Federal Reserve Bank of Dallas

Mean Group and Pooled Mixed-Frequency Estimators of Responses of Low-Frequency Variables to High-Frequency Shocks

Alexander Chudik and Lutz Kilian

Working Paper 2603 **February 2026**

Research Department

https://doi.org/10.24149/wp2603

Mean Group and Pooled Mixed-Frequency Estimators of Responses of Low-Frequency Variables to High-Frequency Shocks*

Alexander Chudik† and Lutz Kilian‡

February 10, 2026

Abstract

This paper proposes mean group and pooled estimators of impulse responses based on mixed-frequency auxiliary distributed lag (DL), autoregressive distributed lag (ARDL), or vector autoregressive distributed lag (VARDL) estimating equations. Our setup assumes that the data are generated by a high-frequency VAR process. While the shock of interest is directly observed at high frequency, the outcome variable is only observed as a temporally aggregated variable at a lower frequency. We derive the asymptotic distributions of the six proposed estimators. Monte Carlo experiments show that pooled estimators generally perform better than the corresponding mean group estimators for relevant sample sizes. An empirical illustration to the pass-through from daily wholesale gasoline price shocks to monthly consumer price inflation illustrates the usefulness of the proposed methods.

Keywords: Mixed frequencies, temporal aggregation, impulse responses, shock sequences, DL, ARDL, VARDL

JEL Classification: C22, Q43

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Dallas or the Federal Reserve System.

†Alexander Chudik, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX 75201. E-mail: alexander.chudik@gmail.com.

‡Lutz Kilian, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX 75201. E-mail: lkilian2019@gmail.com.

1 Introduction

A recurrent situation faced by applied researchers is that exogenous shocks are observed at daily or weekly frequency, while the outcome variable of interest is measured at monthly or quarterly frequency. There is a large literature on quantifying the responses of macroeconomic aggregates at a given frequency to exogenous shocks measured at the same frequency (see, e.g., Kilian and Lütkepohl, 2017). This literature leaves unanswered the question of how to quantify the responses of macroeconomic aggregates measured at a low frequency to shocks observed at higher frequency.

For example, a policymaker may face an exogenous shock to the daily price of gasoline on a given day of the current month or a sequence of such shocks extending over several days. The question of interest to the policymaker is how this sequence of daily shocks is expected to change current and future realizations of monthly aggregates such as inflation or industrial production growth, given the historical relationship between gasoline prices and the economy. In general, this response will depend on when in the month the sequence of daily gasoline price shocks occurs. Shocks earlier in the month will have a larger impact on the current monthly inflation than shocks occurring later in the month, while the response of future monthly inflation will be the larger, all else equal, the later in the current month the daily shocks occur.

When working with economic time series sampled at different frequencies it has been common, in practice, to aggregate high-frequency variables to a common low frequency and to analyze the joint process sampled at the low frequency. This approach, however, may distort the co-movements among the model variables and bias the estimates of the impulse responses. In this paper, we address this concern by postulating the existence of a vector autoregressive data generating process at high frequency. The high-frequency variables described by this process need not be observed, except for the daily shock of interest. The high-frequency data are temporally aggregated to generate the corresponding low-frequency data. We propose three alternative estimators of the responses of the observed temporally aggregated variables to a sequence of high frequency shocks and derive their asymptotic distribution. Each of these estimators may be implemented as a mean group or pooled estimator, so there are six ways of proceeding in applied work.

Monte Carlo evidence shows that for realistic sample sizes pooled estimators tend to have lower RMSE than mean group estimators, reflecting their greater parsimony. They also tend to be more accurate in inference. The pooled autoregressive-distributed lag (ARDL) and pooled vector

autoregressive-distributed lag (VARDL) estimators tend to be slightly more accurate than the pooled distributed lag (DL) estimator.

An empirical illustration focuses on sequences of daily wholesale gasoline price shocks associated with four exogenous events in global oil markets. These episodes illustrate how different the path of daily gasoline prices may be over a month. We trace the effects of these shock sequences on monthly headline inflation, core inflation and retail gasoline price inflation. Our estimates show that the effect on headline inflation is short-lived and the response of core inflation tends to be muted, arguing against important indirect effects on inflation.

