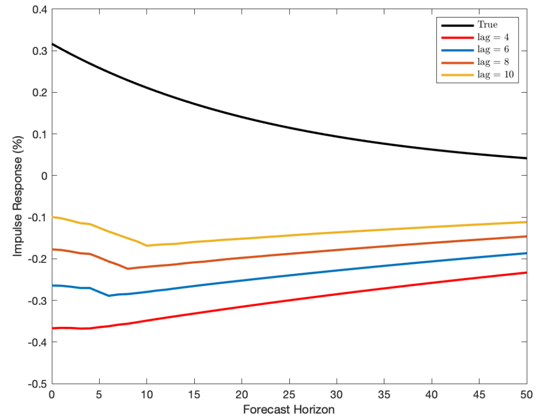
levels specification than in the differences specification. In the latter, increasing s_{max} results in a mean estimated impulse response function that is quantitatively more similar to the true response. The bias in the estimated LRLPs with $s_{max} = 100$ is negligible at all but the longest forecast horizons. In the differences specification, the estimated response changes relatively little with s_{max}. Taken together, this suggests that including 100 moving average terms is quantitatively sufficient in the construction of $\widehat{B}$ to eliminate the bias from omitted moving average coefficients.

Mirroring the sensitivity of the proposed estimator to the choice of s_{max}, the choice of included lag length appears to be less of a concern for the levels specification than for the differences specification. In fact, Figure 5c shows that the levels specification shows virtually no adjustment in

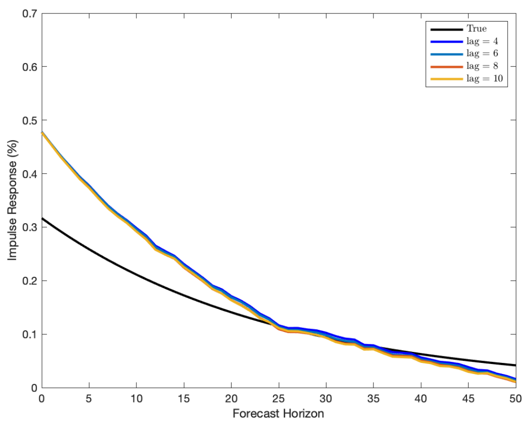
Figure 5: Sensitivity to Choice of Lag Length



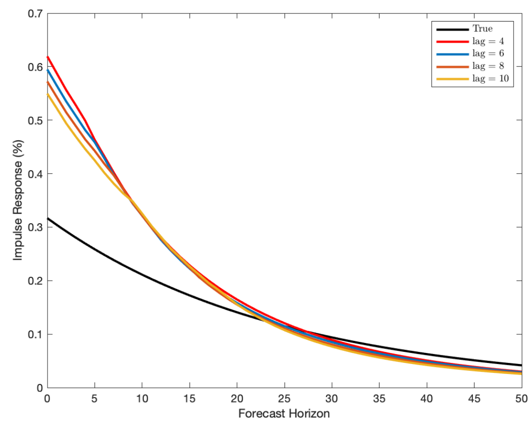
(a) Labor in Differences: LRLP Method



(b) Labor in Differences: SVAR Method



(c) Labor in Levels: LRLP Method



(d) Labor in Levels: SVAR Method

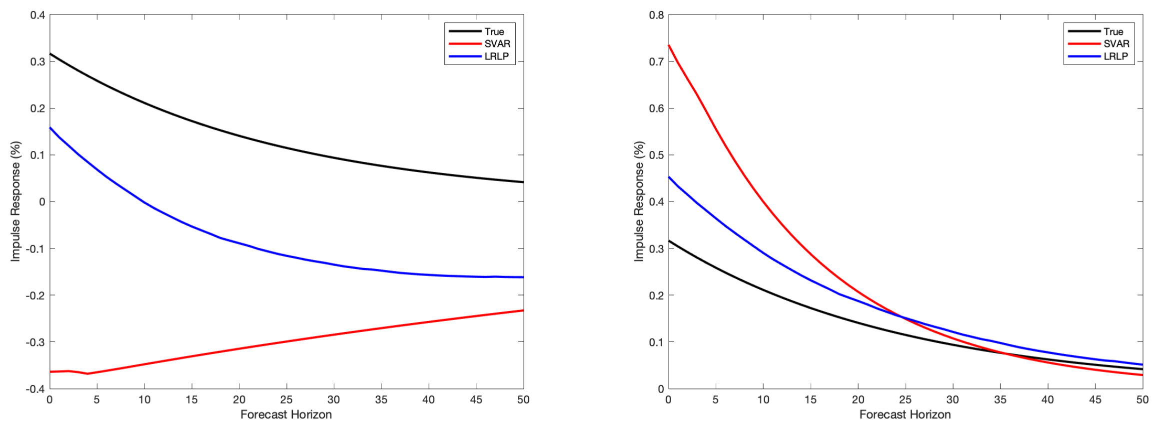
the mean estimated response. Figure 5a instead shows that the differences specification does show improvements with the included lag length. Despite these sensitivities, however, it is notable that the shape of the proposed estimator remains reflective of the true response regardless of the choice of s_{max} and included lag length. The shape of estimates from the SVAR on the other hand differ not only with the measure of hours used, but also with the choice of included lag length in the case of the differences specification.[19]

[19]Bias reductions from increasing s_{max} and the included lag length are of course accompanied by reductions in efficiency for both estimators.

