



Federal Reserve
Bank of Dallas

Estimation of Impulse Response Functions When Shocks are Observed at a Higher Frequency than Outcome Variables

Alexander Chudik and Georgios Georgiadis

Globalization Institute Working Paper 356

Research Department

<https://doi.org/10.24149/gwp356>

Working papers from the Federal Reserve Bank of Dallas are preliminary drafts circulated for professional comment. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Estimation of Impulse Response Functions When Shocks are Observed at a Higher Frequency than Outcome Variables*

Alexander Chudik[†] and Georgios Georgiadis[‡]

March 15, 2019

Abstract

This paper proposes mixed-frequency distributed-lag (MFDL) estimators of impulse response functions (IRFs) in a setup where (i) the shock of interest is observed, (ii) the impact variable of interest is observed at a lower frequency (as a temporally aggregated or sequentially sampled variable), (iii) the data-generating process (DGP) is given by a VAR model at the frequency of the shock, and (iv) the full set of relevant endogenous variables entering the DGP is unknown or unobserved. Consistency and asymptotic normality of the proposed MFDL estimators is established, and their small-sample performance is documented by a set of Monte Carlo experiments. The proposed approach is then applied to estimate the daily pass-through of changes in crude oil prices observed at a daily frequency to U.S. gasoline consumer prices observed at a weekly frequency. We find that the pass-through is fast, with about 28% of the crude oil price changes passed through to retail gasoline prices within five working days, and that the speed of the pass-through has increased over time.

Keywords: Mixed frequencies, temporal aggregation, impulse response functions, estimation and inference, VAR models.

JEL Classification: C22

*We would like to thank Lutz Kilian, Karel Mertens, and Michael Plante for helpful comments on earlier versions of this paper. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas, Federal Reserve System or the European Central Bank.

[†]Alexander Chudik, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, Texas, alexander.chudik@dal.frb.org.

[‡]Georgios Georgiadis, European Central Bank, 60640 Frankfurt am Main, Germany, georgios.georgiadis@ecb.int.

1 Introduction

This paper is concerned with estimating impulse-response functions (IRFs) when the outcome variable of interest is observed at a lower frequency than the shock of interest. One empirical question that can be addressed in this framework, and that is explored in this paper, is the daily pass-through of changes in crude oil price shocks observed at the daily frequency on U.S. retail gasoline prices (at the pump) observed only at the weekly frequency. There are naturally other interesting empirical questions that can be addressed with the framework proposed in this paper. For example, many monetary policy shock measures are constructed at the monthly frequency while real GDP measures, and in some economies even inflation, are only observed at the quarterly frequency. The approach we propose in this paper is also useful because the commonly used method of temporally aggregating shocks to match the frequency at which the outcome variable is observed yields inconsistent estimates of the temporally aggregated impulse response coefficients.

The problem approached in this paper relates mainly to two strands of the literature. The first strand of the literature considers estimation of IRFs when the shock of interest is observed, which is relatively straightforward and a number of different approaches can be used to accomplish this, such as the distributed lag or (vector) autoregressive distributed lag specifications (Kilian, 2008a and 2008b, and Romer and Romer, 2010, among others), or the local-projection approach by Jorda (2005).¹ The approaches considered in this literature all assume that the data on the shocks and the outcome variables are observed at the same frequency. Unfortunately, as mentioned above, it is sometimes the case that data on shock measures are observed at a higher frequency than those on outcome variables. Against this background, in this paper we propose a framework which allows to estimate consistently the high-frequency effects of a shock that is observed at a higher frequency than the outcome variable.

The second strand of the literature is concerned with the modelling of mixed-frequency data, see surveys by Forni and Marcellino (2013) and Forni, Ghysels, and Marcellino (2013), and the more recent contributions by Forni and Marcellino (2016) and Bacchiocchi, Bastianin, Missale, and Rossi (2018). While conceptually very closely related, the question of estimating IRFs at the high frequency when the shock is observed at a higher frequency than the outcome variable of

¹Choi and Chudik (2019) provide Monte Carlo comparisons of these approaches.

interest has not yet been considered, to the best of our knowledge. We borrow from this strand of the literature the idea that the true IRFs can be approximated by a flexible function that features a small number of unknown parameters. We document that while not necessary for consistent estimation of the IRFs, such parametrization can substantially improve small sample performance.

More specifically, we propose two mixed-frequency distributed lag (MFDL) estimators: While the *unrestricted* MFDL estimator does not impose any functional form on the IRF, the *restricted* MFDL estimator assumes that the true IRFs can be approximated by a flexible function that features a small number of unknown parameters and is therefore more parsimonious. We establish asymptotic normality for both estimators, and document their small sample performance by means of Monte Carlo experiments. The results from the Monte Carlo experiments suggest that the unrestricted MFDL estimator can in some cases be quite noisy for empirically relevant sample sizes. In contrast, the restricted MFDL estimator performs reasonably well.

To illustrate the usefulness of the proposed estimators, we explore the daily pass-through of changes in crude oil prices observed at the daily frequency to U.S. retail gasoline prices (at the pump) observed at the weekly frequency. Using the full sample, we find that the crude oil price pass-through is quite fast, with 28% of the crude oil price changes passed through to retail gasoline prices within five working days. Sub-sample rolling-window estimations suggest that the pass through has changed over time, with a faster pass-through observed in more recent periods compared with the 1990s. In the most recent sub-sample, almost 40% of the pass-through is estimated to materialize within five working days.

The rest of this paper is organized as follows. Section 2 introduces the model, defines the IRFs, and derives asymptotic properties of the MFDL estimators. Section 3 presents Monte Carlo experiments. Section 4 presents estimates of the daily crude oil price pass-through, and the last section concludes. Proofs, extensions, summary of notations and additional supplementary material are provided in the Appendix.

2 The MFDL estimators

This section first describes the data-generating process (DGP) and the main assumptions (Subsection 2.1), and then defines the objects of interests (Subsection 2.2), before proposing the MFDL

estimators and providing their asymptotic properties (Subsection 2.3).

2.1 Data-generating process and assumptions

Consider the following DGP for the $n \times 1$ dimensional vector of variables \mathbf{z}_t ,

$$\mathbf{z}_t = \mathbf{\Phi}\mathbf{z}_{t-1} + \mathbf{u}_t, \quad \text{for } t = 0, 1, 2, \dots, T, \quad (1)$$

where $\mathbf{\Phi}$ is an $n \times n$ matrix of coefficients, \mathbf{u}_t is an $n \times 1$ vector of reduced-form errors. For simplicity of exposition but without any loss of generality, one lag and no deterministic terms in (1) are assumed. Introducing additional lags and/or deterministic terms in (1) is straightforward but bears a loss of notational and expositional clarity. The DGP given by (1) is a standard VAR(1) model, which can be obtained, for instance, as the solution of a log-linearized dynamic stochastic general equilibrium model describing the evolution of the endogenous variables.

Let \mathbf{z}_t be, without any loss of generality, partitioned as $\mathbf{z}_t = (x_t, \mathbf{q}_t)'$, where x_t is the outcome variable of interest and \mathbf{q}_t is an $(n - 1) \times 1$ vector of remaining variables. Notice that it is not assumed that the remaining variables in \mathbf{q}_t are observed. It is not assumed that n is known, as is often the case in empirical work, in which the choice of variables is a contentious issue. Furthermore, partition \mathbf{u}_t as

$$\mathbf{u}_t = \mathbf{a}e_t + \mathbf{v}_t, \quad (2)$$

where \mathbf{v}_t is the component of \mathbf{u}_t that is uncorrelated with the shock of interest e_t . Notice that in this paper we assume that the shock e_t in the decomposition (2) is observed, and therefore we do not contribute to the important problem of shock identification.

The decomposition given by (2) is quite general, and it nests two important examples. The first example is the structural model

$$\mathbf{u}_t = \mathbf{A}\boldsymbol{\xi}_t,$$

where $\boldsymbol{\xi}_t$ is an $l \times 1$ vector of structural shocks that contains e_t as one of its elements, and \mathbf{A} is an $n \times l$ matrix of coefficients, and the number of structural shocks, l , could be greater, equal, or smaller than the number of variables, n . In this example, let the vector of structural shocks, without any loss of generality, be partitioned as $\boldsymbol{\xi}_t = (e_t, \boldsymbol{\zeta}_t)'$. Partitioning accordingly $\mathbf{A} = [\mathbf{a}, \mathbf{A}_{-1}]$,

and assuming that structural shocks are mutually uncorrelated, we obtain (2) by noting that $\mathbf{u}_t = \mathbf{A}\boldsymbol{\xi}_t = \mathbf{a}e_t + \mathbf{A}_{-1}\boldsymbol{\zeta}_t = \mathbf{a}e_t + \mathbf{v}_t$, with $\mathbf{v}_t = \mathbf{A}_{-1}\boldsymbol{\zeta}_t$ being uncorrelated with e_t . The IRF of a unit shock to e_t (formally defined in Section 2.2 below) is the structural IRF of the structural shock e_t .

The second example is letting e_t be one of the reduced-form shocks in $\mathbf{u}_t = (u_{1t}, u_{2t}, \dots, u_{nt})'$. In particular, let $e_t = u_{jt}$ for some given $j \in \{1, 2, \dots, n\}$, and let $E(\mathbf{u}_t | e_t) = \mathbf{a}e_t$. Then, we can define $\mathbf{v}_t \equiv \mathbf{u}_t - E(\mathbf{u}_t | e_t)$. Hence, we obtain (2), where, by construction, e_t and \mathbf{v}_t are uncorrelated. The IRF of a unit shock to $e_t = u_{jt}$ in this example corresponds to the generalized IRF of a unit shock to u_{jt} , as considered by Koop et al. (1996) and Pesaran and Shin (1998).

We assume that the variable of interest, x_t , is observed only at a lower frequency than the shock e_t . Let $m > 1$ be the number of high-frequency time periods contained in one low-frequency period. For instance, if high frequency is monthly and low frequency is quarterly, then $m = 3$. Define

$$t_s = ms \tag{3}$$

and $T_m = [T/m] - 1$, where $[T/m]$ is the integer part of T/m . Thus, T is the number of high-frequency time periods for which e_t is observed, and T_m is the number of low-frequency time periods for which the temporally aggregated time series

$$\bar{x}_{w,s} = \sum_{h=0}^{m-1} w_h x_{t_s-h}, \text{ for } s = 1, 2, \dots, T_m, \tag{4}$$

is observed. For example, a simple averaging of the high-frequency observations contained in one low-frequency time period is represented by equal temporal aggregation weights $w_h = 1/m$, for $h = 0, 1, \dots, m - 1$. Sequential sampling is represented by setting one of the weights to 1 and others to 0. When low frequency is quarterly and high frequency is monthly, sequentially sampling the last month of each quarter is represented by $w_0 = 1$, and $w_1 = w_2 = 0$. Using the notation $t_s = ms$ allows us to conveniently refer to the high-frequency periods $t_s - (m - 1), t_s - (m - 2), \dots, t_s$ within the low-frequency period s . For example, when the low frequency is annual and the high-frequency is monthly, then $t_s - h$ for $h = 0, 1, \dots, 11$, refers to the h^{th} month of year s before December.

Remark 1 *The framework we suggest below can also be used for the case of more complex temporal*

aggregation schemes than in (4), for example $\bar{x}_{w,q,s} = \sum_{h=0}^{qm-1} w_{qh} x_{t_s-h}$, for $s = 1, 2, \dots, T_m$, and some fixed $q > 1$. The case of $q = 2$ can be of particular interest, since it represents the case of first-differences of the temporally aggregated data. This extension is relegated to Appendix A.3.

It is useful to define m sequentially sampled shock series:

$$\bar{e}_{sh} = e_{t_s-h}, \text{ for } h = 0, 1, 2, \dots, m-1. \quad (5)$$

It is clear that observations \bar{e}_{sh} , for $h = 0, 1, 2, \dots, m-1$, are available for $s = 1, 2, \dots, T_m$ and so is $\bar{x}_{w,s}$.

The following assumptions are postulated. We use c and K to denote generic small and large positive constants that do not depend on the sample size T . These constants can take different values at different instances in the paper.

ASSUMPTION 1 $|\lambda_1(\Phi)| \leq \rho < 1$, where $\lambda_1(\Phi)$ is the largest eigenvalue of Φ .

ASSUMPTION 2 Innovations \mathbf{u}_t are given by (2), where $e_t \sim IID(0, \sigma_e^2)$, $c < \sigma_e^2 < K$, and $\|\mathbf{a}\| < K$. In addition, $\mathbf{v}_t \sim IID(\mathbf{0}_{k \times k}, \Xi)$, and e_t is orthogonal to $v_{t'}$ for any $t \neq t'$. The fourth moments of e_t and individual elements of \mathbf{v}_t exist.

Assumption 1 is the standard stationarity condition on the coefficient matrix Φ , when n does not increase with T . Assumption 2 implies uncorrelatedness of e_t and \mathbf{v}_t , which is important for the consistency of the estimators proposed in this paper.