Our work relates to several strands of the literature. One is the literature on temporal aggregation (see, e.g., Lütkepohl, 1987, Rossana and Seater, 1995, Marcellino, 1999, Breitung and Swanson, 2002, and Swanson and Granger, 2012), which in turn gave rise to the development of mixed-frequency VAR models (see Foroni, Ghysels, and Marcellino, 2013, for a review). Much of this literature has focused on forecasting rather than structural impulse response analysis, but there are exceptions. For example, Ghysels (2016) introduced a class of mixed-frequency VAR models that allows the estimation of the impact of high frequency shocks on low-frequency variables when treating the daily shock as predetermined. In related work, Foroni and Marcellino (2014, 2016) found that responses from common low-frequency VAR models that ignore temporal aggregation can be substantially biased and that this bias can be alleviated by the use of mixed-frequency VAR models.

Compared to mixed-frequency VAR models, our approach does not require the user to fully specify the VAR model generating the data, which is an important advantage. Estimation as well as inference using approaches that rely on the full specification of the underlying model is conditional on the model being correctly specified. However, what variables belong in the data generating process and what are the true lag orders is unknown in practice. Unless this uncertainty is recognized and explicitly incorporated into the analysis, statistical inference will be invalid. In practice, it is difficult to take model uncertainty into account in a systematic manner and, as a result, this is rarely done in empirical research. Our approach is not subject to this drawback. Moreover, our approach is more parsimonious, it tends to perform reasonably well even in small samples, it lends itself to establishing the asymptotic properties of the estimator of the impulse responses, and it easily allows for the consideration of sequences of high-frequency shocks.

The work methodologically most closely related to ours is Chudik and Georgiadis (2022), which also does not require full specification of the model. There are four key differences. First, our paper is about the impact of high-frequency shocks on observed low-frequency variable(s), whereas Chudik and Georgiadis (2022) focused on the impact of high-frequency shocks on the underlying latent high-frequency variable(s). Consequently, our paper does not assume that the temporal aggregation weights for the construction of low-frequency variables are known to the researcher, whereas Chudik and Georgiadis (2022) do. Second, we generalize the mixed-frequency DL estimating equations proposed by Chudik and Georgiadis (2022) to mixed-frequency ARDL and VARDL estimating equations. We show that estimators based on ARDL and parsimonious VARDL specifications are prefered because they tend to achieve lower root mean square error. Third, our paper develops pooled estimators that remain applicable in settings where the approach of Chudik and Georgiadis (2022) is not feasible due to a larger frequency mismatch. Our Monte Carlo evidence shows that pooled estimators for realistic sample sizes tend to achieve smaller root mean square error despite these estimators not being asymptotically efficient. Finally, unlike earlier work we focus on the impact of sequences of shocks.

Our empirical analysis is relevant for a recent literature examining the impact of gasoline price shocks on headline and core inflation, especially since the pandemic era, based on low-frequency VAR models. Examples include Kilian and Zhou (2022a, 2022b, 2023), Diab and Karaki (2023), Vatsa and Pino (2024), Gründler and Scharler (2025), and Karaki and Chaar (2026). For a review of this literature see Kilian and Zhou (2025). Our analysis confirms the findings in this literature that the response of headline inflation is short-lived and that the pass-through to core inflation tends to be muted, even after allowing for temporal aggregation. It adds to this literature a more precise estimate of the effects of daily shocks on monthly inflation.

The remainder of the paper is organized as follows. Section 2 introduces the model, notation, and defines the impulse response functions (IRFs) of interest. Section 3 introduces mean group and pooled mixed-frequency DL, ARDL and VARDL estimators, and derives their asymptotic properties. Section 4 provides Monte Carlo evidence. Section 5 presents the empirical application and Section 6 concludes. Appendix provides proofs and additional details for the empirical application. Throughout the paper, matrices are denoted by upper case bold letters. Vectors are denoted by lower case bold letters. All vectors are column vectors.

2 Model

We assume an $n \times 1$ dimensional vector of stationary high-frequency variables $\mathbf{z}_t$ generated by

$$\mathbf{z}_t = \mathbf{\Phi}\mathbf{z}_{t-1} + \mathbf{u}_t, \tag{1}$$

where $\mathbf{\Phi}$ is an $n \times n$ matrix of coefficients and $\mathbf{u}_t$ is an $n \times 1$ vector of serially uncorrelated reduced-form errors. We postulate one lag without loss of generality, since any VAR(p) model can be expressed as a VAR(1) model in companion form.[1] Without loss of generality, $\mathbf{z}_t$ may be partitioned as $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$, where x_t is the outcome variable of interest and $\mathbf{y}_t$ is the $(n-1) \times 1$ vector of the remaining variables. It is not assumed that the variables in $\mathbf{y}_t$ or the dimension n are necessarily observed.