4.2 Removing Small Sample Bias

To investigate the relative performance of each method without contaminating the results with small sample bias, I repeat the above exercise with 1,000 simulated time series of 10,000 quarters in length. In these exercises, the LRLPs and SVARs suffer from downward bias AR coefficients to a quantitatively insignificant degree. Figure 6a to Figure 6b mirror Figure 1a and Figure 1b and show the results of the large sample exercise. I focus on the benchmark specifications as the alternative specifications follow the same qualitative trends as in the previous section.

Figure 6: Comparison to True Response without Small Sample Bias



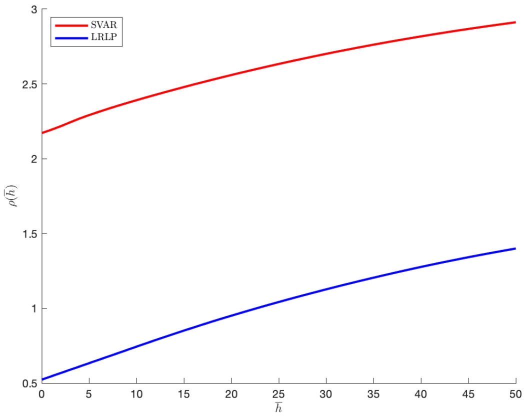
(a) Differences Specification

(b) Levels Specification

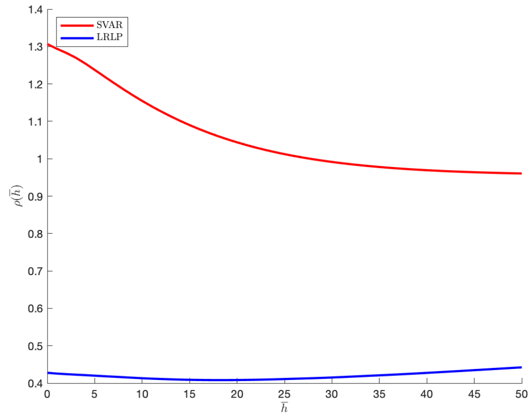
The contemporaneous response of the LRLP method in differences and levels is 16 basis points and 45 basis points, respectively. Relative to the estimated contemporaneous responses in the small sample case, the former is shifted up by 3 basis points and the latter is shifted downwards by 3 basis points. The non-uniformity of these shifts is consistent with Erceg et al. (2005), who show that the way in which small sample bias of estimated AR coefficients translates to structural parameters estimated using long-run restrictions is dependent on the estimated econometric model. Said differently, downward biased AR coefficients do not necessarily translate into downward bias structural parameters when applying long-run restrictions. The contemporaneous response of the SVAR method is instead -36 basis points and 74 basis points, respectively.

Figure 7 displays only $\mathbb{E}\rho\left(\widehat{IR}, IR, \overline{h}\right)$ as $\rho\left(\mathbb{E}\widehat{IR}, IR, \overline{h}\right)$ is quantitatively similar. Comparison of Figure 7b and Figure 3b highlights one apparent advantage of the SVAR method. Because the bias of the SVAR with labor in levels at long forecast horizons is small, its efficiency gains outweigh the bias reductions of the LRLPs at long horizons in this specification. In practice, however, the true bias of each estimator is unknown and the response of hours need not decay to 0 at long horizons. Furthermore, LRLPs outperform the SVAR in terms of mean squared error at forecast horizons

Figure 7: Integrated Mean Squared Error without Small Sample Bias



(a) Labor in Differences: $\mathbb{E}\rho\left(\widehat{IR}, IR, \bar{h}\right)$



(b) Labor in Levels: $\mathbb{E}\rho\left(\widehat{IR}, IR, \bar{h}\right)$

of most interest to business cycle researchers (i.e. several years following a shock), and in terms of bias at all forecast horizons herein considered. Given these considerations, the superior bias properties of the LRLP method should be preferred to the efficiency gains of SVARs.

4.3 Choosing s_{max}

Up until now, I have remained silent on how to choose the free parameter, s_{max}, used in the construction of the contemporaneous correlation matrix. While it is possible to use a variety of information criteria to choose the lag length for autoregressive econometric models, the same is not true for the case of s_{max}.[20] Traditional information criteria choose a model parameter value, e.g. the lag length, to minimize some function of the model likelihood function with a penalty term for over-parameterization. Because these criteria depend on the likelihood function, they require the ability to compare observation values to those predicted by the estimated econometric model. In the case of choosing s_{max}, this would require comparing the contemporaneous correlation matrix implied by the estimated LRLP model to the "observed" contemporaneous correlation matrix. This object, however, is unobservable and is itself the subject of the estimator herein proposed. Thus, no comparison is feasible and standard information criteria are not applicable in this context.