2.2 Objects of interest and further definitions

Under Assumptions 1-2, x_t is represented by the sum of the following two moving-average processes,

$$x_t = \sum_{\ell=0}^{\infty} b_{\ell} e_{t-\ell} + \varepsilon_t, \quad (6)$$

where

$$b_{\ell} = \mathbf{s}'_{n,1} \Phi^{\ell} \mathbf{a}, \text{ for } \ell = 0, 1, 2, \dots, \quad (7)$$

$\mathbf{s}_{n,1} = (1, 0, \dots, 0)'$ is an $n \times 1$ selection vector that selects the first element, and

$$\varepsilon_t = \sum_{\ell=0}^{\infty} \mathbf{s}'_{n,1} \mathbf{\Phi}^{\ell} \mathbf{v}_{t-\ell}. \quad (8)$$

The sequence $\{b_{\ell}\}_{\ell=0}^{\infty}$ is the IRF of a unit shock to e_t on the variable of interest, x_t , at the high frequency at which the shock is observed, namely

$$b_{\ell} = E(x_{t+\ell} | e_t = 1, \mathcal{I}_{t-1}) - E(x_{t+\ell} | \mathcal{I}_{t-1}), \text{ for } \ell = 0, 1, \dots, \quad (9)$$

where $\mathcal{I}_{t-1} = \{\mathbf{z}_{t-1}, \mathbf{z}_{t-2}, \dots\}$ is an information set featuring all variables up to $t-1$. Our main goal is to estimate the high-frequency IRF coefficients b_{ℓ} defined by (9) and given by (7).

In order to estimate $\{b_{\ell}\}$, it is going to be useful to consider the impact of a shock observed at the high frequency on the low-frequency, temporally aggregated outcome variable, namely the coefficients

$$d_{rh} = E(\bar{x}_{w,s+r} | e_{t_s-h} = 1, \mathcal{I}_{t_s-h-1}) - E(\bar{x}_{w,s+r} | \mathcal{I}_{t_s-h-1}), \text{ for } r = 0, 1, \dots, \text{ and } h = 0, 1, \dots, m-1, \quad (10)$$

where t_s is given by (3). Note that the parameter h in (10) refers to the timing of the shocks within the low-frequency period, namely occurring h high-frequency periods before the end of the low-frequency period. The parameter r in (10) refers to the low-frequency horizon.

Using (4) and (9) in (10), we obtain the following mapping between the high-frequency IRFs $\{b_{\ell}, \ell = 0, 1, 2, \dots\}$, and the low-frequency IRFs $\{d_{rh}, r = 0, 1, 2, \dots\}$ for $h = 0, 1, \dots, m-1$.

$$d_{0,h} = \sum_{q=0}^h w_q b_{h-q}, \text{ for } h = 0, 1, \dots, m-1, \quad (11)$$

$$d_{r,h} = \sum_{q=0}^{m-1} w_q b_{mr+h-q}, \text{ for } r = 1, 2, \dots \text{ and } h = 0, 1, \dots, m-1. \quad (12)$$

2.3 Estimation of IRF coefficients

In order to propose estimators of the high-frequency IRFs $\{b_{\ell}\}$, we derive the following representation for $\bar{x}_{w,s}$.

Lemma 1 Let \mathbf{z}_t be given by model (1), and suppose Assumptions 1-2 hold. Consider the temporally aggregated variable $\bar{x}_{w,s} = \sum_{h=0}^{m-1} w_h x_{t_s-h}$ as defined in (4), where $t_s = ms$, w_h for $h = 0, 1, \dots, m-1$ are the temporal aggregation weights, and $m > 1$ is the number of high-frequency periods contained in one low-frequency period. Then,

$$\bar{x}_{w,s} = \sum_{h=0}^{m-1} \sum_{r=0}^{\infty} d_{rh} \bar{e}_{s-r,h} + \bar{\varepsilon}_{w,s}, \quad (13)$$

where d_{rh} is given by (11)-(12), \bar{e}_{sh} is defined in (5), namely $\bar{e}_{sh} = e_{t_s-h}$, for $h = 0, 1, 2, \dots, m-1$, and $\bar{\varepsilon}_{w,s} = \sum_{h=0}^{m-1} w_h \varepsilon_{t_s-h}$ with ε_t given by (8).

Representation (13) shows that the temporally aggregated variable $\bar{x}_{w,s}$ defined in (4) and observed at the low frequency can be written as a weighted sum of current and lagged shocks that are observed at the high frequency. The low-frequency IRF coefficients d_{rh} , given by (11)-(12), are functions of the temporal aggregation weights $\{w_h, h = 0, 1, \dots, m-1\}$ and of the high-frequency IRF coefficients $\{b_\ell, \ell = 0, 1, 2, \dots\}$.

For the estimation of the low-frequency IRF coefficients $\{d_{rh}\}$, we propose the following auxiliary MFDL regression based on a truncated version of (13),

$$\bar{x}_{w,s} = \sum_{h=0}^{m-1} \sum_{r=0}^p d_{rh} \bar{e}_{s-r,h} + \vartheta_{ps}, \quad (14)$$

where

$$\vartheta_{ps} = \bar{\varepsilon}_{w,s} + \sum_{h=0}^{m-1} \sum_{r=p+1}^{\infty} d_{rh} \bar{e}_{s-r,h}, \quad (15)$$

and p is a chosen truncation lag. Regression (14) can be compactly written as

$$\bar{x}_{w,s} = \mathbf{d}'_{(p)} \bar{\mathbf{e}}_{(p),s} + \vartheta_{ps}, \quad (16)$$

where $\mathbf{d}_{(p)} = (\mathbf{d}'_0, \mathbf{d}'_1, \dots, \mathbf{d}'_p)'$ is $(p+1)m \times 1$ dimensional vector of coefficients partitioned into $(p+1) \times 1$ dimensional subvectors

$$\mathbf{d}_\ell \equiv (d_{r,0}, d_{r,1}, \dots, d_{r,m-1})' \text{ for } r \geq 0, \quad (17)$$

and $\bar{\mathbf{e}}_{(p),s} = (\bar{\mathbf{e}}'_s, \bar{\mathbf{e}}'_{s-1}, \dots, \bar{\mathbf{e}}'_{s-p})'$ is $(p+1)m \times 1$ dimensional vector of sequentially sampled shock series, in which $\bar{\mathbf{e}}_s = (\bar{e}_{s,0}, \bar{e}_{s,1}, \dots, \bar{e}_{s,m-1})'$. The subscript (p) is used to distinguish $\mathbf{d}_{(p)}$ from its $p+1$ components \mathbf{d}_r , for $r = 0, 1, \dots, p$, defined in (17). It is clear that the low-frequency IRF coefficients d_{rh} , $h = 0, 1, \dots, m-1$ and $r = 0, 1, \dots, p$, can be consistently estimated by least squares (LS), since ϑ_{ps} is uncorrelated with $\bar{\mathbf{e}}_{(p),s}$. Moreover, a consistent estimate of the high-frequency IRF coefficients $\{b_\ell\}$ can be obtained from the consistent estimates of $\{d_{rh}\}$, using the mapping in (11)-(12). In particular, given the truncation lag order p in (14), the mapping between the low-frequency IRF coefficients $\{d_{rh}\}$ and the high-frequency IRF coefficients $\{b_\ell\}$ given by (11)-(12) can be compactly written as

$$\mathbf{d}_{(p)} = \begin{pmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_p \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{W}_b & \mathbf{0}_{m \times m} & \mathbf{0}_{m \times m} & \cdots & \mathbf{0}_{m \times m} \\ \mathbf{W}_a & \mathbf{W}_b & \mathbf{0}_{m \times m} & \cdots & \mathbf{0}_{m \times m} \\ \mathbf{0}_{m \times m} & \mathbf{W}_a & \mathbf{W}_b & & \mathbf{0}_{m \times m} \\ \vdots & & \ddots & \ddots & \\ \mathbf{0}_{m \times m} & \mathbf{0}_{m \times m} & \cdots & \mathbf{W}_a & \mathbf{W}_b \end{pmatrix}}_{\mathbf{W}_p} \cdot \underbrace{\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{(p+1)m} \end{pmatrix}}_{\mathbf{b}_p}, \quad (18)$$

where

$$\mathbf{W}_a \equiv \begin{pmatrix} 0 & w_{m-1} & w_{m-2} & w_{m-3} & \cdots & w_1 \\ 0 & 0 & w_{m-1} & w_{m-2} & \cdots & w_2 \\ 0 & 0 & 0 & w_{m-1} & & w_3 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & & w_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \mathbf{W}_b \equiv \begin{pmatrix} w_0 & 0 & 0 & \cdots & 0 \\ w_1 & w_0 & 0 & \cdots & 0 \\ w_2 & w_1 & w_0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ w_{m-1} & w_{m-2} & w_{m-3} & \cdots & w_0 \end{pmatrix}.$$

Assuming \mathbf{W}_p is invertible, using (18) we obtain

$$\mathbf{b}_p = \mathbf{W}_p^{-1} \mathbf{d}_{(p)}. \quad (19)$$

Remark 2 By the lower-triangularity of \mathbf{W}_p , $\det(\mathbf{W}_p) = w_0^{(p+1)m}$ it follows that \mathbf{W}_p is invertible

if and only if $w_0 \neq 0$. However, if $w_0 = w_1 = \dots = w_j = 0$ for some $0 < j < m$ in the temporal aggregation of the outcome variable x_t , then we can simply re-define \mathbf{W}_p by considering a modified version of the auxiliary regression (14) that does not include the j most recent regressors $\bar{e}_{s,0}, \bar{e}_{s,1}, \dots, \bar{e}_{s,j}$.

Consider the following unrestricted MFDL estimator of $\mathbf{b}_p = (b_0, b_1, \dots, b_{(p+1)m-1})'$,

$$\hat{\mathbf{b}}_p = \mathbf{W}_p^{-1} \hat{\mathbf{d}}_{(p)}, \quad (20)$$

where $\hat{\mathbf{d}}_{(p)}$ is the LS estimator based on the auxiliary regression (16) given by

$$\hat{\mathbf{d}}_{(p)} = (\mathbf{E}'\mathbf{E})^{-1} \mathbf{E}'\bar{\mathbf{x}}, \quad (21)$$

in which $\bar{\mathbf{x}} = (\bar{x}_{p+1}, \bar{x}_{p+2}, \dots, \bar{x}_{T_m})'$ is the vector of observations on the dependent variable $\bar{x}_{w,s}$ in (16), and \mathbf{E} is the matrix of observations on $\bar{\mathbf{e}}_{(p),s}$ in (16), namely $\mathbf{E} = (\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_p)$, where $\mathbf{E}_\ell = (\bar{\mathbf{e}}'_{p-\ell+1}, \bar{\mathbf{e}}'_{p-\ell+2}, \dots, \bar{\mathbf{e}}'_{T_m-\ell})'$, for $\ell = 0, 1, \dots, p$, and $\bar{\mathbf{e}}'_s$, for $s = 1, 2, \dots, T_m$ is defined below (16). The dimensions of \mathbf{E} and $\bar{\mathbf{x}}$ depend on the sample size, T , as well as the truncation lag, p , but we omit these subscripts to simplify the notation. Similarly, the vectors $\hat{\mathbf{b}}_p$ and $\hat{\mathbf{d}}_{(p)}$ depend on T , but the subscript T is again not used to simplify the notation.

The asymptotic distribution of $\hat{\mathbf{b}}_p$ is established in the next theorem.

Theorem 1 *Let \mathbf{z}_t be generated by the DGP in (1), and suppose Assumptions 1-2 hold. Let the matrix \mathbf{W}_p defined in (18) be invertible, and consider the estimator $\hat{\mathbf{b}}_p$ defined in (20). Let $T \rightarrow \infty$, and suppose m and p are fixed. Then,*

$$\sqrt{T/m-p} \left(\hat{\mathbf{b}}_p - \mathbf{b}_p \right) \rightarrow^d N(\mathbf{0}, \mathbf{\Omega}_p), \quad (22)$$

where $\mathbf{b}_p = (b_0, b_1, \dots, b_{(p+1)m-1})'$ is the vector of high-frequency IRF coefficients defined in (9), and

$$\mathbf{\Omega}_p = \sigma_e^{-2} \mathbf{W}_p^{-1} \left(\sum_{\ell=-p}^p \gamma_{p\ell} \mathbf{H}_{p,\ell} \right) \mathbf{W}_p^{-1'}, \quad (23)$$

in which $\sigma_e^2 = E(e_t^2)$ is the variance of e_t , $\gamma_{p\ell} = E(\vartheta_{ps}\vartheta_{p,s-\ell})$ is the autocovariance function of

ϑ_{ps} defined in (15), $\mathbf{H}_{p,\ell} = \mathbf{G}_{p\ell} \otimes \mathbf{I}_m$, \mathbf{I}_m is an $m \times m$ identity matrix, $\mathbf{G}_{p\ell}$ is a $(p+1) \times (p+1)$ dimensional shift matrix with its (i, j) -th element given by $\delta_{i+\ell, j}$, and

$$\delta_{i,j} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}, \quad (24)$$

is the Kronecker delta.