The shock of interest, denoted as e_t, is assumed to be observed and

$$E(\mathbf{u}_t | e_t) = \mathbf{a}e_t, \tag{2}$$

where $\mathbf{a}$ is an $n \times 1$ vector of constants. Then the reduced-form shock $\mathbf{u}_t$ can be partitioned as

$$\mathbf{u}_t = \mathbf{a}e_t + \boldsymbol{\eta}_t, \tag{3}$$

where $\boldsymbol{\eta}_t \equiv \mathbf{u}_t - E(\mathbf{u}_t | e_t)$. The shock e_t may be a "structural" shock, but it does not have to be. It also could be a shock without any structural economic interpretation, such as a reduced-form shock, or some rotation of reduced-form shocks, so long as the assumption $E(\mathbf{u}_t | e_t) = \mathbf{a}e_t$ is met (see Chudik and Georgiadis, 2022, for further discussion).

We assume that the outcome variable, x_t, is observed only at a lower frequency than the shock e_t. In particular, let $m > 1$ be the number of high-frequency time periods contained in one low-frequency period. We refer to the last high-frequency period within the low-frequency period s by t_s, defined as

$$t_s = ms, \text{ for } s = 1, 2, ..., T_m, \tag{4}$$

with $T_m = [T/m]$, where $[T/m]$ is the integer part of T/m. Thus, T is the number of high-frequency

[1] To simplify the exposition, we omit deterministic terms in (1). Introducing deterministic terms come at the expense of added notation without any helpful insights.

time periods for which e_t is observed, and T_m is the number of low-frequency time periods.

We use the following notation to define sequentially sampled variables,

$$x_{is} := x_{t_s-i}, \text{ for } i = 0, 1, 2, ..., m-1, \text{ and } s = 1, 2, ..., T_m, \tag{5}$$

and similarly $e_{is} := e_{t_s-i}$. Temporal averages are defined as

$$\bar{x}_{ws} = \sum_{i=0}^{m-1} w_i x_{is}, \text{ for } s = 1, 2, ..., T_m, \tag{6}$$

where w_i are possibly unknown temporal aggregation weights. $\bar{x}_{ws}$ is the temporally aggregated version of the outcome variable, which is observed only at the low frequency. For example, simple averaging of the high-frequency observations contained in one low-frequency time period is represented by equal temporal aggregation weights $w_i = 1/m$, for $i = 0, 1, ..., m-1$. The definition in equation (6) allows for the temporal aggregation of both stock and flow variables.[2]

The following assumption summarizes which objects in our setup are observed and which are unobserved.

ASSUMPTION 1 *The researcher observes the high-frequency scalar shock e_t in (1), for $t = 1, 2, ..., T$, and the low-frequency temporally aggregated variable $\bar{x}_{ws}$, for $s = 1, 2, ..., T_m$, defined by (6). The aggregation weights $\{w_i\}$, the dimension of $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$, the underlying lag order of VAR model (1), the variables entering $\mathbf{y}_t$ and their realizations are all potentially unobserved by the researcher.*

Remark 1 *The reason that we do not require the full specification of the high-frequency VAR process in Assumption 1 is that e_t is observed.*

The following standard assumptions are imposed on the high-frequency VAR model (1). We use K to denote a generic, finite positive constant that does not depend on the sample size T. This constant can take on different values in different instances.

ASSUMPTION 2 $|\lambda_1(\mathbf{\Phi})| < 1$, *where $\lambda_1(\mathbf{\Phi})$ is the largest eigenvalue of $\mathbf{\Phi}$.*

[2] Our analysis can be extended to allow for temporal aggregation scheme in (6) to span multiple low-frequency periods. For expository purposes and to simplify the notation, we restrict attention to the aggregation scheme in (6).