Instead, recognize that increasing s_{max} implies a bias variance tradeoff. Figure 6 shows that the bias in the estimated contemporaneous correlation matrix decreases as s_{max} rises. Increasing the number of estimated parameter values used in Eq. 9, however, increases the uncertainty in their sum and therefore the implied contemporaneous correlation matrix. To balance this tradeoff,

[20] Hacker and Hatemi-J (2008), Ozcicek and Mcmillin (1999), and Lütkepohl (1985) provide a review and comparisons of the selection criteria typically employed.

practitioners can therefore estimate the contemporaneous correlation matrix for a range of s_{max} and choose the specification with the minimum s_{max} such that the estimated contemporaneous correlation matrix becomes both qualitatively and quantitatively similar to its behavior as s_{max} becomes large. Note that in large samples, an arbitrarily large s_{max} should be chosen as the bias-variance tradeoff is quantitatively negligible.

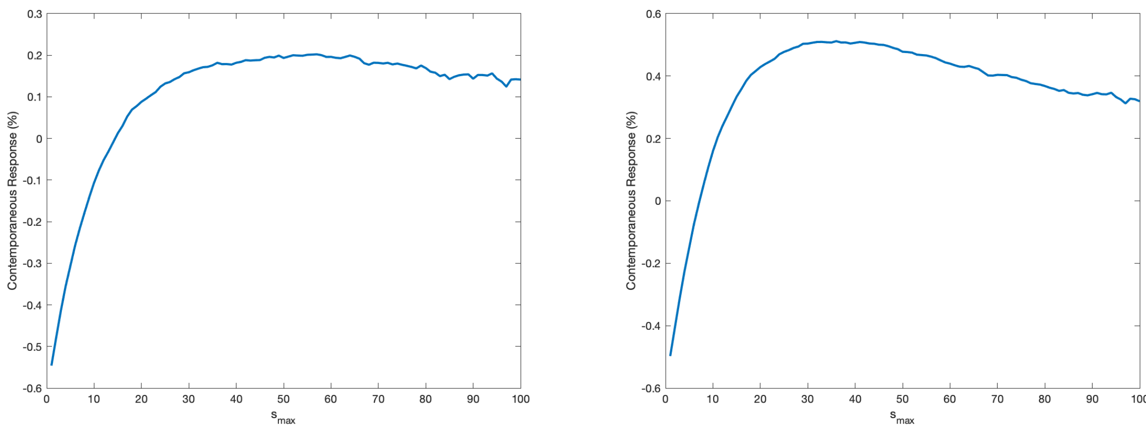
Figure 8 shows this procedure for the benchmark specification in the empirically relevant sample size discussed above and used in Section 5. The first panel shows the mean behavior across simulations for the contemporaneous response as a function of s_{max} when labor is included in differences. The second panel shows the same when labor is included in levels. When s_{max} is low, both specifications incorrectly suggest that hours decline on impact, a finding echoed by a large strand of the literature. As s_{max} increases, both specifications quickly estimate a contemporaneous response that is qualitatively similar to the model implied response, i.e. that hours initially rise in response to a positive productivity shock. In fact, the estimated response becomes positive after increasing the maximum forecast horizon to only 10-15 quarters depending on whether hours are included in levels or differences. In the differences specification, the contemporaneous response plateaus at a maximum forecast horizon of roughly 25 quarters. In the levels specification, the contemporaneous response also briefly plateaus at around 25 quarters and subsequently slowly declines to become quantitatively indistinguishable from the model implied response. Figure 8 suggests that restricting s_{max} to be between 20 and 30 quarters yields a good approximation to the model implied contemporaneous response both qualitatively and quantitatively. Moreover, a comparison of both specifications suggests that the true response–32 basis points–is between approximately +20 and +40 basis points. In contrast, comparison of the SVAR specifications yields a range of roughly -40 basis points to +60 basis points. Both of these facts serve to further highlight the improvements of the proposed estimator over the traditional SVAR approach discussed in the previous two subsections.

5 Estimating the Response of Aggregate Hours

Having established the advantages of the proposed estimator, I now revisit the large literature estimating the response of aggregate hours to a productivity shock. As previously discussed, Galí (1999), Christiano et al. (2003), Fernald (2007), Francis and Ramey (2005), Galí and Rabanal (2005), Canova et al. (2010), and Saijo (2019) all present estimates from a structural VAR.[21] Christiano et al. (2004), Basu et al. (2006), and Sims (2011) embed a constructed TFP series in a VAR directly rather than identifying TFP shocks using the estimated VAR itself and come to

[21] Saijo (2019) also presents estimates using the local projections method using the TFP series constructed by Fernald (2014).