The asymptotic variance matrix $\mathbf{\Omega}_p$ can be consistently estimated by

$$\hat{\mathbf{\Omega}}_p = \hat{\sigma}_e^{-2} \mathbf{W}_p^{-1} \left(\sum_{\ell=-p}^p \hat{\gamma}_{p\ell} \mathbf{H}_{p,\ell} \right) \mathbf{W}_p^{-1'}, \quad (25)$$

where $\hat{\sigma}_e^2 = T^{-1} \sum_{t=1}^T e_t^2$, and $\hat{\gamma}_{p\ell}$ is the sample autocovariance of $\hat{\vartheta}_{ps}$, namely

$$\hat{\gamma}_{p\ell} = \frac{1}{T_m - p} \sum_{s=p+1}^{T_m} \hat{\vartheta}_{ps} \hat{\vartheta}_{p,s-\ell}, \quad (26)$$

and $\hat{\vartheta}_{ps} = \bar{x}_{w,s} - \mathbf{W}_p^{-1} \hat{\mathbf{b}}'_{(p)} \bar{\mathbf{e}}_{(p),s}$ are the residuals from (16).

It is interesting to point out that using the temporally aggregated shock series

$$\bar{e}_{w,s} = \sum_{h=0}^{m-1} w_h e_{t_s-h}, \text{ for } s = 1, 2, \dots, T_m. \quad (27)$$

in a regression of the temporally aggregated outcome variable of interest $\bar{x}_{w,s}$ on $\{\bar{e}_{w,s}, \bar{e}_{w,s-1}, \dots, \bar{e}_{w,s-p}\}$ will in general not yield a consistent estimate of the temporally aggregated IRF defined as

$$\bar{b}_{w,s} = \sum_{h=0}^{m-1} w_h b_{sm-h}, \text{ for } s = 0, 1, \dots$$

We illustrate this in an example.

Example 1 Consider $m = 3$ and $p = 0$. Suppose the object of interest is the temporally aggregated impact response in the low frequency, namely $\bar{b}_{w,0} = w_2 b_0 + w_1 b_1 + w_0 b_2$. Using the aggregated shock $\bar{e}_{w,s}$ defined in (27) in a regression of $\bar{x}_{w,s}$ on $\bar{e}_{w,s}$ will not yield a consistent estimate of $\bar{b}_{w,0}$.

Instead, the inconsistency (large- T bias) of this estimator is

$$inc = \frac{E(\bar{e}_{w,s}\bar{x}_{w,s})}{E(\bar{e}_{w,s}^2)} - \bar{b}_{w,0},$$

which, after substituting the expression for $\bar{e}_{w,s}$ and (14) for $\bar{x}_{w,s}$ and after some algebra is given by

$$inc = b_0(1 - w_0) + b_1 \left(\frac{w_1 w_0 + w_2 w_1}{w_0^2 + w_1^2 + w_2^2} - w_1 \right) + b_2 \left(\frac{w_0}{w_0^2 + w_1^2 + w_2^2} - w_2 \right).$$

It can be seen that the sign and the magnitude of this inconsistency depend on the true high-frequency IRF coefficients $\{b_0, b_1, b_2\}$ and the aggregation weights $\{w_0, w_1, w_2\}$ in a non-trivial way. In the case in which $b_0 = 1$, $b_1 = \phi$, $b_2 = \phi^2$ and $w_0 = w_1 = w_2 = 1$, which corresponds to a DGP with an AR(1) with autoregressive parameter ϕ at the high frequency and the temporally aggregated variable being the sum of the data at the high frequency. In this case, for $\phi = 0.75$ we have $\bar{b}_{w,0} = 2.3125$ and $inc = -0.625$, which is sizeable.

2.3.1 Potential small sample drawbacks of the unrestricted MFDL estimator

As can be seen from (20) and (23), the inverse of the aggregation matrix \mathbf{W}_p is important for the asymptotic and small sample properties of the unrestricted MFDL estimator $\widehat{\mathbf{b}}_p$. It is useful to clarify the role of \mathbf{W}_p for two common aggregation schemes: end-of-period sequential sampling with $\mathbf{w} = (1, 0, 0, \dots, 0)'$, period sums with $\mathbf{w} = (1, 1, \dots, 1)'$ or period averages with $\mathbf{w} = (1/m, 1/m, \dots, 1/m)'$.

In the former case, \mathbf{W}_p is an identity matrix, which is a very convenient case since \mathbf{W}_p^{-1} also is an identity matrix so that $\widehat{\mathbf{b}}_p$ coincides with $\widehat{\mathbf{d}}_{(p)}$.

The latter case, however, is more challenging. Under $\mathbf{w} = \boldsymbol{\tau} = (1, 1, \dots, 1)'$, the first $m - 1$ subdiagonals and the main diagonal of the matrix \mathbf{W}_p consist of ones, which results in a rather inconvenient structure of \mathbf{W}_p^{-1} . In particular, the diagonal elements of $\mathbf{W}_p^{-1}\mathbf{W}_p'^{-1}$ linearly increase along the diagonal, meaning that the variance of the low-frequency IRF coefficient estimates increase with the IRF horizon ℓ . Given that the variance of $\widehat{\mathbf{b}}_p$ depends on $\mathbf{W}_p^{-1}\mathbf{W}_p'^{-1}$ in (23), this implies that the low-frequency IRF coefficients are less precisely estimated at longer horizons. This is illustrated in the upper part of Figure A1, which plots the diagonal elements of $\mathbf{W}_p^{-1}\mathbf{W}_p'^{-1}$ for

$m = 3$. Furthermore, for any selection vector \mathbf{s}_ℓ that selects the ℓ -th element, the vector $\mathbf{W}_p^{-1}\mathbf{s}_\ell$ exhibits a cyclical pattern from ℓ onward. This suggests that a one-off large sampling error for one of the elements in $\widehat{\mathbf{d}}_{(p)}$ can cause a cyclical pattern in $\widehat{\mathbf{b}}_p$ even if the true low-frequency IRF coefficients do not exhibit such a pattern. This is illustrated in the lower panel of Figure A1, which plots elements of $\mathbf{W}_p^{-1}\mathbf{s}_1$ as well as $\mathbf{W}_p^{-1}\boldsymbol{\tau}$ for $m = 3$. This cyclical accumulation of estimation errors in $\widehat{\mathbf{d}}_{(p)}$ can be an important small-sample drawback under some DGPs and sample sizes, as we document below in the Monte Carlo experiments.

Since the unrestricted MFDL estimator (22) can be noisy in small samples, we propose next a restricted MFDL estimator, which improves on the performance of the unrestricted estimator by assuming that the true IRFs can be well approximated by a sufficiently general function featuring a smaller set of unknown parameters.

2.3.2 Restricted MFDL estimator

We postulate the following assumption.

ASSUMPTION 3 *Let $\mathbf{b}_p = \mathbf{f}_p(\boldsymbol{\psi}_0)$, where $\mathbf{f}_p(\boldsymbol{\psi}) : \Theta \rightarrow \mathbb{R}^k$ is continuously differentiable, $k = (p + 1)m$, the support Θ is a compact set, and $\mathbf{f}_p(\boldsymbol{\psi}) = \mathbf{b}_p$ if and only if $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, where $\boldsymbol{\psi}_0$ is a $q \times 1$ vector of unknown parameters and $q \leq k$.*

We propose the minimum distance estimator

$$\hat{\boldsymbol{\psi}} = \arg \min_{\boldsymbol{\psi} \in \Theta} \left\| \widehat{\mathbf{b}}_p - \mathbf{f}_p(\boldsymbol{\psi}) \right\|_{\widehat{\boldsymbol{\Omega}}_p^{-1}}, \quad (28)$$

where $\widehat{\mathbf{b}}_p$ is the unrestricted MFDL estimator, Θ is a compact set of admissible values of $\boldsymbol{\psi}$, $\|\mathbf{x}\|_{\widehat{\boldsymbol{\Omega}}_p^{-1}} = \mathbf{x}'\widehat{\boldsymbol{\Omega}}_p^{-1}\mathbf{x}$, and $\widehat{\boldsymbol{\Omega}}_p$ is a consistent estimator of $\boldsymbol{\Omega}_p$ defined in (25), which we assume is invertible. The asymptotic distribution of $\mathbf{f}_p(\hat{\boldsymbol{\psi}})$ is established in the next theorem.

Theorem 2 *Let \mathbf{z}_t be generated by the DGP in (1), and suppose Assumptions 1-3 hold. Let the matrices \mathbf{W}_p and $\widehat{\boldsymbol{\Omega}}_p$, defined in (18) and (25), respectively, be invertible, and consider the estimator*

$$\tilde{\mathbf{b}}_p = \mathbf{f}_p(\hat{\boldsymbol{\psi}}), \quad (29)$$

where $\hat{\boldsymbol{\psi}}$ is given by (28). Let $T \rightarrow \infty$, and suppose m and p are fixed. Then,

$$\sqrt{T/m-p} \left(\tilde{\mathbf{b}}_p - \mathbf{b}_p \right) \rightarrow^d N(\mathbf{0}, \boldsymbol{\Sigma}_p), \quad (30)$$

where

$$\boldsymbol{\Sigma}_p = \mathbf{J}_p \left(\mathbf{J}'_p \boldsymbol{\Omega}_p^{-1} \mathbf{J}_p \right)^{-1} \mathbf{J}'_p, \quad (31)$$

\mathbf{J}_p is the $k \times q$ dimensional Jacobian of \mathbf{f}_p at $\boldsymbol{\psi}_0$,

$$\mathbf{J}_p \equiv \mathbf{J}_p(\boldsymbol{\psi}_0) = \left. \frac{\partial \mathbf{f}_p(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}'} \right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}, \quad (32)$$

and $\boldsymbol{\Omega}_p$ is a $k \times k$ variance matrix given by (23).

Remark 3 Any continuously differentiable function could be considered for $\mathbf{f}_p(\cdot)$. One possible choice is the function with elements given by the elements of the polynomial $\psi^{-1}(L, q)$, where

$$\psi(L, q) \equiv \psi_0 + \psi_1 L + \dots + \psi_{q-1} L^{q-1} \quad (33)$$

is invertible. Examples of other flexible choices include Exponential Almon Lag parametrization, or Beta Lag parametrization employed in the mixed-frequency literature, see, for example, Section 2.3.1 of Foroni and Marcellino (2013) for a description of these and other alternatives.

Remark 4 It can be seen from (31) that when the number of unknown parameters in $\boldsymbol{\psi}$ is the same as the dimension of \mathbf{b}_p , namely when $q = k$, then $\boldsymbol{\Sigma}_p = \boldsymbol{\Omega}_p$ and no improvement in the asymptotic variance is obtained. Consequently, the restricted MFDL estimator $\tilde{\mathbf{b}}_p$ can be asymptotically more efficient only if $q < k$.

Replacing $\boldsymbol{\Omega}_p$ with its consistent estimator $\hat{\boldsymbol{\Omega}}_p$ given by (25), and \mathbf{J}_p with its consistent estimator $\hat{\mathbf{J}}_p = \mathbf{J}_p(\hat{\boldsymbol{\psi}})$, the asymptotic variance matrix $\boldsymbol{\Sigma}_p$ given by (31) can be consistently estimated by

$$\hat{\boldsymbol{\Sigma}}_p = \hat{\mathbf{J}}_p \left(\hat{\mathbf{J}}'_p \hat{\boldsymbol{\Omega}}_p^{-1} \hat{\mathbf{J}}_p \right)^{-1} \hat{\mathbf{J}}'_p. \quad (34)$$

Remark 5 It is possible to improve the asymptotic variance of the unrestricted and the restricted

MFDL estimators by augmenting the auxiliary regressions (16) with variables that are correlated with the residual term ϑ_{ps} , but uncorrelated with the regressors in $\bar{\mathbf{e}}_{(p),s} = (\bar{\mathbf{e}}'_s, \bar{\mathbf{e}}'_{s-1}, \dots, \bar{\mathbf{e}}'_{s-p})'$. One candidate for the augmentation of (16) is the dependent variable lagged by (at least) $p + 1$ periods, which (under Assumptions 1-2) is uncorrelated with the regressors but correlated with the dependent variable, and could therefore improve the fit of the auxiliary regression. However, whether such augmentation is helpful or not in finite samples will depend on the magnitude of the improvement in the fit, which might be small when p is large and when, at the same time, the dependent variable is not very persistent. When the target variable is very persistent, or even integrated of order one, $I(1)$ for short, then the augmentation of the auxiliary regression by the dependent variable lagged by $p + 1$ periods will be necessary.²

3 Monte Carlo experiments

This section investigates the small sample performance of the unrestricted and restricted MFDL estimators, $\hat{\mathbf{b}}_{(p)}$ and $\tilde{\mathbf{b}}_{(p)}$, under two temporal aggregation weight schemes. Subsection 3.1 sets out the design of the experiments, and Subsection 3.2 summarizes the findings.