ASSUMPTION 3 *Innovations* $\mathbf{u}_t$ *are given by (3), where* $E(\mathbf{u}_t|e_t) = \mathbf{a}e_t$, $\boldsymbol{\eta}_t = \mathbf{u}_t - E(\mathbf{u}_t|e_t)$, $e_t \sim IID\left(0, \sigma_e^2\right)$, $\sigma_e^2 > 0$, $E\left(e_t^4\right) = \kappa_4$, $\|\mathbf{a}\| < K$, *and* $\boldsymbol{\eta}_t \sim IID(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$. e_t *is independently distributed of* $\boldsymbol{\eta}_{t'}$ *for all* t, t'. *There exists* $\epsilon > 0$ *such that* $E|e_t|^{4+\epsilon} < K$ *and* $E\|\boldsymbol{\eta}_t\|^{4+\epsilon} < K$.

Remark 2 *Assumption 2 rules out variables integrated of order 1. We assume that such variables are appropriately transformed to stationarity prior to the analysis.*

Remark 3 *Assumption 3 does not take a stand on the economic interpretation of* e_t. *Causal interpretation of the effects of this shock require* x_t *to be pre-determined with respect to* $\mathbf{y}_t$.

Remark 4 *To simplify the derivations of asymptotic distributions, Assumption 3 assumes innovations are i.i.d. This rules out heteroskedastic shocks. We propose estimators of standard errors that are robust to heteroskedasticity, and we allow for heteroskedastic shocks in Monte Carlo simulations.*

2.1 Definitions of impulse response functions of interest and additional notations

Under Assumptions 2-3, we have

$$x_t = \sum_{\ell=0}^{\infty} b_\ell e_{t-\ell} + \varepsilon_t, \tag{7}$$

where

$$b_\ell = \mathbf{s}_{n1}' \boldsymbol{\Phi}^\ell \mathbf{a}, \text{ for } \ell = 0, 1, 2, ..., \tag{8}$$

and

$$\varepsilon_t = \sum_{\ell=0}^{\infty} \mathbf{s}_{n1}' \boldsymbol{\Phi}^\ell \boldsymbol{\eta}_{t-\ell}. \tag{9}$$

in which $\mathbf{s}_{n1} = (1, 0, ..., 0)'$ is an $n \times 1$ selection vector that selects the first element. Consider the high-frequency IRF of a unit disturbance to the high-frequency shock e_t on the outcome variable of interest, x_t,

$$IRF(x, \ell) = b_\ell = E(x_{t+\ell}|e_t = 1, \mathcal{I}_{t-1}) - E(x_{t+\ell}|\mathcal{I}_{t-1}), \text{ for } \ell = 0, 1, ..., \tag{10}$$

where $\mathcal{I}_{t-1} = \{\mathbf{z}_{t-1}, \mathbf{z}_{t-2}, ...\}$ is the information set including all variables up to the time period $t-1$. When aggregation weights $\{w_i\}$ are unknown, it is not possible to estimate b_ℓ.

The main objective in this paper is the estimation of the response of the low-frequency variable to a high-frequency shock sequence

$$\mathcal{I}_{es}\left(\boldsymbol{\omega}\right)=\left\{e_{is}=\omega_i,\ \text{for } i=0,1,...,m-1\right\},$$

where $\boldsymbol{\omega}=\left(\omega_0,\omega_1,...,\omega_{m-1}\right)'$, and, for convenience, we will normalize $\sum_{i=0}^{m-1}\omega_i=1$. The low-frequency IRF corresponding to the sequence of shocks $\mathcal{I}_{es}\left(\boldsymbol{\omega}\right)$ is given by

$$IRF\left(\bar{x}_w,\omega,r\right)=\bar{d}_{\omega r}=E\left(\bar{x}_{w,s+r}|\,\mathcal{I}_{es}\left(\boldsymbol{\omega}\right),\mathcal{I}_{t_{s-1}}\right)-E\left(\bar{x}_{w,s+r}|\,\mathcal{I}_{t_{s-1}}\right),\ \text{for } r=0,1,...\ . \tag{11}$$

Let h be the maximum horizon of interest, and collect the low-frequency impulse-response coefficients $\bar{d}_{\omega r}$, for $r=0,1,2,...,h$, in an $(h+1)\times 1$ vector $\bar{\mathbf{d}}_\omega=\left(\bar{d}_{\omega 0},\bar{d}_{\omega 1},...,\bar{d}_{\omega h}\right)'$. Our objective is the estimation of $\bar{\mathbf{d}}_\omega$. These impulse-response coefficients can be estimated for any choice of $\boldsymbol{\omega}$. We do not assume that the aggregation weights $\{w_i\}$ are known, and no *a priori* restrictions are imposed on the specification of the shock sequence $\{\omega_i\}$.