Figure 8: Contemporaneous Response vs. s_{max}



(a) Labor in Differences (b) Labor in Levels

conflicting conclusions. With the exception of Christiano et al. (2003) and Christiano et al. (2004), the overwhelming conclusion of these papers is that aggregate hours decline–at least on impact–in response to an unanticipated rise in productivity.[22] This contrasts with RBC models à la Kydland and Prescott (1982), King and Rebelo (1999), and the extensions thereof, which predict a rise in hours in response to a positive technology shock. The direct, and often explicit, implication of this literature is therefore that this class of models is unreliable.[23]

Because much of the evidence to date has been obtained by imposing long-run restrictions on estimated VARs, it is subject to the critiques previously discussed. I therefore, estimate the response of aggregate hours to a productivity shock using the proposed estimator. I obtain output per hour (labor productivity) and labor hours data for the non-farm private business sector spanning the 1948Q1-2019Q2 period from the BLS. Labor hours are normalized by the non-institutionalized civilian population aged 16 and over in any given quarter.[24] Labor productivity is first differenced in all empirical specifications. There has been significant debate in the literature as to whether or not labor hours should enter in levels or in first differences with no clear consensus to date. While Dickey-Fuller tests fail to reject the null hypothesis that log-labor hours have a unit root, I do not take a stand on which approach is correct and instead present results when logged hours are included in both levels and first differences. Time series of the logarithm of the raw data are shown in Figure 9.

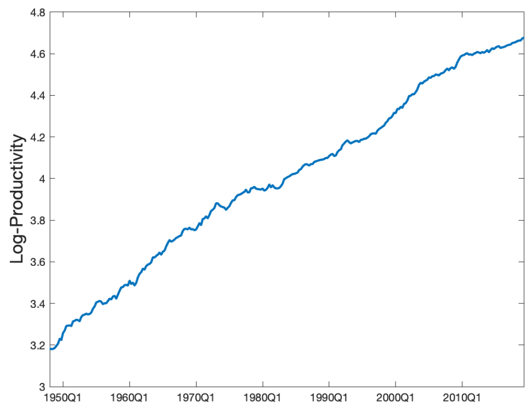
Furthermore, Fernald (2007) and Canova et al. (2010) show that controlling for low-frequency

[22]Sims (2011) finds that hours rise in response to temporary TFP shocks but fall in response to permanent TFP shocks.

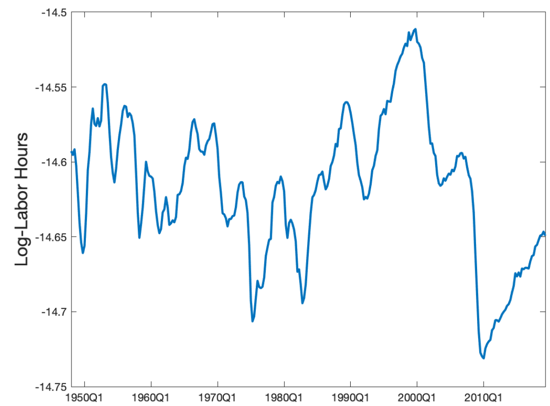
[23]See Kilian and Lütkepohl (2017) for a more detailed summary of the methodological components of this debate.

[24]The series ID of the BLS data used are PRS85006093, PRS85006033, and LNU00000000, respectively.

Figure 9: BLS Data



(a) Log-Labor Productivity



(b) Log-Labor Hours per Capita

changes in labor productivity growth has first order effects on estimated impulse response functions. To control for such considerations, I test for potential differences in subsample means of labor productivity growth in the following way. First, a Chow (1960) test is conducted on the original data for each potential break date until the first break date is found. A second Chow test is then performed on every possible break date after including the previous break dates as regressors. This process is continued until no new break dates are found. Once all break dates have been found, I de-mean labor productivity growth for each subsample (i.e. between break dates). Two break dates were found using this approach: 1973Q1 and 1974Q3.[25]

For comparison, I estimate the response of hours using both the proposed estimator and an SVAR. The lag-length of the local projections is chosen to match that in the SVAR using the Akaike (1974) information criteria with a maximum possible lag length of 10. While some have argued that the the Schwarz (1978) information criteria outperforms alternatives (See e.g. Hacker and Hatemi-J, 2008; Lütkepohl, 1985), long-run restrictions are sensitive to lag-truncation bias. As a result, the over parameterized specifications typically chosen by the Akaike (1974) information criteria may be preferable when relying on long-run restrictions for identification. Both information criteria yield the same result in this case. The imposed maximum possible lag length is not binding for any specification.