3.1 Data-generating process

We consider $n = k = 2$ and generate $\mathbf{z}_t = (x_t, q_t)'$ based on a VAR model (1) augmented with deterministic terms:

$$\mathbf{z}_t = (\mathbf{I}_2 - \mathbf{\Phi}) \boldsymbol{\mu}_z + \mathbf{\Phi} \mathbf{z}_{t-1} + \mathbf{A} \boldsymbol{\xi}_t, \text{ with } \boldsymbol{\xi}_t = (\xi_{1t}, \xi_{2t})' \sim IIDN(\mathbf{0}, \mathbf{I}_2), \quad (35)$$

for $t = -M + 1, -M + 2, \dots, 0, 1, \dots, T$, where $\boldsymbol{\mu}_z = (1, 1)'$, and the starting values $\mathbf{z}_{-M} = \mathbf{0}$. We set $M = 100$ and discard the first M time periods as burn-in periods. The matrix of the autoregressive coefficients, $\mathbf{\Phi}$, and the matrix of parameters \mathbf{A} are set equal to

$$\mathbf{\Phi} = \begin{pmatrix} 0.6 & 0.1 \\ 0.2 & 0.5 \end{pmatrix}, \text{ and } \mathbf{A} = \begin{pmatrix} 1 & -0.2 \\ 0.2 & 1 \end{pmatrix}. \quad (36)$$

²The theory developed in this paper covers the stationary case. The $I(1)$ case could be considered by adding the $(p + 1)$ -th lag of the dependent variable in the auxiliary regression (14). We leave the $I(1)$ case for future research.

We assume that ξ_{1t} is observed for $t = 1, 2, \dots, T$. In addition, the aggregate series $\bar{x}_{w,s} = \sum_{h=0}^{m-1} w_h x_{t_s-h}$ is observed for $s = 1, 2, \dots, T_m$. We set $m = 3$, which corresponds to a setting in which there are quarterly observations on the variable of interest and monthly data on the shock of interest. Furthermore, we consider two aggregation schemes, $\mathbf{w}_a = (1, 0, 0)'$ and $\mathbf{w}_b = (1, 1, 1)'$. The former represents the sequential sampling, or “end-of-period” measurement; the latter represents the sums of the m high-frequency periods, and, when dividing by m , “average-of-period” measurement.

We investigate the bias and root mean square error (RMSE) of the unrestricted and restricted MFDL estimators $\widehat{\mathbf{b}}_{(p)}$ and $\widetilde{\mathbf{b}}_{(p)}$ defined in (20), and (29). We set $p = 3$, and approximate the true high-frequency IRF coefficients in \mathbf{b}_p with $\psi^{-1}(L, q)$ with $q = 5$, where $\psi(L, q)$ is an invertible q -order polynomial given by (33). In addition, we investigate the size and power properties of the tests based on the two MFDL estimators. The variances of $\widehat{\mathbf{b}}_{(p)}$ and $\widetilde{\mathbf{b}}_{(p)}$ are estimated using (25) and (34), respectively. We consider $T_m = \{50, 100, 200, 500\}$ and conduct $R = 4,000$ Monte Carlo replications. In the context of the quarterly-monthly setting mentioned above, this choice of sample sizes corresponds to datasets that span 12.5 to 125 years of quarterly (monthly) observations on the outcome variable (shock) of interest. For later reference, the true IRFs are represented by the thick black lines in Figures A2-A3.

3.2 Monte Carlo findings

The results from the Monte Carlo experiments are presented in Table 1 for the temporal aggregation weights $\mathbf{w}_a = (1, 0, 0)'$ and in Table 2 for $\mathbf{w}_b = (1, 1, 1)'$.

The findings for the aggregation scheme \mathbf{w}_a in Table 1 show little or no bias for all values of T_m . The RMSE of the restricted and unrestricted MFDL estimators decreases with T_m . The RMSE of the two estimators is quite similar for shorter IRF horizons (0 to 3), whereas the RMSE of the restricted MFDL estimator is substantially smaller, about only half, compared with the unrestricted estimator for longer horizons. The size of the tests based on the two estimators is quite similar, with slight over-rejections in the range of 11 to 15 percent in case of $T_m = 50$. These size distortions disappear as T_m increases. The power findings qualitatively correspond to the RMSE findings and document significantly better power of the tests based on the restricted MFDL estimator for longer

IRF horizons ($h > 3$).

The small sample performance significantly deteriorates in case of the temporal aggregation scheme $\mathbf{w}_b = (1, 1, 1)'$, as can be expected against the background of discussion in Subsection 2.3.1. While the deterioration in the magnitude of the bias is overall not serious, the RMSE findings reported in Table 2 document a substantial deterioration, especially in the case of the unrestricted MFDL estimator at longer IRF horizons. The RMSE of $\widehat{\mathbf{b}}_{(p)}$ increases by a factor of two to four, and, in addition, we see that the RMSE of the unrestricted MFDL estimator now considerably increases with the IRF horizon, which is a consequence of increasing diagonal elements of $\mathbf{W}_p^{-1}\mathbf{W}_p'^{-1}$. In contrast, the restricted MFDL estimator $\widetilde{\mathbf{b}}_{(p)}$ performs much better than its unrestricted counterpart, as its RMSE tends to decline with higher IRF horizons, and is smaller for all choices of T_m and h than in case of the unrestricted MFDL estimator. The differences between the RMSE findings of the two estimators are very large for higher IRF horizons. For instance, for $h = 10$, the RMSE of the unrestricted MFDL estimator is 5 to 18 times as large as that of the restricted MFDL estimator. The tests based on the two estimators are somewhat oversized for $T_m = 50$, but the size distortions disappear with an increase in the sample size. Corresponding to the findings for the RMSE, the power of the tests based on the restricted MFDL estimator is in all cases higher, and in many cases substantially higher, than for the unrestricted MFDL estimator.³

To shed light on the concern on the possibility of a spuriously cyclical behavior of $\widehat{\mathbf{b}}_{(p)}$, in Figures A2-A3 we plot the first 50 estimates of $\mathbf{b}_{(p)}$ for the restricted and unrestricted MFDL estimators for the two cases of temporal aggregation schemes \mathbf{w}_a and \mathbf{w}_b . We see that the IRF estimates appear rather smooth under \mathbf{w}_a , for both estimators, but some draws of the unrestricted MFDL estimator exhibit a somewhat cyclical behavior. This is a small sample issue, since the RMSE declines with an increase in sample size and consequently any estimation error eventually becomes negligible. However, in finite samples, this is a drawback. The restricted MFDL estimator (lower panel of Figure A3) mitigates substantially this undesirable property of the unrestricted MFDL estimator.

Overall, the results from the Monte Carlo experiments document that the restricted MFDL estimator performs satisfactorily. Moreover, the restricted MFDL estimator seems to always perform better than the unrestricted MFDL estimator, which could suffer from serious small sample

³In the Appendix we also report results from Monte Carlo experiments in which the regression (16) is augmented with the dependent variable lagged by $p + 1$ periods in order to improve the fit.

performance issues, depending on the temporal aggregation weights. Consequently, we recommend to use the restricted MFDL estimator when the aggregation weights $\{w_\ell\}$ do not correspond to the sequential sampling scheme \mathbf{w}_a . We follow this recommendation in the empirical section below.

4 Daily crude oil price pass-through

In order to illustrate the usefulness of the estimators developed in this paper, we estimate the pass-through of changes in the Brent crude oil price observed at the daily frequency to U.S. retail prices for regular grade gasoline observed at the weekly frequency. Specifically, we obtain daily data on European Free Market Brent Crude Oil price (\$/Barrel) reported by the Financial Times, which we label as P_t . The daily series is available for working days only (Mondays to Fridays). We transform the daily series in logs and compute first differences, $e_t = \Delta \ln(P_t)$. We find that e_t does not exhibit any significant serial correlation, and consequently we treat e_t as the observed (white noise) reduced-form price shock. This is in line with approximating $\ln(P_t)$ by a daily random-walk. Since we cannot distinguish between the underlying reasons for the daily crude oil price changes (which could be due to a combination of supply, demand or speculative considerations), we refer to e_t as an oil price shock, without giving it a structural interpretation. This is in line with the price-pass-through literature, which does not distinguish the underlying sources of price movements. In turn, we obtain weekly data on U.S. Retail Gasoline Prices (Regular Grade, Monday prices, Cents/Gallon), which, after taking logs and first differences, we denote by $\bar{x}_{w,s}$. Our low-frequency period is therefore one week consisting of $m = 5$ working days. The observed retail price series $\bar{x}_{w,s}$ corresponds to the sum of daily log-differences, and therefore the aggregation weights vector \mathbf{w} is a vector of ones. Our sample spans the time period from 16 January 1991 to 22 May 2017, which gives $T_m = 1374$ weeks and $T = 6870$ working days.

Using the daily data on crude oil prices and the weekly data on U.S. retail prices for gasoline, we estimate the pass-through of a unit crude oil price shock at the *daily* frequency. We use the restricted MFDL estimator with $p = 4$ weeks, and set the lag of the polynomial that approximates the true high-frequency IRF coefficients in (33) to $q = 5$. Figure 1 reports estimates and 95% confidence intervals of the cumulative IRF coefficients for horizons of $h = 0, 1, \dots, 20$ days. The estimates suggest that about 12% of the crude oil price increase is passed through to retail gasoline

prices on the impact day ($h = 0$), about 28% of the crude oil price increase is passed through five working days after the shock ($h = 5$), and about 48% twenty working days after the shock ($h = 20$).

To shed some light on whether the pattern of the crude oil price pass-through has been stable throughout our sample period, we re-estimate the pass-through using rolling window with a length of $T_m = 500$ weeks. Figure 2 reports the rolling-window estimates of the cumulative pass-through for the horizons of $h = 0, 2, 5, 10$ and 20 days. The findings show that the pass-through has increased over time for each of the chosen horizons, with the exception of the last two years, when pass-through estimates at horizons of $h \geq 2$ days decreased slightly.

5 Conclusion

This paper considers the estimation of IRFs in settings in which the outcome variables of interest are observed at a lower frequency than the shock. We propose a restricted and an unrestricted MFDL estimator, derive their asymptotic distribution, and document by means of Monte Carlo experiments that only the restricted MFDL estimator has satisfactory small-sample performance across different temporal aggregation weights. Against the background of these findings, we employ the restricted MFDL estimator to estimate the daily pass-through of changes in crude oil prices in the U.S. observed at the daily frequency on U.S. retail gasoline prices observed at the weekly frequency. There are many other interesting empirical questions that can be addressed within the framework we propose in this paper. One example concerns the effects of exchange rate changes observed at high frequency on consumer prices observed at best at monthly frequency. Another example relates to the effects of monetary policy shocks constructed at high frequency on GDP observed at best at the quarterly frequency.

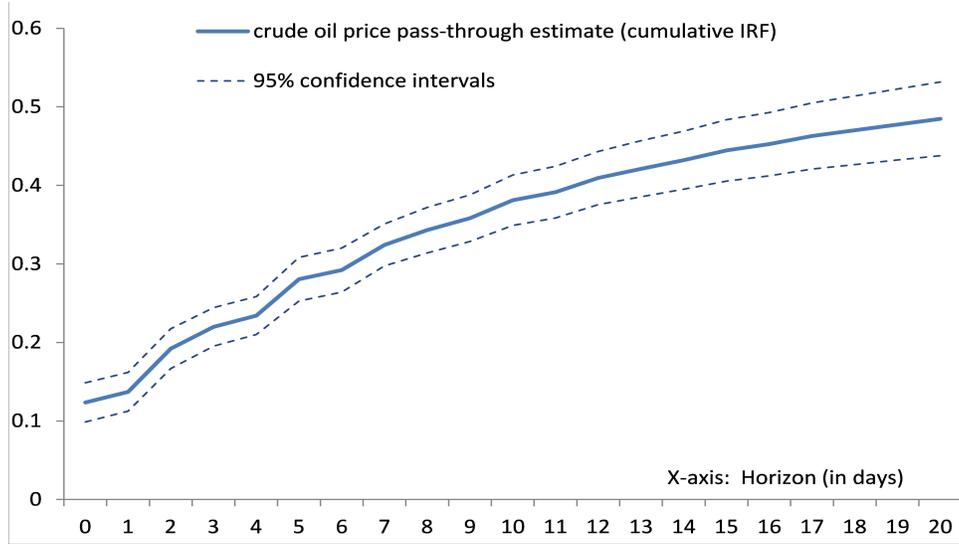


Figure 1: Full sample crude oil price pass-through estimates (cumulative IRF): Point estimates and 95% confidence intervals.