For future reference, we also define the IRF of a unit disturbance to a single high-frequency shock to e_t on the observed aggregated outcome variable of interest, $\bar{x}_{ws}$,

$$IRF\left(\bar{x}_w,i,r\right)=d_{ir}=E\left(\bar{x}_{w,s+r}|\,e_{is}=1,\mathcal{I}_{t_{s-1}}\right)-E\left(\bar{x}_{w,s+r}|\,\mathcal{I}_{t_{s-1}}\right), \tag{12}$$

for $r=0,1,...$, and $i=0,1,...,m-1$, where, given (1) and Assumptions 2-3, we have

$$d_{i0} = \sum_{q=0}^{i} w_q b_{i-q},\ \text{for } i=0,1,...,m-1, \tag{13}$$

$$d_{ir} = \sum_{q=0}^{m-1} w_q b_{mr+i-q},\ \text{for } r=1,2,...,h,\ \text{and } i=0,1,...,m-1. \tag{14}$$

The IRF coefficient $\bar{d}_{\omega r}$ given by (11) can, in the context of the model in this paper, be written as a weighted average of d_{ir}, namely

$$\bar{d}_{\omega r}=\sum_{i=0}^{m-1}\omega_i d_{ir},\ \text{for } r=0,1,2,.... \tag{15}$$

For future reference, the temporally aggregated version of the moving average representation

(7) is given by

$$\bar{x}_{ws} = \sum_{i=0}^{m-1} \sum_{r=0}^{\infty} d_{ir} e_{i,s-r} + \bar{\varepsilon}_{ws}, \tag{16}$$

where $\bar{\varepsilon}_{ws} = \sum_{i=0}^{m-1} w_i \varepsilon_{is}$ (see also Lemma 1 of Chudik and Georgiadis, 2022). Using vector notation, collect the coefficients $\{d_{ir}, r = 0, 1, ..., h, i = 0, 1, ..., m-1\}$ in an $(h+1)\, m \times 1$ coefficient vector $\mathbf{d} = (\mathbf{d}_0', \mathbf{d}_1', ..., \mathbf{d}_h')'$, where $\mathbf{d}_r = (d_{0r}, d_{1r}, ..., d_{m-1,r})'$ for $r = 0, 1, ..., h$. Then

$$\bar{\mathbf{d}}_\omega = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right) \mathbf{d}, \tag{17}$$

where $\mathbf{I}_{h+1}$ is $(h+1) \times (h+1)$ identity matrix.

3 Mean Group and Pooled Mixed-Frequency Estimators

We propose estimators of $\bar{\mathbf{d}}_\omega$ based on distributed lag (DL), autoregressive distributed lag (ARDL), and vector autoregressive distributed lag (VARDL) estimating equations below. For each of the these estimating equations, we consider mean group and pooled estimators.

3.1 Estimators based on DL estimating equations

3.1.1 Mean Group DL estimator

Chudik and Georgiadis (2022) introduced a mixed-frequency distributed lag (DL) estimator of d_{ir} based on the Least Squares (LS) estimate of the auxiliary regression

$$\bar{x}_{ws} = \sum_{i=0}^{m-1} \sum_{r=0}^{h} d_{ir} e_{i,s-r} + \vartheta_{hs}, \tag{18}$$

which is obtained by truncating (16), where

$$\vartheta_{hs} = \bar{\varepsilon}_{ws} + \sum_{i=0}^{m-1} \sum_{r=h+1}^{\infty} d_{ir} e_{i,s-r}. \tag{19}$$

Define the $(h+1)\, m \times 1$ dimensional vector of sequentially sampled errors and their lags $\mathbf{e}_{hs} = \left(\mathbf{e}_s', \mathbf{e}_{s-1}', ..., \mathbf{e}_{s-h}'\right)'$ with $\mathbf{e}_s = (e_{0s}, e_{1s}, ..., e_{m-1,s})'$. Using vector notations, (18) can be compactly

written as

$$\bar{x}_{ws} = \mathbf{d}'\mathbf{e}_{hs} + \vartheta_{hs}. \tag{20}$$

The DL estimator of $\mathbf{d}$, denoted by $\widehat{\mathbf{d}} = \left(\hat{\mathbf{d}}_0', \hat{\mathbf{d}}_1', ..., \hat{\mathbf{d}}_h'\right)'$ where $\hat{\mathbf{d}}_r = \left(\hat{d}_{0r}, \hat{d}_{1r}, ..., \hat{d}_{m-1,r}\right)'$ for $r = 0, 1, 2, ...h$, is given by

$$\widehat{\mathbf{d}} = \left(\mathbf{E}'\mathbf{E}\right)^{-1}\mathbf{E}'\bar{\mathbf{x}}_w. \tag{21}$$