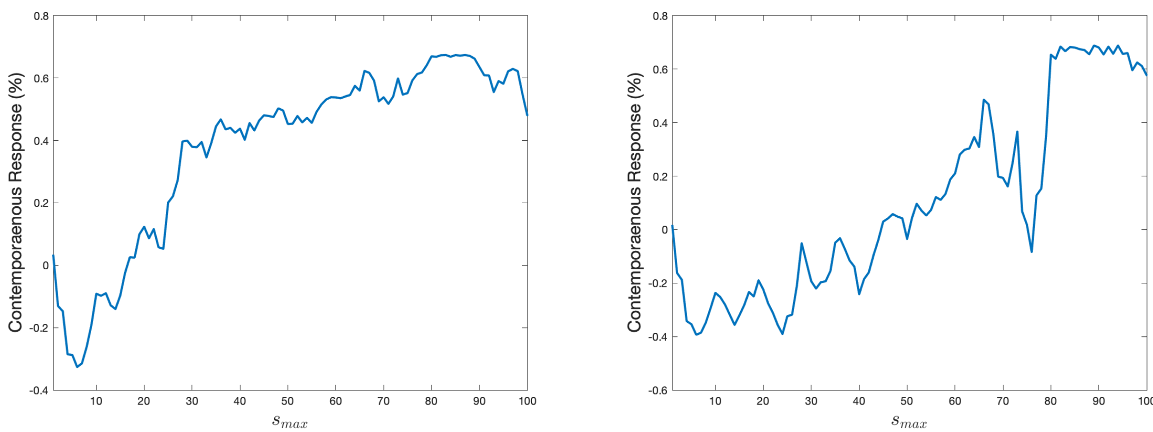
I construct 90% bootstrapped confidence for the SVARs using the bias-corrected bootstrap algorithm described in Kilian (1998). I use 1,000 replications to estimate the bias in the first step

[25]Using the partial sample Wald statistic test suggested in Andrews (1993) and Andrews (2003) does not yield a statistically significant break date using the critical values therein presented, though 1972Q4 is close to significant for some presented critical values and virtually identical to first break date implied by the iterative Chow (1960) procedure. Imposing a break only in 1972Q4 does not qualitatively change my results.

and 2,000 replications to estimate the confidence bands in the second step. An appropriate method for bootstrapping local projections is less clear. While Jordà (2009) discusses the importance of developing bootstrap methods for the local projections method, a blocks-of-blocks bootstrap approach for each forecast horizon, s, described in Kilian and Kim (2011) is the only proposal to date. Unlike this paper, their estimates of the contemporaneous correlation matrix are obtained from a VAR. In addition, this block-of-blocks bootstrap does no adequately consider the correlation between local projections coefficients across forecast horizons. Moreover, Kilian and Kim (2011) show through Monte-Carlo simulation that their proposed bootstrap may yield less than nominal coverage. This issue is more or less severe depending on the data generating process used to generate the Monte-Carlo samples. In light of these shortcomings, I instead bootstrap the long-run local projections by drawing from the asymptotic joint distribution of the reduced form coefficients and applying the same bias correction method as for the bootstrapped SVARs.

Finally, as described in Section 2, I must specify a choice of s_{max} to recover the structural parameters and therefore the impulse response function. As discussed previously, the choice of s_{max} directly affects the number of moving average coefficients included in the construction of the long-run impact of each shock. If s_{max} is too small, I may be missing important dynamic relationships of the data generating process. Thus, I follow the heuristic approach discussed in Section 4.3 and estimate the contemporaneous correlation matrix for a wide range of s_{max}. Figure 10 shows the estimated contemporaneous response for each specification.

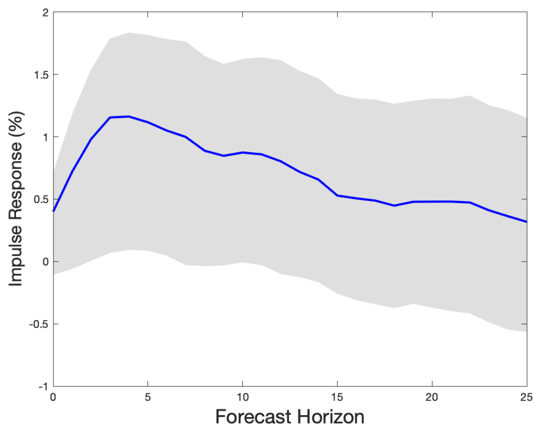
Figure 10: Sensitivity of Results to s_{max}



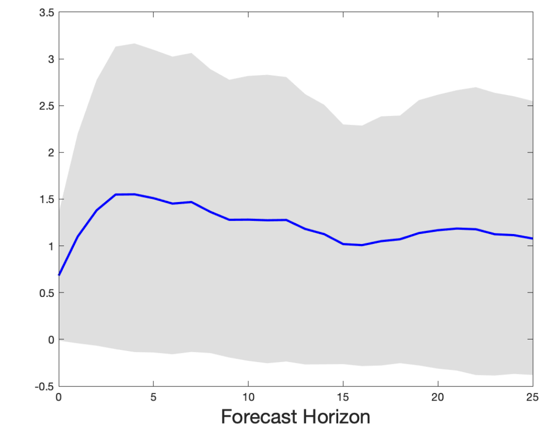
(a) Labor in Levels (b) Labor in Differences