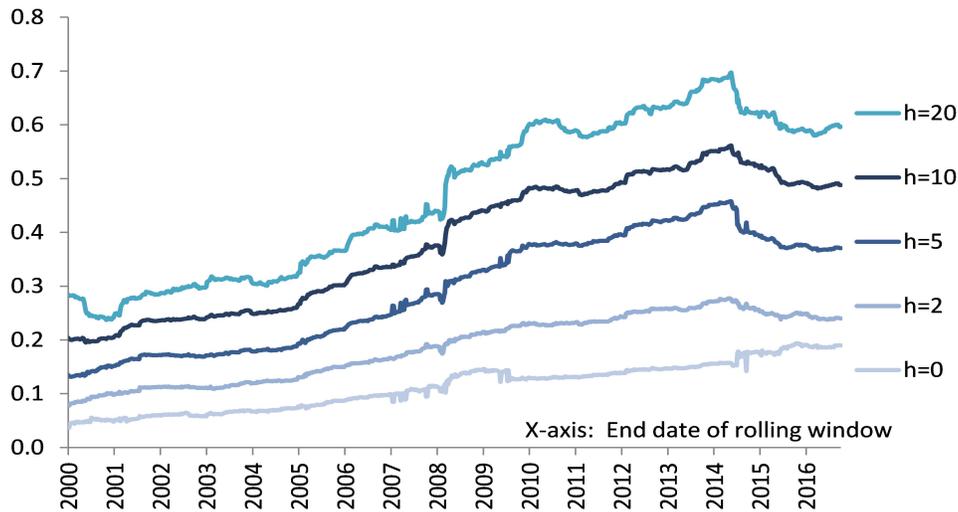


Figure 2: Rolling window estimates of the crude oil price pass-through at chosen horizons ($h = 0, 2, 5, 10,$ and 20 days).

Table 1: Monte Carlo findings using aggregation weights \mathbf{w}_a

Dependent variable is observed as temporally aggregated using the weights $\mathbf{w}_a = (1, 0, 0)'$.

$b_\ell :$	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Bias ($\times 100$)											
T_m	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	-0.13	-0.11	0.04	-0.05	-0.06	-0.10	-0.07	-0.05	-0.04	-0.03	-0.11
100	-0.02	-0.04	-0.01	-0.02	-0.03	-0.06	-0.07	0.02	-0.01	0.02	-0.09
200	-0.02	0.01	0.01	0.02	-0.01	0.02	-0.07	-0.01	-0.04	-0.01	-0.06
500	-0.01	0.00	0.00	0.02	-0.02	0.01	0.00	-0.01	-0.02	-0.01	-0.01
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	-0.17	-0.11	0.00	-0.01	-0.10	-0.16	0.12	-0.15	0.04	-0.11	-0.11
100	-0.03	-0.03	-0.02	-0.03	-0.06	0.01	0.08	-0.07	-0.03	-0.09	-0.10
200	-0.03	0.01	0.00	0.01	-0.04	0.07	0.07	-0.04	-0.04	-0.08	-0.09
500	-0.01	0.01	-0.01	0.01	-0.07	0.10	0.08	-0.01	-0.04	-0.07	-0.07
RMSE ($\times 100$)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	3.79	3.78	3.76	3.73	3.76	3.80	3.79	3.76	3.79	3.77	3.70
100	2.37	2.38	2.35	2.37	2.32	2.41	2.37	2.34	2.37	2.39	2.36
200	1.60	1.62	1.60	1.58	1.64	1.63	1.59	1.59	1.63	1.61	1.58
500	0.99	1.01	0.98	0.98	0.98	0.98	0.99	0.97	0.99	0.97	1.00
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	3.87	3.88	3.92	3.84	3.69	1.90	1.79	1.74	1.52	1.30	1.05
100	2.39	2.42	2.37	2.38	2.24	1.09	1.07	1.05	0.93	0.79	0.65
200	1.61	1.62	1.60	1.57	1.57	0.74	0.72	0.70	0.63	0.53	0.44
500	0.98	1.01	0.98	0.97	0.94	0.45	0.44	0.43	0.39	0.33	0.28
Size ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	11.5	11.8	11.5	11.3	12.3	12.1	12.1	11.5	11.9	11.5	11.1
100	7.6	7.8	7.1	7.4	7.4	8.1	7.3	6.9	7.9	7.6	7.4
200	6.3	6.6	6.3	6.2	7.2	6.3	6.0	6.1	6.5	6.3	6.0
500	5.0	6.1	5.2	5.6	5.2	5.3	5.3	5.2	5.3	5.3	5.7
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	13.3	14.0	15.3	15.0	14.6	13.5	12.8	12.5	12.0	12.2	12.0
100	8.4	8.7	8.3	8.5	8.3	7.8	7.6	7.9	7.3	7.5	7.4
200	6.8	6.9	6.5	6.5	7.7	6.6	6.7	6.5	6.3	6.6	6.7
500	5.1	6.3	5.5	5.6	5.6	6.3	5.6	5.3	5.3	5.6	6.0
Power ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	39.3	39.6	41.6	40.0	39.1	39.4	40.5	40.6	39.7	39.2	39.1
100	62.4	62.3	62.0	63.3	63.1	60.9	61.2	63.4	62.2	63.1	61.7
200	89.2	89.2	89.5	90.2	88.5	89.1	88.9	90.1	89.0	88.8	88.2
500	99.8	99.9	99.9	100.0	99.8	99.9	99.9	99.9	100.0	99.9	99.9
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	40.6	42.0	43.7	43.6	44.9	85.1	90.5	88.7	95.3	97.0	99.5
100	62.1	62.6	63.3	64.2	67.2	99.5	99.7	99.7	99.9	100.0	100.0
200	89.2	89.4	90.3	90.1	91.1	100.0	100.0	100.0	100.0	100.0	100.0
500	99.9	99.9	99.9	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0

Notes: $\mathbf{z}_t = (x_t, q_t)'$ is generated by VAR model (35), namely $\mathbf{z}_t = (\mathbf{I}_2 - \Phi) \mu_z + \Phi \mathbf{z}_{t-1} + \mathbf{A} \xi_t$, with $\xi_t = (\xi_{1t}, \xi_{2t})' \sim IIDN(\mathbf{0}, \mathbf{I}_2)$, $\mu_z = (1, 1)'$ and the coefficient matrices Φ and \mathbf{A} are given by (36). The aggregate series $\bar{x}_{w,s} = \sum_{h=0}^{m-1} w_h x_{t_s-h}$ is observed for $s = 1, 2, \dots, T_m$, and the shock ξ_{1t} is observed for $t = 1, 2, \dots, T$, where $T_m = T/m$, $t_s = s \cdot m$, and $m = 3$. Unrestricted MFDL estimator $\hat{b}_{\ell,p}$, given by (20), is computed using $p = 3$. The restricted MFDL estimator $\tilde{b}_{\ell,p}$, given by (29), is computed using $p = 3$ and IRF coefficients are approximated by the coefficients of $\psi^{-1}(L, q)$ with $q = 5$. The IRF coefficients are given by $b_\ell = \mathbf{s}'_{2,1} \Phi^\ell \mathbf{a}_1$, for $\ell = 0, 1, \dots$, where $\mathbf{s}_{2,1} = (1, 0)'$, and $\mathbf{a}_1 = \mathbf{A} \mathbf{s}_{2,1}$. The power rejection frequencies are reported using the alternative $H_1 : b_\ell = \mathbf{s}'_{2,1} \Phi^\ell \mathbf{a}_1 - 0.05$. See Section 3 for additional details.

Table 2: Monte Carlo findings using aggregation weights \mathbf{w}_b Dependent variable is observed as temporally aggregated using the weights $\mathbf{w}_b = (1, 1, 1)'$.

$b_\ell :$	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Bias ($\times 100$)											
T_m	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	-0.24	0.01	0.10	-0.34	0.20	0.08	-0.55	0.38	0.04	-0.57	0.29
100	-0.16	0.01	0.03	0.01	-0.03	0.00	-0.17	0.19	-0.06	-0.20	0.20
200	-0.02	0.03	-0.02	-0.01	-0.03	0.04	-0.13	0.03	0.08	-0.11	-0.02
500	-0.04	0.00	0.02	0.05	-0.05	0.01	0.03	-0.08	-0.01	0.01	-0.03
	Restricted MFDL estimator $b_{\ell,p}$										
50	-0.03	-0.10	-0.53	0.67	-0.05	-0.94	0.97	-0.70	0.14	0.00	0.05
100	-0.07	-0.03	-0.20	0.41	-0.21	-0.34	0.46	-0.31	0.08	-0.01	-0.05
200	0.07	-0.08	-0.09	0.26	-0.24	-0.10	0.26	-0.16	0.03	-0.03	-0.06
500	0.01	-0.07	0.01	0.18	-0.20	0.02	0.13	-0.06	-0.01	-0.05	-0.06
RMSE ($\times 100$)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	7.24	10.29	10.26	12.04	13.38	13.56	14.75	15.20	15.72	16.69	16.85
100	4.51	6.48	6.46	7.42	8.52	8.46	8.87	9.40	9.42	9.65	10.05
200	3.05	4.32	4.37	4.94	5.56	5.68	5.85	6.17	6.16	6.26	6.43
500	1.91	2.70	2.71	3.06	3.46	3.43	3.58	3.75	3.71	3.85	3.86
	Restricted MFDL estimator $b_{\ell,p}$										
50	6.86	9.47	8.98	9.27	6.81	5.07	4.36	3.99	3.39	3.26	2.93
100	4.14	5.48	5.33	5.26	3.45	1.96	1.47	1.23	1.05	0.93	0.71
200	2.79	3.60	3.51	3.57	2.20	0.92	0.71	0.62	0.54	0.46	0.36
500	1.72	2.21	2.11	2.09	1.33	0.47	0.36	0.34	0.32	0.28	0.23
Size ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	11.2	12.1	11.7	12.3	11.8	12.9	13.7	12.8	14.4	15.3	14.0
100	7.1	7.7	7.3	7.1	8.6	8.7	8.2	9.3	9.1	8.7	9.2
200	5.9	6.0	6.2	5.9	5.8	6.8	6.6	6.9	6.5	6.6	7.2
500	5.6	5.6	6.2	4.8	6.1	5.7	5.2	5.9	5.1	5.6	5.6
	Restricted MFDL estimator $b_{\ell,p}$										
50	15.1	18.0	20.3	19.1	14.4	6.2	6.5	7.8	7.0	7.4	7.8
100	8.1	9.1	10.5	9.6	7.1	2.2	2.7	3.9	4.8	5.7	5.6
200	6.7	7.3	8.3	7.7	5.5	1.5	2.6	3.8	4.9	5.4	5.2
500	5.9	5.4	6.5	5.5	5.2	2.4	3.6	4.6	5.6	6.1	6.5
Power ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	19.6	15.9	16.6	14.7	15.6	15.0	15.4	14.7	16.5	16.8	15.6
100	23.2	16.4	16.7	14.3	14.0	13.6	12.8	12.9	13.6	12.8	13.2
200	40.3	23.6	23.8	18.7	17.4	17.2	15.3	15.9	15.8	15.2	14.7
500	75.1	48.1	47.5	39.5	31.5	32.5	29.9	27.8	28.2	28.7	25.9
	Restricted MFDL estimator $b_{\ell,p}$										
50	25.0	24.2	29.1	20.5	35.7	54.3	74.4	64.3	87.7	86.2	92.5
100	28.2	23.1	28.1	20.6	44.0	82.2	98.3	93.5	98.5	99.5	99.4
200	47.8	33.6	39.0	34.4	62.9	96.3	100.0	99.8	100.0	100.0	100.0
500	85.1	63.2	68.0	72.2	93.3	100.0	100.0	100.0	100.0	100.0	100.0

Notes: See notes to Table 1.

References

- Bacchiocchi, E., A. Bastianin, A. Missale, and E. Rossi (2018). Structural analysis with mixed-frequency data: A midas-svar model of us capital flows. arXiv preprint, <https://arxiv.org/abs/1802.00793>.
- Choi, C.-Y. and A. Chudik (2019). Estimating impulse response functions when the shock series is observed. Federal Reserve Bank of Dallas, Globalization and Monetary Policy Working paper No. 353, <https://www.doi.org/10.24149/gwp353>.
- Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press.
- Forni, C., E. Ghysels, and M. Marcellino (2013). Mixed-frequency vector autoregressive models? In *VAR Models in Macroeconomics—New Developments and Applications: Essays in Honor of Christopher A. Sims*, pp. 247–272. Emerald Group Publishing Limited.
- Forni, C. and M. Marcellino (2013). A survey of econometric methods for mixed-frequency data. Norges Bank Research Working Paper No. 2013|06, also available at SSRN: <http://dx.doi.org/10.2139/ssrn.2268912>.
- Forni, C. and M. Marcellino (2016). Mixed frequency structural vector auto-regressive models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179(2), 403–425. <https://doi.org/10.1111/rssa.12120>.
- Jorda, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review* 95, 161–182. <https://doi.org/10.1257/0002828053828518>.
- Kilian, L. (2008a). A Comparison of the Effects of Exogenous Oil Supply Shocks on Output and Inflation in the G7 Countries. *Journal of the European Economic Association* 6(1), 78–121. <https://doi.org/10.1162/JEEA.2008.6.1.78>.
- Kilian, L. (2008b). Exogenous oil supply shocks: How big are they and how much do they matter for the u.s. economy? *Review of Economics and Statistics* 90, 216–240. <https://doi.org/10.1162/rest.90.2.216>.
- Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74, 119–147. <https://doi.org/10.1016/0304->

4076(95)01753-4.