$\bar{\mathbf{x}}_w = (\bar{x}_{w,h+1}, \bar{x}_{w,h+2}, ..., \bar{x}_{w,T_m})'$ is the $(T_m - h)\times 1$ vector of observations on the dependent variable $\bar{x}_{ws}$ in (20), and $\mathbf{E}$ is the $(T_m - h) \times (h+1)\, m$ matrix of observations on regressors $\bar{\mathbf{e}}_{hs}$ in (20), namely $\mathbf{E} = (\mathbf{E}_0, \mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_h)$, where $\mathbf{E}_r = (\mathbf{e}_{h-r+1}, \mathbf{e}_{h-r+2}, ..., \mathbf{e}_{T_m-r})'$, for $r = 0, 1, ..., h$.

Estimation of $\bar{d}_{\omega r}$ involves plugging the DL estimates $\hat{d}_{ir}$ into (15), yielding

$$\widehat{\bar{d}}_{\omega r}^{MG} = \sum_{i=0}^{m-1} \omega_i \hat{d}_{ir}, \text{ for } r = 0, 1, 2, ...h, \tag{22}$$

or $\widehat{\bar{\mathbf{d}}}_{\omega}^{MG} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right)\hat{\mathbf{d}}$. We refer to this approach as the mean group DL estimator. The following theorem establishes the asymptotic distribution of $\widehat{\bar{\mathbf{d}}}_{\omega}^{MG}$ as $T \to \infty$.

Proposition 1 *Suppose $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$ is given by (1), and Assumptions 2-3 hold. Consider the mean group DL estimator $\widehat{\bar{\mathbf{d}}}_{\omega}^{MG} = \left(\widehat{\bar{d}}_{\omega 0}^{MG}, \widehat{\bar{d}}_{\omega 1}^{MG}, ..., \widehat{\bar{d}}_{\omega h}^{MG}\right)'$ given by (22) and let $T_{mh} = T/m - h$. Then, for fixed m and h, as $T \to \infty$,*

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}_{\omega}^{MG} - \bar{\mathbf{d}}_{\omega}\right) \to_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{MG}\right), \tag{23}$$

where

$$\boldsymbol{\Omega}_{MG} = \sigma_e^{-2}\left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right)\Xi\left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}\right), \tag{24}$$

$\sigma_e^2 = E\left(e_t^2\right)$ is the variance of e_t, $\mathbf{I}_{h+1}$ is $(h+1)\times(h+1)$ identity matrix,

$$\boldsymbol{\Xi} = \sum_{\ell=-h}^{h} \gamma_{h\ell}\mathbf{H}_{h\ell}, \tag{25}$$

$\gamma_{h\ell} = E\left(\vartheta_{hs}\vartheta_{h,s-\ell}\right)$ is the autocovariance function of ϑ_{hs} defined in (19), $\mathbf{H}_{h\ell} = \mathbf{G}_{h\ell} \otimes \mathbf{I}_m$, $\mathbf{G}_{h\ell}$ is

$(h+1)\times(h+1)$ *dimensional shift matrix with its* (i,j)*-th element given by* $\delta_{i+\ell,j}$*, and*

$$\delta_{ij}=\begin{cases}1 & \text{for } i=j\\ 0 & \text{for } i\neq j\end{cases}, \tag{26}$$

is the Kronecker delta.

All proofs are presented in Appendix.

Under homoskedasticity, $\mathbf{\Omega}_{MG}$ can be consistently estimated as

$$\hat{\mathbf{\Omega}}_{MG}=\hat{\sigma}_e^{-2}\left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}'\right)\hat{\mathbf{\Xi}}\left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}\right), \tag{27}$$

where $\hat{\sigma}_e^2=T^{-1}\sum_{t=1}^{T}e_t^2$, $\hat{\mathbf{\Xi}}=\sum_{\ell=-h}^{h}\hat{\gamma}_{h\ell}\mathbf{H}_{h\ell}$, and $\hat{\gamma}_{h\ell}$ are the sample autocovariances of $\hat{\vartheta}_{hs}=\bar{x}_{ws}-\hat{\mathbf{d}}'\mathbf{e}_{hs}$, in which $\hat{\mathbf{d}}$ is given by (21). In the case of heteroskedastic errors, the following heteroskedasticity-consistent estimator can be used