The first column of Figure 10 shows the contemporaneous response when log-labor is included in levels. The second column shows the estimated contemporaneous response when labor is included in first differences. The overall relationship between s_{max} and the estimated contemporaneous response depicted above suggests that the contemporaneous response rises with s_{max}. The nonsmoothness of this relationship results from the fact that the AR coefficients of the local projections method can be choppy relative to their VAR counterparts, a fact highlighted in Ramey (2016). Despite these fluctuations, the tendency of the contemporaneous response to rise with s_{max} is clear. In the levels specification, a maximum forecast horizon of 30 is chosen. A larger s_{max} is required for the differences specification. In this case, I choose a maximum forecast horizon of 85, though a less conservative choice of 65 provides quantitatively similar results.

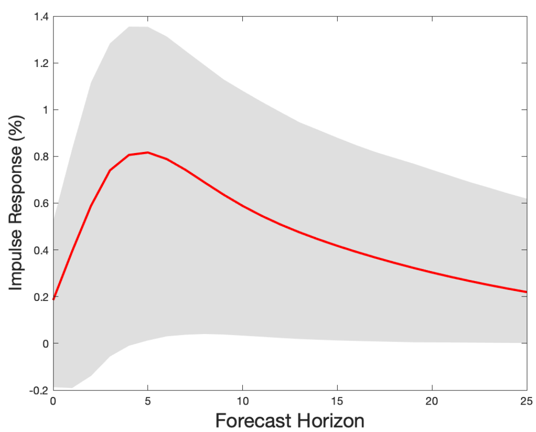
Figure 11: Estimated Impulse Response of Hours to a Productivity Shock



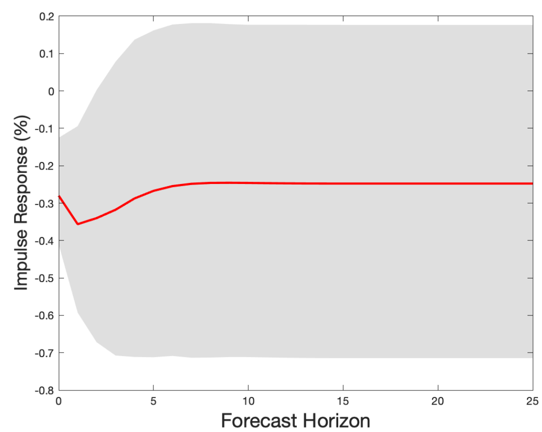
(a) LRLP Method: Labor in Levels



(b) LRLP Method: Labor in Differences



(c) SVAR Method: Labor in Levels

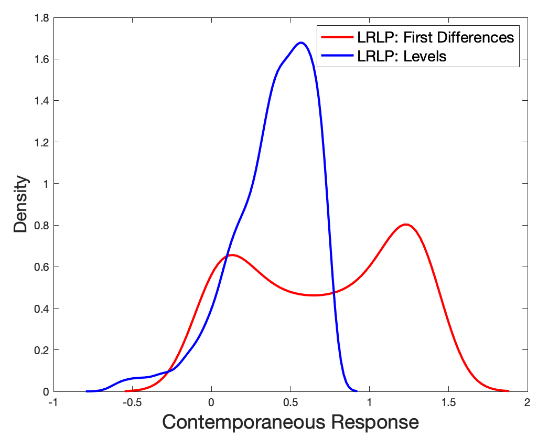


(d) SVAR Method: Labor in Differences

Figure 11 shows the estimated impulse response of hours to a technology shock. The first column shows results for the long-run local projections and SVAR in levels and the second column shows the results for the these two estimators when log-labor hours is included in first differences. The results from the proposed estimator in Figure 11a and Figure 11b show that labor hours rise on impact in response to a positive technology shock and follow a hump-shaped response thereafter regardless of whether labor hours are included in levels or first differenced. The striking feature of the proposed estimator is that, regardless of the chosen specification, the estimated response of hours is qualitatively the same. The SVAR method on the other hand predicts two qualitatively different responses depending on the specification used. The SVAR in levels, however, tightly matches both the level LRLP and the differenced LRLP. These results echo the findings of Christiano et al. (2003) that conclusions drawn from the SVAR with labor hours in levels should be preferred.

In all cases except for the SVAR in first differences, the bootstrapped confidence bands for the contemporaneous response contain 0. Hence, I cannot reject the hypothesis that hours do not respond, or even feature a very small negative response, to a positive productivity shock. Still there is substantially more probability mass to the right of zero than to the left of zero. Figure 12 shows the bootstrapped distribution of the contemporaneous response for the long-run local projections. For the long-run local projections in levels, approximately 91.8% of the probability mass lies above 0. For the long-run local projections with hours included in first differences, approximately 90.6% of the probability mass lies above 0. Taken together, the results of this section provide new evidence that hours in fact rise in response to a technology shock and that the standard RBC model may in fact be consistent with the data.