Pesaran, M. H. and Y. Shin (1998). Generalised impulse response analysis in linear multivariate models. *Economics Letters* 58, 17–29. [https://doi.org/10.1016/s0165-1765\(97\)00214-0](https://doi.org/10.1016/s0165-1765(97)00214-0).

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review* 100, 763–801. <https://doi.org/10.3386/w13264>.

A Appendix

A.1 Notation

Matrices are denoted by bold upper case letters and vectors by bold lower case letters. All vectors are column vectors. $O(\cdot)$ and $o(\cdot)$ denote the Big O and Little o notations, respectively. If $\{f_n\}_{n=1}^{\infty}$ is any real sequence and $\{g_n\}_{n=1}^{\infty}$ is a sequences of positive real numbers, then $f_n = O(g_n)$ if there exists a positive finite constant K such that $|f_n|/g_n \leq K$ for all n . $f_n = o(g_n)$ if $f_n/g_n \rightarrow 0$ as $n \rightarrow \infty$. Convergence in probability and in distribution are denoted as \xrightarrow{p} and \xrightarrow{d} , respectively.

A.2 Proofs

Proof of Lemma 1. Consider x_{t_s} first. Using (6), we have for $t = t_s$,

$$x_{t_s} = \sum_{\ell=0}^{\infty} b_{\ell} e_{t_s-\ell} + \varepsilon_{t_s} = \sum_{q=0}^{m-1} \sum_{\ell=0}^{\infty} b_{\ell m+q} e_{t_s-\ell m-q} + \varepsilon_{t_s}.$$

Using $t_s = ms$, we obtain $t_s - \ell m - q = t_{(s-\ell)} - q$ and therefore $e_{t_s-\ell m-q} = \bar{e}_{s-\ell,q}$, and we can write x_{t_s} as

$$x_{t_s} = \sum_{q=0}^{m-1} \sum_{\ell=0}^{\infty} b_{\ell m+q} \bar{e}_{s-\ell,q} + \varepsilon_{t_s}. \quad (\text{A.1})$$

Consider x_{t_s-h} , for $h = 1, 2, \dots, m-1$, next. Using (6), we have for $t = t_s - h$,

$$x_{t_s-h} = \sum_{\ell=0}^{\infty} b_{\ell} e_{t_s-h-\ell} + \varepsilon_{t_s-h}. \quad (\text{A.2})$$

Consider the first component on the right side of (A.2),

$$\begin{aligned} \sum_{\ell=0}^{\infty} b_{\ell} e_{t_s-\ell-h} &= \sum_{\ell=0}^{m-h-1} b_{\ell} e_{t_s-h-\ell} + \sum_{\ell=m-h}^{\infty} b_{\ell} e_{t_s-\ell-h} \\ &= \sum_{q=0}^{m-h-1} b_q \bar{e}_{s,h+q} + \sum_{q=0}^{m-1} \sum_{\ell=1}^{\infty} b_{m\ell+q-h} \bar{e}_{s-\ell,q}, \end{aligned}$$

and therefore

$$x_{t_s-h} = \sum_{q=0}^{m-h-1} b_q \bar{e}_{s,h+q} + \sum_{q=0}^{m-1} \sum_{\ell=1}^{\infty} b_{m\ell+q-h} \bar{e}_{s-\ell,q} + \varepsilon_{t_s-h}. \quad (\text{A.3})$$

The representation for $\bar{x}_{w,s}$ can now be readily obtained by substituting (A.1) and (A.3) in (4),

$$\bar{x}_{w,s} = w_0 \sum_{q=0}^{m-1} \sum_{\ell=0}^{\infty} b_{\ell m+q} \bar{e}_{s-\ell,q} + \sum_{h=1}^{m-1} w_h \left(\sum_{q=0}^{m-h-1} b_q \bar{e}_{s,h+q} + \sum_{q=0}^{m-1} \sum_{\ell=1}^{\infty} b_{m\ell+q-h} \bar{e}_{s-\ell,q} \right) + \bar{e}_{w,s}, \quad (\text{A.4})$$

where $\bar{e}_{w,s} = \sum_{h=0}^{m-1} w_h \varepsilon_{t_s-h}$. Collecting the individual terms, we obtain the representation (13). ■

Proof of Theorem 1. Let $T_{mp} = T/m - p$. Using (20) and $\mathbf{b}_{(p)} = \mathbf{W}_p^{-1} \mathbf{d}_{(p)}$, we have

$$\sqrt{T_{mp}} \left(\hat{\mathbf{b}}_{(p)} - \mathbf{b}_{(p)} \right) = \sqrt{T_{mp}} \mathbf{W}_p^{-1} \left(\hat{\mathbf{d}}_{(p)} - \mathbf{d}_{(p)} \right),$$

and substituting (21) for $\hat{\mathbf{d}}_{(p)}$, we obtain

$$\sqrt{T_{mp}} \left(\hat{\mathbf{b}}_{(p)} - \mathbf{b}_{(p)} \right) = \mathbf{W}_p^{-1} \left(\frac{\mathbf{E}'\mathbf{E}}{T_{mp}} \right)^{-1} \frac{\mathbf{E}'\boldsymbol{\vartheta}_p}{\sqrt{T_{mp}}}, \quad (\text{A.5})$$

where \mathbf{E} is $T_{mp} \times (p+1)m$ dimensional matrix of observations on the $(p+1)m$ regressors in $\bar{\mathbf{e}}_{(p),s} = (\bar{\mathbf{e}}'_s, \bar{\mathbf{e}}'_{s-1}, \dots, \bar{\mathbf{e}}'_{s-p})'$, and $\boldsymbol{\vartheta}_p = (\vartheta_{p,p+1}, \vartheta_{p,p+2}, \dots, \vartheta_{p,T_m})'$ with ϑ_{ps} , for $s = p+1, p+2, \dots, T_m$, given by (15). We focus on the individual terms on the right side of (A.5). Note that under Assumption 2, $e_t \sim IID(0, \sigma_e^2)$, for some $c < \sigma_e^2 < K$, and the fourth order moments of e_t are finite. Hence (for a fixed m and p), $\mathbf{E}'\mathbf{E}/T_{mp} \rightarrow_p \sigma_e \mathbf{I}_{(p+1)m}$, and therefore also

$$\left(\frac{\mathbf{E}'\mathbf{E}}{T_{mp}} \right)^{-1} \rightarrow_p \sigma_e^{-2} \mathbf{I}_{(p+1)m}, \quad (\text{A.6})$$

as $T \rightarrow \infty$. We focus on the term $\mathbf{E}'\boldsymbol{\vartheta}_p/\sqrt{T_{mp}}$ next. Let us define $\mathbf{y}_s = \bar{\mathbf{e}}_{(p),s} \vartheta_{ps}$, where we suppressed the subscript p to simplify the notations, but it is understood that the vector \mathbf{y}_s depends on p . Then we can write $\mathbf{E}'\boldsymbol{\vartheta}_p/\sqrt{T_{mp}}$ as

$$\frac{\mathbf{E}'\boldsymbol{\vartheta}_p}{\sqrt{T_{mp}}} = \frac{1}{\sqrt{T_{mp}}} \sum_{s=p+1}^{T_m} \mathbf{y}_s$$

Note that $E(\mathbf{y}_s) = \mathbf{0}$, since $\bar{\mathbf{e}}_{(p),s} = (\bar{\mathbf{e}}'_s, \bar{\mathbf{e}}'_{s-1}, \dots, \bar{\mathbf{e}}'_{s-p})'$ is independently distributed of ϑ_{ps} , and both have zero mean. Also, ϑ_{ps} is serially correlated, but $\bar{\mathbf{e}}_s$ is independently distributed over t .

Hence, \mathbf{y}_s is independently distributed from $\mathbf{y}_{s'}$ for $|s - s'| > p$. Let

$$\Psi \equiv \sum_{\ell=-p}^p E(\mathbf{y}_s \mathbf{y}'_{s+\ell}) = \sum_{\ell=-p}^p \gamma_{p\ell} \mathbf{H}_{p,\ell},$$

where $\gamma_{p\ell}$ and $\mathbf{H}_{p,\ell}$ are defined in the statement of this theorem, and we have again suppressed the subscript p . Let also

$$\Psi_{T_{mp}} = E \left(\frac{1}{T_{mp}} \sum_{s=p+1}^{T_m} \sum_{s'=p+1}^{T_m} \mathbf{y}_s \mathbf{y}'_{s'} \right),$$

and note that (by independence of \mathbf{y}_s and $\mathbf{y}_{s'}$ for $|s - s'| > p$)

$$\Psi_{T_{mp}} = \Psi + O(T_{mp}^{-1}). \quad (\text{A.7})$$

In addition, define the stochastic vector sequence $V_s = (e_{t_s}, e_{t_s-1}, \dots, e_{t_s-m+1}, \mathbf{v}'_{t_s}, \mathbf{v}'_{t_s-1}, \dots, \mathbf{v}'_{t_s-m+1})'$, and for any $(p+1)m$ dimensional vector \mathbf{a} satisfying $\|\mathbf{a}\| = 1$ define the stochastic array

$$X_{T_{mp},s} = \mathbf{a}' \Psi_{T_{mp}}^{-1/2} \mathbf{y}_s / c_{T_{mp}},$$

for $s = p+1, p+2, \dots, T_{mp}$, where $c_{T_{mp}}$ is a sequence of constants set equal to $c_{T_{mp}} = 1/\sqrt{T_{mp}}$. Under Assumptions 1-2, we have (i) \mathbf{y}_s and therefore also $X_{T_{mp},s}$ is L_2 -NED of any size on $\{V_s\}$, (ii) V_s is *IID* over s with bounded fourth order moments, (iii) $\sup_{T_{mp},t} \|X_{T_{mp},s}/c_{T_{mp}}\|_r < \infty$ for some $r > 2$, and (iv) $\sup_{T_{mp}} T_{mp} c_{T_{mp}}^2 = 1 < \infty$. Hence conditions of Corollary 24.7 of Davidson (1994) are satisfied and it follows that

$$\sum_{s=p+1}^{T_{mp}} X_{T_{mp},s} \rightarrow_d N(0, 1).$$

Since this result holds for any $\|\mathbf{a}\| = 1$, we obtain (using Cramer-Wold theorem)

$$\frac{1}{\sqrt{T_{mp}}} \sum_{s=p+1}^{T_m} \Psi_{T_{mp}}^{-1/2} \mathbf{y}_s \rightarrow_d N(0, \mathbf{I}_{(p+1)m}),$$

and therefore, noting (A.7),

$$\frac{\mathbf{E}'\boldsymbol{\vartheta}_p}{\sqrt{T_{mp}}} \rightarrow_d N(0, \boldsymbol{\Psi}). \quad (\text{A.8})$$

Using now (A.6) and (A.8) in (A.5), we obtain (22). This completes the proof. ■

Proof of Theorem 2. Consider the mean-value expansion of $\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\hat{\boldsymbol{\psi}})$ around $\boldsymbol{\psi}_0$,

$$\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\hat{\boldsymbol{\psi}}) = \hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\boldsymbol{\psi}_0) - [\mathbf{J}_p(\bar{\boldsymbol{\psi}})](\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0),$$

where $\bar{\boldsymbol{\psi}}$ lies, element by element, between $\hat{\boldsymbol{\psi}}$ and $\boldsymbol{\psi}_0$, and

$$\mathbf{J}_p(\bar{\boldsymbol{\psi}}) = \left. \frac{\partial \mathbf{f}_p(\boldsymbol{\psi})}{\partial \boldsymbol{\psi}'} \right|_{\boldsymbol{\psi}=\bar{\boldsymbol{\psi}}}.$$

Pre-multiplying from the left by $\mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}$ gives

$$\mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\hat{\boldsymbol{\psi}})] = \mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\boldsymbol{\psi}_0)] - \mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\mathbf{J}_p(\bar{\boldsymbol{\psi}})](\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0).$$

But $\mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\hat{\boldsymbol{\psi}})] = \mathbf{0}$ are the first order conditions of the minimization problem (28), and therefore

$$\mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\mathbf{J}_p(\bar{\boldsymbol{\psi}})](\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0) = \mathbf{J}'_p(\hat{\boldsymbol{\psi}})\boldsymbol{\Omega}_p^{-1}[\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\boldsymbol{\psi}_0)].$$