$$\hat{\mathbf{\Omega}}_{MG,HC}=\left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}'\right)\hat{\mathbf{V}}\left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}\right), \tag{28}$$

where $\hat{\mathbf{V}}=T_{mh}\left(\mathbf{E}'\mathbf{E}\right)^{-1}\mathbf{E}'diag\left(\hat{\vartheta}_{h,h+1},\hat{\vartheta}_{h,h+2},...,\hat{\vartheta}_{h,T_m}\right)\mathbf{E}\left(\mathbf{E}'\mathbf{E}\right)^{-\mathbf{1}}$.

3.1.2 Pooled DL estimator

The downside of the mean group DL estimator is that the number of coefficients to be estimated in (20) is $(h+1)m$ (abstracting from the deterministic terms, such as the intercept). As a result, the mean group DL estimator is not feasible when the number of regressors exceeds the regression sample size, namely when $(h+1)m>T_m-h$. When $\widehat{\overline{\mathbf{d}}}_{\omega}^{MG}$ is feasible but the product $(h+1)m$ is relatively large, this estimator is unlikely to perform well. In our empirical application in Section 5, $m=21$ (working days in a month), which is quite large. We therefore propose a pooled alternative to the mean group DL estimator next, which is more parsimonious, feasible for $T_m>2h+1$, and tends to be more accurate for practically relevant sample sizes.

Consider pooling the coefficients d_{ir}, for $i=0,1,...,m-1$. Since we are interested in weighted averages of d_{ir}, namely coefficients $\bar{d}_{\omega r}$ given by (15), we shall employ appropriately weighted

pooling to ensure consistent estimation of $\bar{d}_{\omega r}$. To this end, let

$$\tilde{\omega}_i = \frac{\omega_i}{\sum_{j=0}^{m-1} \omega_j^2}, \tag{29}$$

and define the corresponding weighted aggregate of sequentially sampled shocks

$$\tilde{e}_s := \bar{e}_{\tilde{\omega} s} = \sum_{i=0}^{m-1} \tilde{\omega}_i e_{is}. \tag{30}$$

For notational brevity, below we use $\tilde{e}_s$ as opposed to the more cumbersome $\bar{e}_{\tilde{\omega} s}$. The pooled estimation is based on the auxiliary DL regression

$$\bar{x}_{ws} = \sum_{r=0}^{h} \bar{d}_{\omega r} \tilde{e}_{s-r} + \vartheta_{hs}^*, \tag{31}$$

where

$$\vartheta_{hs}^* = \sum_{i=0}^{m-1} \sum_{r=0}^{h} \Delta_{ir} e_{i,s-r} + \vartheta_{hs}, \tag{32}$$

ϑ_{hs} is given by (19), and

$$\Delta_{ir} = d_{ir} - \tilde{\omega}_i \bar{d}_{\omega r}. \tag{33}$$

It is helpful to re-write (31), in matrix notation, as

$$\bar{x}_{ws} = \bar{\mathbf{d}}_{\omega}' \tilde{\mathbf{e}}_s + \vartheta_{hs}^*, \tag{34}$$

where, $\tilde{\mathbf{e}}_s = (\tilde{e}_s, \tilde{e}_{s-1}, ..., \tilde{e}_{s-h})'$. The corresponding LS estimator, which we denote as pooled DL, is given by

$$\hat{\mathbf{d}}_{\omega}^P = \left(\tilde{\mathbf{E}}' \tilde{\mathbf{E}} \right)^{-1} \tilde{\mathbf{E}}' \bar{\mathbf{x}}_w, \tag{35}$$

where $\tilde{\mathbf{E}}$ is the $(T_m - h) \times (h+1)$ matrix of observations on $\tilde{\mathbf{e}}_s$ in (34), namely $\tilde{\mathbf{E}} = (\tilde{\mathbf{e}}_{h+1}, \tilde{\mathbf{e}}_{h+2}, ..., \tilde{\mathbf{e}}_{T_m})'$, and as before $\bar{\mathbf{x}}_w = (\bar{x}_{w,h+1}, \bar{x}_{w,h+2}, ..., \bar{x}_{w,T_m})'$ is the vector of observations on the dependent variable $\bar{x}_{ws}$. The following proposition establishes the asymptotic normality of the pooled DL estimator.