Figure 12: Bootstrapped Distribution of Contemporaneous Response



6 Conclusion

In this paper, I extend the local projections method to identify structural shocks through long-run restrictions. I show that my proposed estimator yields significant reductions in bias relative to SVARs both on impact and for most forecast horizons. Using Monte-Carlo evidence, I show that the proposed estimator is less sensitive than standard SVARs identified with long-run restrictions to the choice of included lag length and order of integration of the endogenous variables. Moreover, using my proposed estimator, I provide evidence that, in contrast to much of the evidence based on SVARs, aggregate labor hours rise in response to positive productivity shocks and follow a hump shaped profile thereafter. In fact, I show that over 90% of the bootstrapped probability mass indicates that hours rise on impact in response to a positive productivity shock. This result provides new empirical support for the standard real business cycle model.

I also highlight several previously unexplored issues of importance for future research. First, the bias reductions of the proposed estimator are illustrated only in the context of long-run restrictions. Structural identification in empirically relevant sample sizes using, for example, sign restrictions may also be improved upon by relying on local projections rather than estimated VARs. Additionally, current methods of bootstrapping with time series data either perform poorly for local projections or do not appropriately accommodate structural identification using local projections. Both of these issues require additional research.

References

Akaike, Hirotugu, "A new look at the statistical model identification," in "Selected Papers of Hirotugu Akaike," Springer, 1974, pp. 215–222.

Andrews, Donald W. K., "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 1993, *61* (4), 821–856.

_ , "Tests for Parameter Instability and Structural Change with Unknown Change Point: A Corrigendum," *Econometrica*, 2003, *71* (1), 395–397.

Basu, Susanto, John G. Fernald, and Miles S. Kimball, "Are Technology Improvements Contractionary?," *American Economic Review*, December 2006, *96* (5), 1418–1448.

Blanchard, Olivier Jean and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review*, 1989, *79* (4), 655–673.

Brugnolini, Luca, "About Local Projection Impulse Response Function Reliability," Working Paper 440, CEIS 2018.

Canova, Fabio, David Lopez-Salido, and Claudio Michelacci, "The effects of technology shocks on hours and output: a robustness analysis," *Journal of Applied Econometrics*, 2010, *25* (5), 755–773.

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan, "Business Cycle Accounting," *Econometrica*, 2007, *75* (3), 781–836.

Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan, "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Monetary Economics*, 2008, *55* (8), 1337 – 1352.

Choi, Chi-Young and Alexander Chudik, "Estimating impulse response functions when the shock series is observed," *Economics Letters*, 2019, *180*, 71 – 75.

Chow, Gregory C., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 1960, *28* (3), 591–605.

Christiano, Lawrence J, Martin Eichenbaum, and Robert Vigfusson, "What Happens After a Technology Shock?," Working Paper 9819, National Bureau of Economic Research July 2003.

Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson, "The Response of Hours to a Technology Shock: Evidence based on Direct Measures of Technology," *Journal of the European Economic Association*, 2004, *2* (2-3), 381–395.

_ , _ , and _ , "Alternative Procedures for Estimating Vector Autoregressions Identified with Long-run Restrictions," *Journal of the European Economic Association*, 2006, *4* (2/3), 475–483.

Cooley, Thomas F. and Mark Dwyer, "Business cycle analysis without much theory A look at structural VARs," *Journal of Econometrics*, 1998, *83* (1), 57 – 88.

Erceg, Christopher J., Luca Guerrieri, and Christopher Gust, "Can Long-Run Restrictions Identify Technology Shocks?," *Journal of the European Economic Association*, 2005, *3* (6), 1237–1278.

Faust, Jon and Eric M. Leeper, "When Do Long-Run Identifying Restrictions Give Reliable Results?," *Journal of Business and Economic Statistics*, 1997, *15* (3), 345–353.

Fernald, John, "A quarterly, utilization-adjusted series on total factor productivity," Working Paper 2012-19, Federal Reserve Bank of San Francisco 2014.

Fernald, John G., "Trend breaks, long-run restrictions, and contractionary technology improvements," *Journal of Monetary Economics*, 2007, *54* (8), 2467 – 2485.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson, "ABCs (and Ds) of Understanding VARs," *American Economic Review*, June 2007, *97* (3), 1021–1026.

Francis, Neville and Valerie A. Ramey, "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited," *Journal of Monetary Economics*, 2005, *52* (8), 1379 – 1399.

Galí, Jordi and Pau Rabanal, "Technology shocks and aggregate fluctuations: How well does the real business cycle model fit postwar US data?," *NBER macroeconomics annual*, 2005, *19*, 225–288.