Under Assumption 3, we have $\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\boldsymbol{\psi}_0) = \hat{\mathbf{b}}_{(p)} - \mathbf{b}_{(p)}$, and using result (22) of Theorem 1, we have $\sqrt{T/m-p}[\hat{\mathbf{b}}_{(p)} - \mathbf{f}_p(\boldsymbol{\psi}_0)] \rightarrow_d N(\mathbf{0}, \boldsymbol{\Omega}_p)$. Since also $\mathbf{J}_p(\bar{\boldsymbol{\psi}})$ is continuous by Assumption 3, we obtain

$$\sqrt{T/m-p}(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0) \rightarrow_d N\left[\mathbf{0}, (\mathbf{J}'_p\boldsymbol{\Omega}_p^{-1}\mathbf{J}_p)^{-1}\right].$$

Noting that $\tilde{\mathbf{b}}_p = \mathbf{f}_p(\hat{\boldsymbol{\psi}})$ and $\mathbf{b}_p = \mathbf{f}_p(\boldsymbol{\psi}_0)$, where $\mathbf{f}_p(\cdot)$ satisfies Assumption 3, it now follows that $\sqrt{T/m-p}(\tilde{\mathbf{b}}_p - \mathbf{b}_p) \rightarrow_d N\left[\mathbf{0}, \mathbf{J}_p(\mathbf{J}'_p\boldsymbol{\Omega}_p^{-1}\mathbf{J}_p)^{-1}\mathbf{J}'_p\right]$. This completes the proof. ■

A.3 Extension of aggregation schemes

This part of Appendix explores the extension to the case where the aggregate variable is defined as

$$\bar{x}_{w_q,s} = \sum_{h=0}^{qm-1} w_{qh} x_{t_s-h}, \text{ for } s = 1, 2, \dots, T_m,$$

for some finite $q > 1$. The case of $q = 2$ is of particular interest, since it represents the case of first-differences of the temporally aggregated data, as illustrated in the next example.

Example 2 Consider the case with $m = 3$, which corresponds to quarterly-monthly case and suppose that y_t represents logs of GDP in a month $t = 1, 2, \dots, T$. Then, quarterly GDP series is given by

$$\bar{y}_{ws} = \sum_{h=0}^{m-1} w_h y_{t_s-h} = \sum_{h=0}^2 y_{t_s-h}, \text{ for } s = 1, 2, \dots, T_m,$$

where $w_h = 1$ for $h = 0, 1, 2$ and $T_m = [T/3]$. Suppose that a researcher observes the first differences of \bar{y}_{wt} , which can be represented as a temporal aggregate of the monthly growth series as

$$\begin{aligned} \bar{x}_{w_q,s} &= \bar{y}_{ws} - \bar{y}_{w,s-1} \\ &= \left(\sum_{h=0}^{m-1} w_h y_{t_s-h} \right) - \left(\sum_{h=0}^{m-1} w_h y_{t_s-h-1} \right) \\ &= \sum_{h=0}^{qm-1} w_{q,h} \Delta y_{t_s} = \sum_{h=0}^{6-1} w_{2,h} \Delta y_{t_s}, \end{aligned}$$

where $\Delta y_{t_s} = y_{t_s} - y_{t_s-1}$, $q = 2$, $m = 3$, and $(w_{2,0}, w_{2,1}, \dots, w_{2,5})' = (1, 2, 3, 2, 1, 0)'$.

The earlier analysis can be readily extended to cases with $q > 1$. Consider for the simplicity of exposition $p = hq$ for some $h = 0, 1, 2, \dots$, and define

$$\mathbf{W}_{b,q} \equiv \begin{pmatrix} w_0 & 0 & 0 & \cdots & 0 \\ w_1 & w_0 & 0 & \cdots & 0 \\ w_2 & w_1 & w_0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ w_{qm-1} & w_{qm-2} & w_{qm-3} & \cdots & w_0 \end{pmatrix}, \text{ and } \mathbf{W}_{a,q} \equiv \begin{pmatrix} 0 & w_{qm-1} & w_{qm-2} & w_{qm-3} & \cdots & w_1 \\ 0 & 0 & w_{qm-1} & w_{qm-2} & \cdots & w_2 \\ 0 & 0 & 0 & w_{qm-1} & & w_3 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & & w_{qm-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Then (18) holds with \mathbf{W}_p defined as

$$\mathbf{W}_{hq} = \begin{pmatrix} \mathbf{W}_b & \mathbf{0}_{qm \times qm} & \mathbf{0}_{qm \times qm} & \cdots & \mathbf{0}_{qm \times qm} \\ \mathbf{W}_a & \mathbf{W}_b & \mathbf{0}_{qm \times qm} & \cdots & \mathbf{0}_{qm \times qm} \\ \mathbf{0}_{qm \times qm} & \mathbf{W}_a & \mathbf{W}_b & & \mathbf{0}_{qm \times qm} \\ \vdots & & \ddots & \ddots & \\ \mathbf{0}_{qm \times qm} & \mathbf{0}_{qm \times qm} & \cdots & \mathbf{W}_a & \mathbf{W}_b \end{pmatrix}, \quad (\text{A.9})$$

and the previous theorems continue to apply. In practice it is not required that p is a multiple of q , and the expression for \mathbf{W}_p can be easily obtained by removing the appropriate number of rows of \mathbf{W}_p from the bottom. Specifically, Let $h = [(p + 1) / q] + 1$ and $n_r = (hq - p)qm$, where $[(p + 1) / q]$ is denote the integer part of $(p + 1) / q$. Then n_r rows need to be removed from the matrix \mathbf{W}_{hq} defined above.

A.4 Augmenting the auxiliary regression with the dependent variable lagged by $p + 1$ periods

Tables 3 and 4 report results from Monte Carlo experiments in which the regression in (16) is augmented with the dependent variable lagged by $p + 1$ periods in order to improve the fit.

There are no clear systematic differences across the results for the cases with and without augmentation with the lagged dependent variable under the aggregation weights \mathbf{w}_b . The results for the aggregation weights \mathbf{w}_a suggest that in smaller samples both the unrestricted and the restricted MFDL estimators exhibit larger RMSE, although the finite-sample bias is reduced. For the smaller values of T , the size performance has deteriorated for both estimators, while the power has improved. For the larger sample sizes, findings are very similar to the case in which the regression is not augmented with the lagged dependent variable.

Table 3: Monte Carlo findings in experiments with augmentation by lagged dependent variable and aggregation weights \mathbf{w}_a

Dependent variable is observed as temporally aggregated using the weights $\mathbf{w}_a = (1, 0, 0)'$ and the auxiliary regression (16) is augmented by $(p + 1)$ -th lag of the dependent variable.

$b_\ell :$	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Bias ($\times 100$)											
T_m	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	-0.08	-0.10	0.06	0.00	-0.05	-0.09	0.01	-0.01	-0.03	-0.01	-0.09
100	0.01	-0.02	0.00	0.00	-0.02	-0.05	-0.03	0.04	0.00	0.04	-0.08
200	-0.01	0.01	0.01	0.03	-0.01	0.02	-0.06	0.00	-0.03	0.00	-0.05
500	-0.01	0.00	0.00	0.02	-0.02	0.01	0.01	-0.01	-0.01	-0.01	-0.01
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	-0.12	-0.09	0.01	0.07	-0.10	-0.14	0.17	-0.12	0.07	-0.08	-0.09
100	0.00	-0.02	-0.02	-0.01	-0.05	0.03	0.10	-0.05	-0.01	-0.08	-0.09
200	-0.02	0.02	0.00	0.02	-0.04	0.07	0.07	-0.03	-0.04	-0.08	-0.09
500	-0.01	0.01	-0.01	0.01	-0.06	0.10	0.08	0.00	-0.03	-0.07	-0.07
RMSE ($\times 100$)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	3.90	3.90	3.86	3.82	3.86	3.89	3.87	3.88	3.94	3.85	3.81
100	2.40	2.40	2.37	2.39	2.34	2.43	2.39	2.35	2.39	2.41	2.39
200	1.60	1.62	1.61	1.58	1.65	1.64	1.59	1.59	1.64	1.61	1.57
500	0.99	1.01	0.98	0.98	0.98	0.98	0.99	0.97	0.98	0.98	1.00
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	3.98	4.02	4.05	3.97	3.77	1.95	1.86	1.79	1.57	1.33	1.07
100	2.41	2.44	2.40	2.41	2.26	1.10	1.08	1.05	0.93	0.79	0.65
200	1.61	1.63	1.61	1.58	1.58	0.74	0.72	0.70	0.63	0.53	0.44
500	0.98	1.01	0.98	0.97	0.93	0.45	0.44	0.43	0.38	0.33	0.28
Size ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	13.1	13.4	13.0	13.5	13.9	14.0	13.5	13.4	14.5	13.4	13.0
100	8.1	8.3	7.7	7.8	8.2	8.5	7.8	7.1	8.4	8.3	7.9
200	6.6	7.1	6.3	6.4	7.1	6.6	6.5	6.6	6.8	6.2	5.9
500	5.0	6.2	5.2	5.6	5.2	5.2	5.3	4.9	5.4	5.4	5.5
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	15.1	15.7	17.4	17.2	16.0	15.2	14.8	14.4	13.9	14.0	13.9
100	8.8	9.3	8.9	8.7	8.7	8.5	8.0	8.4	7.8	7.8	7.9
200	7.2	6.9	6.7	6.7	8.1	6.9	6.6	6.5	6.6	6.9	7.0
500	5.2	6.1	5.4	5.5	5.7	6.2	5.8	5.4	5.1	5.4	6.0
Power ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	41.1	42.2	42.5	41.7	40.5	40.7	42.4	41.5	41.7	41.3	40.8
100	62.5	62.8	62.4	63.4	63.6	61.9	62.2	64.0	62.7	63.8	62.2
200	89.5	89.1	89.6	90.1	88.9	89.1	89.3	90.0	88.8	89.0	88.6
500	99.9	99.8	99.9	100.0	99.9	99.9	99.9	99.9	100.0	99.9	99.9
	Restricted MFDL estimator $\tilde{b}_{\ell,p}$										
50	42.7	43.5	45.7	45.7	45.9	85.3	90.6	89.0	95.8	97.0	99.4
100	63.0	62.8	64.0	64.4	67.2	99.7	99.7	99.7	100.0	100.0	100.0
200	89.3	89.6	90.4	90.2	91.0	100.0	100.0	100.0	100.0	100.0	100.0
500	99.9	99.9	99.9	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0

Notes: See notes to Table 1.

Table 4: Monte Carlo findings in experiments with augmentation by lagged dependent variable and aggregation weights \mathbf{w}_b

Dependent variable is observed as temporally aggregated using the weights $\mathbf{w}_b = (1, 1, 1)'$ and the auxiliary regression (16) is augmented by $(p + 1)$ -th lag of the dependent variable.

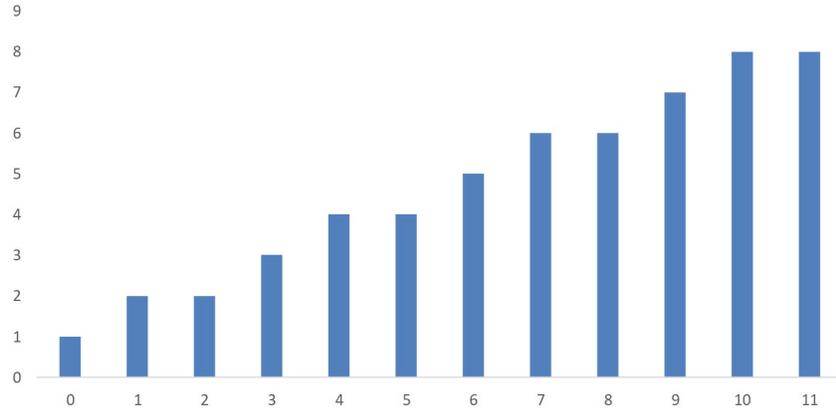
$b_\ell :$	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}
Bias ($\times 100$)											
T_m	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	-0.11	0.04	0.11	-0.26	0.21	0.15	-0.49	0.41	0.12	-0.61	0.36
100	-0.08	0.01	0.04	0.05	-0.04	-0.01	-0.08	0.18	-0.07	-0.15	0.20
200	0.02	0.03	-0.03	0.04	-0.03	0.01	-0.07	0.03	0.03	-0.05	-0.01
500	-0.03	0.01	0.01	0.06	-0.04	0.01	0.03	-0.07	-0.01	0.00	-0.01
	Restricted MFDL estimator $\check{b}_{\ell,p}$										
50	0.15	-0.09	-0.64	0.94	-0.09	-1.02	1.19	-0.75	0.17	0.12	0.02
100	0.00	-0.03	-0.19	0.44	-0.20	-0.34	0.49	-0.28	0.10	0.01	-0.03
200	0.10	-0.07	-0.08	0.29	-0.24	-0.10	0.27	-0.15	0.04	-0.03	-0.06
500	0.02	-0.07	0.01	0.19	-0.20	0.02	0.14	-0.05	0.00	-0.05	-0.06
RMSE ($\times 100$)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	7.38	10.52	10.55	12.52	13.87	14.01	15.27	15.96	16.42	17.25	17.76
100	4.52	6.50	6.46	7.39	8.54	8.41	8.86	9.42	9.36	9.71	10.04
200	3.04	4.32	4.36	4.95	5.51	5.62	5.82	6.10	6.07	6.19	6.35
500	1.89	2.69	2.70	3.03	3.42	3.39	3.52	3.68	3.66	3.77	3.79
	Restricted MFDL estimator $\check{b}_{\ell,p}$										
50	7.09	9.71	9.36	9.92	7.28	5.58	5.07	4.52	3.89	3.91	3.52
100	4.17	5.55	5.40	5.31	3.52	2.07	1.52	1.30	1.11	0.98	0.80
200	2.78	3.59	3.52	3.57	2.20	0.92	0.72	0.62	0.53	0.46	0.36
500	1.71	2.20	2.11	2.09	1.32	0.46	0.36	0.34	0.31	0.27	0.23
Size ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	12.7	13.8	13.6	15.0	15.2	14.7	15.6	16.6	17.1	17.8	17.4
100	7.2	8.7	8.0	7.7	9.3	9.0	8.8	9.6	9.5	9.8	10.4
200	6.1	6.4	6.3	6.6	6.0	7.0	6.9	7.1	6.2	7.0	7.5
500	5.6	5.7	6.2	5.1	6.0	5.6	5.2	5.8	5.5	5.8	5.6
	Restricted MFDL estimator $\check{b}_{\ell,p}$										
50	16.9	19.7	22.7	21.7	16.0	7.6	8.4	8.5	8.2	8.3	8.8
100	8.6	9.3	11.3	10.0	7.5	2.5	2.8	4.2	5.1	5.9	6.1
200	6.6	7.3	8.5	8.1	5.4	1.6	2.8	4.0	5.0	5.7	5.5
500	6.0	5.4	6.6	5.5	5.2	2.6	3.6	4.5	5.7	6.2	6.3
Power ($\times 100$, 5% level)											
	Unrestricted MFDL estimator $\hat{b}_{\ell,p}$										
50	21.3	17.5	17.8	17.1	17.6	17.9	17.9	18.2	19.0	19.7	19.6
100	24.6	16.5	17.4	15.2	14.9	14.5	13.2	13.8	13.5	14.1	14.7
200	41.6	24.7	23.8	19.7	17.8	17.6	16.8	16.5	16.7	16.6	15.8
500	75.9	48.9	48.0	40.0	32.9	33.5	31.1	28.7	29.5	29.3	27.7
	Restricted MFDL estimator $\check{b}_{\ell,p}$										
50	27.6	26.5	31.7	22.9	38.5	55.7	75.6	65.1	88.0	86.1	92.1
100	29.5	23.1	29.2	22.2	44.8	82.7	98.2	93.6	98.5	99.5	99.3
200	49.9	34.0	39.6	35.4	63.3	96.3	100.0	99.8	100.0	100.0	100.0
500	85.6	64.2	68.9	72.7	93.4	100.0	100.0	100.0	100.0	100.0	100.0

Notes: See notes to Table 1.

A.5 Additional Figures

Figure A1: Selected features of matrix \mathbf{W}_p for $m = 3$ with aggregation weights $\mathbf{w}_b = (1, 1, 1)'$.

A. Diagonal elements of $(\mathbf{W}_p^{-1}\mathbf{W}_p'^{-1})$.



B. Elements of $\mathbf{W}_p\mathbf{s}_1$ (solid line) and $\mathbf{W}_p\boldsymbol{\tau}$ (dashed line), where $\mathbf{s}_1 = (1, 0, \dots, 0)'$ and $\boldsymbol{\tau} = (1, 1, \dots, 1)'$.

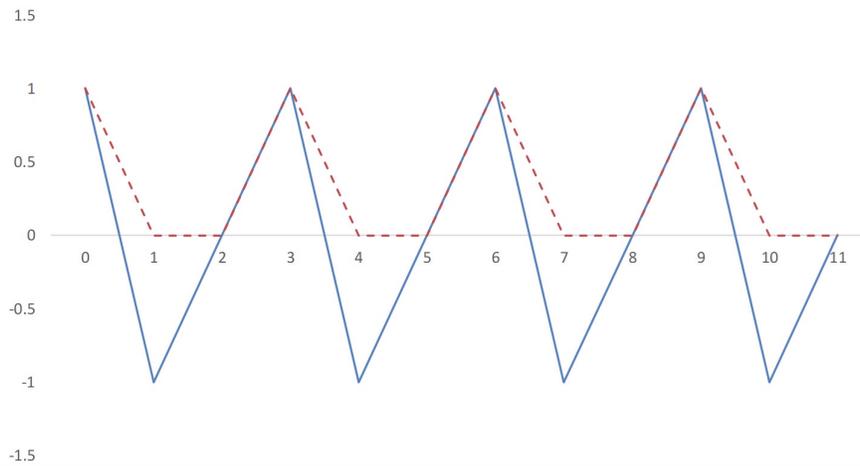
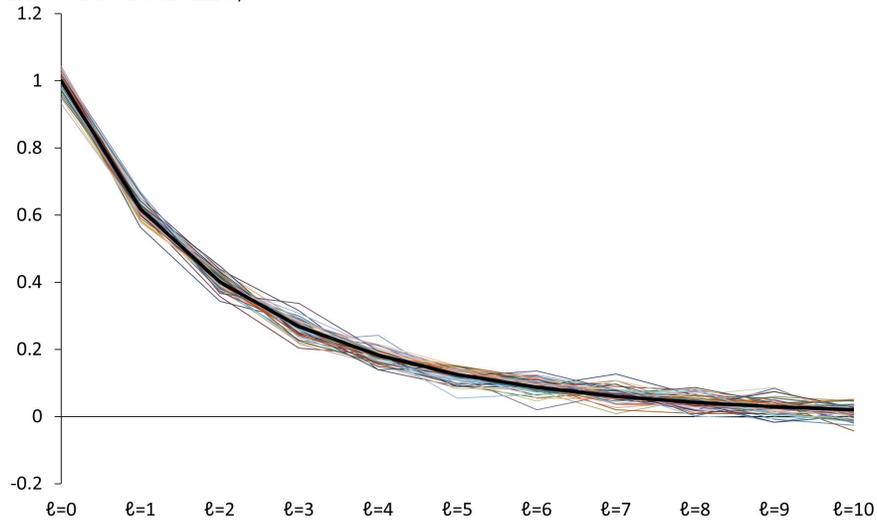
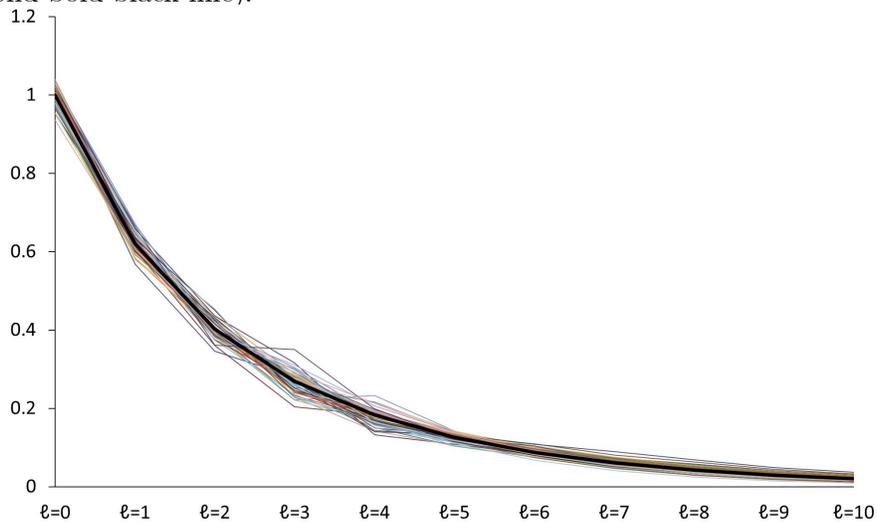


Figure A2: First 50 unrestricted (top chart) and restricted (bottom chart) MFDL estimates of IRF function in MC experiments with aggregation weights $\mathbf{w}_a = (1, 0, 0)'$.

A. The first $r = 1, 2, \dots, 50$ unrestricted IRF estimates $\{\hat{b}_{\ell,p}^{(r)}, \ell = 0, 1, \dots, 10\}$, and the true IRF coefficients (solid bold black line).



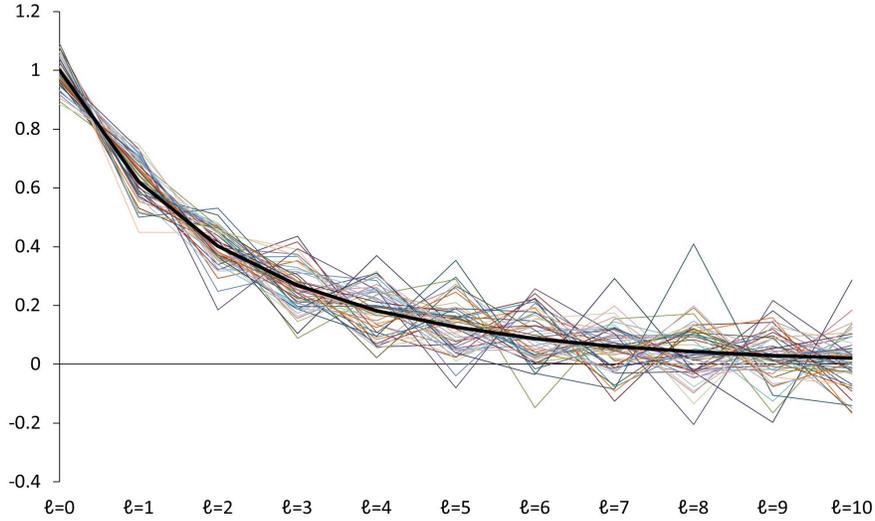
B. The first $r = 1, 2, \dots, 50$ restricted IRF estimates $\{\tilde{b}_{\ell,p}^{(r)}, \ell = 0, 1, \dots, 10\}$, and the true IRF coefficients (solid bold black line).



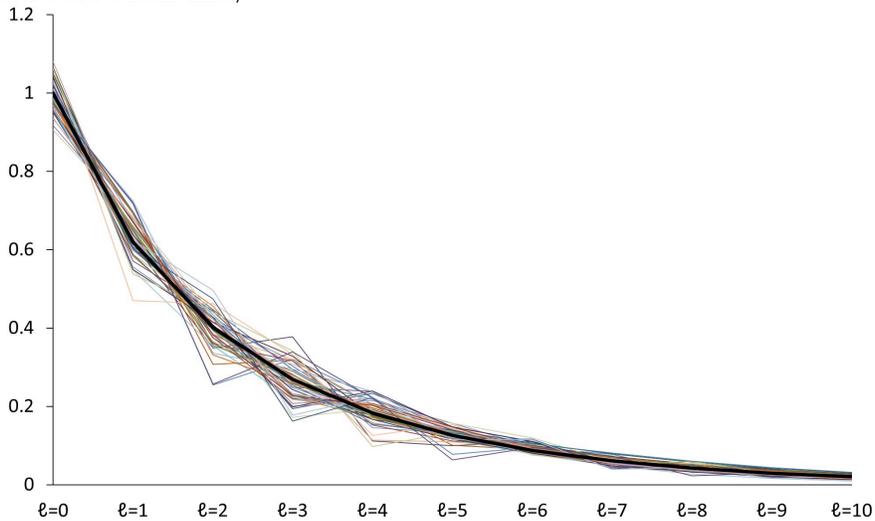
Notes: Horizon $\ell = 0, 1, \dots, 10$ (at the high frequency of the shock) is on the horizontal axis.

Figure A3: First 50 unrestricted (top chart) and restricted (bottom chart) MFDL estimates of IRF function in MC experiments with aggregation weights $\mathbf{w}_b = (1, 1, 1)'$.

A. The first $r = 1, 2, \dots, 50$ unrestricted IRF estimates $\{\hat{b}_{\ell,p}^{(r)}, \ell = 0, 1, \dots, 10\}$, and the true IRF coefficients (solid bold black line).



B. The first $r = 1, 2, \dots, 50$ restricted IRF estimates $\{\tilde{b}_{\ell,p}^{(r)}, \ell = 0, 1, \dots, 10\}$, and the true IRF coefficients (solid bold black line).



Notes: Horizon $\ell = 0, 1, \dots, 10$ (at the high frequency of the shock) is on the horizontal axis.