Galí, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, March 1999, *89* (1), 249–271.

Garín, Julio, Robert Lester, and Eric Sims, "Are Supply Shocks Contractionary at the ZLB? Evidence from Utilization-Adjusted TFP Data," *The Review of Economics and Statistics*, March 2019, *101* (1), 160–175.

Hacker, Scott R and Abdulnasser Hatemi-J, "Optimal lag-length choice in stable and unstable VAR models under situations of homoscedasticity and ARCH," *Journal of Applied Statistics*, 2008, *35* (6), 601–615.

Hamilton, James D., "Historical Causes of Postwar Oil Shocks and Recessions," *The Energy Journal*, 1985, *0* (Number 1), 97–116.

Hurwicz, Leonid, "Least squares bias in time series," *Statistical inference in dynamic economic models*, 1950, *10*, 365–383.

Ingram, Beth Fisher, Narayana R. Kocherlakota, and N.E. Savin, "Explaining business cycles: A multiple-shock approach," *Journal of Monetary Economics*, 1994, *34* (3), 415 – 428.

Ireland, Peter N., "A method for taking models to the data," *Journal of Economic Dynamics and Control*, 2004, *28* (6), 1205 – 1226.

Jordà, Oscar, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, March 2005, *95* (1), 161–182.

_ , "Simultaneous Confidence Regions for Impulse Responses," *The Review of Economics and Statistics*, 2009, *91* (3), 629–647.

_ **and Sharon Kozicki**, "Estimation and Inverence by the Method of Projections Minimum Distance: An Application to the New Keynesian Hybrid Phillips Curve," *International Economic Review*, 2011, *52* (2), 461–487.

Kascha, Christian and Karel Mertens, "Business cycle analysis and VARMA models," *Journal of Economic Dynamics and Control*, 2009, *33* (2), 267 – 282.

Kilian, Lutz, "Small-Sample Confidence Intervals for Impulse Response Functions," *The Review of Economics and Statistics*, 1998, *80* (2), 218–230.

_ **and Helmut Lütkepohl**, *Estimation Subject to Long-Run Restrictions* Themes in Modern Econometrics, Cambridge University Press,

_ **and Yun Jung Kim**, "How Reliable Are Local Projection Estimators of Impulse Responses?," *The Review of Economics and Statistics*, 2011, *93* (4), 1460–1466.

King, Robert, Charles Plosser, James Stock, and Mark Watson, "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 1991, *81* (4), 819–40.

King, Robert G and Sergio T Rebelo, "Resuscitating real business cycles," *Handbook of macroeconomics*, 1999, *1*, 927–1007.

Kydland, Finn E. and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, 1982, *50* (6), 1345–1370.

Lippi, Marco and Lucrezia Reichlin, "VAR analysis, nonfundamental representations, blaschke matrices," *Journal of Econometrics*, 1994, *63* (1), 307 – 325.

Lütkepohl, Helmut, "Comparison of criteria for estimating the order of a vector autoregressive process," *Journal of time series analysis*, 1985, *6* (1), 35–52.

McGrattan, Ellen R., "Measurement with minimal theory," *Federal Reserve Bank of Minneapolis Quarterly Review*, 2010, *33* (1), 2–13.

Ozcicek, Omer and W.D̃ouglas Mcmillin, "Lag length selection in vector autoregressive models: symmetric and asymmetric lags," *Applied Economics*, 1999, *31* (4), 517–524.

Plagborg-Møller, Mikkel and Christian K. Wolf, "Local Projections and VARs Estimate the Same Impulse Responses," Working Paper 2019.

Prescott, Edward C., "Theory ahead of business-cycle measurement," *Carnegie-Rochester Conference Series on Public Policy*, 1986, *25*, 11 – 44.

Ramey, Valerie A., "Identifying Government Spending Shocks: It's all in the Timing*," *The Quarterly Journal of Economics*, 02 2011, *126* (1), 1–50.

Ramey, Valerie A, "Macroeconomic shocks and their propagation," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 71–162.

Ravenna, Federico, "Vector autoregressions and reduced form representations of DSGE models," *Journal of Monetary Economics*, 2007, *54* (7), 2048 – 2064.

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

Saijo, Hikaru, "Technology shocks and hours revisited: Evidence from household data," *Review of Economic Dynamics*, 2019, *31*, 347 – 362.

Schwarz, Gideon, "Estimating the Dimension of a Model," *The Annals of Statistics*, 1978, *6* (2), 461–464.

Shapiro, Matthew D. and Mark W. Watson, "Sources of Business Cycle Fluctuations," *NBER Macroeconomics Annual*, 1988, *3*, 111–148.

Sims, Eric, "Permanent and transitory technology shocks and the behavior of hours: A challenge for DSGE models," *Unpublished manuscript. University of Notre Dame, Notre Dame, IN*, 2011.