Proposition 2 *Suppose $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$ is given by (1), and Assumptions 2-3 hold. Consider the*

pooled DL estimator $\widehat{\bar{\mathbf{d}}}_{\omega}^{P}$ *given by (35) and let* $T_{mh} = T/m - h$*. Then, for fixed* m *and* h*, as* $T \to \infty$*,*

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}_{\omega}^{P} - \bar{\mathbf{d}}_{\omega}\right) \to_{d} N\left(\mathbf{0}, \mathbf{\Omega}_{P}\right), \tag{36}$$

where

$$\mathbf{\Omega}_{P} = \sigma_{e}^{-4}\left\|\boldsymbol{\omega}\right\|^{4}\left(\sigma_{e}^{2}\tilde{\mathbf{\Xi}} + \mathbf{\Sigma}_{\Delta}\right), \tag{37}$$

$\sigma_{e}^{2} = E\left(e_{t}^{2}\right)$ *is the variance of* e_{t}*,* $\left\|\boldsymbol{\omega}\right\| = \sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}$ *is the Euclidean norm of* $\boldsymbol{\omega}$*,*

$$\tilde{\mathbf{\Xi}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right)\mathbf{\Xi}\left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right), \tag{38}$$

$\mathbf{\Xi}$ *is given by (25),*

$$\mathbf{\Sigma}_{\Delta} = \left[\mathbf{\Delta}' \otimes \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right)\right]\mathbf{\Sigma}_{E}\left[\mathbf{\Delta} \otimes \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right)\right], \tag{39}$$

$\tilde{\boldsymbol{\omega}} = \left\|\boldsymbol{\omega}\right\|^{-2}\boldsymbol{\omega}$*,* $\mathbf{\Sigma}_{E}$ *is the asymptotic variance of* $T_{mh}^{-1/2}Vec\left(\mathbf{E}'\mathbf{E} - \sigma_{e}^{2}\mathbf{I}_{(h+1)m}\right)$*,* $\mathbf{\Delta} = \left(\mathbf{\Delta}_{0}', \mathbf{\Delta}_{1}', ..., \mathbf{\Delta}_{h}'\right)'$*,* $\mathbf{\Delta}_{r} = \left(\Delta_{0r}, \Delta_{1r}, ..., \Delta_{m-1,r}\right)'$ *for* $r = 0, 1, ..., h$*, and* Δ_{ir} *is given by (33).*

Under homoskedasticity, $\mathbf{\Omega}_{P}$ can by consistently estimated as

$$\hat{\mathbf{\Omega}}_{P} = \hat{\sigma}_{\tilde{e}}^{-2}\sum_{\ell=-h}^{h}\hat{\gamma}_{h\ell}^{*}\mathbf{H}_{h,\ell}, \tag{40}$$

where $\hat{\sigma}_{\tilde{e}}^{2} = T_{mh}^{-1}\sum_{s=1}^{T_{mh}}\tilde{e}_{s}$, and $\hat{\gamma}_{h\ell}^{*}$ are the sample autocovariances of $\hat{\vartheta}_{hs}^{*} = \bar{x}_{w,s} - \widehat{\bar{\mathbf{d}}}_{\omega}^{P\prime}\tilde{\mathbf{e}}_{s}$, in which $\widehat{\bar{\mathbf{d}}}_{\omega}^{P}$ is given by (35). A heteroskedasticity-consistent variance-covariance matrix estimator is given by

$$\hat{\mathbf{\Omega}}_{P,HC} = T_{mh}\left(\tilde{\mathbf{E}}'\tilde{\mathbf{E}}\right)^{-1}\tilde{\mathbf{E}}'diag\left(\hat{\vartheta}_{h,h+1}^{*}, \hat{\vartheta}_{h,h+2}^{*}, ..., \hat{\vartheta}_{h,T_{m}}^{*}\right)\tilde{\mathbf{E}}\left(\tilde{\mathbf{E}}'\tilde{\mathbf{E}}\right)^{-1}.$$

Remark 5 *The asymptotic variance* $\mathbf{\Omega}_{P}$ *cannot be smaller than* $\mathbf{\Omega}_{MG}$*.* $\mathbf{\Omega}_{P}$ *is adversely affected by 'heterogeneity' of* d_{ir}*, as captured by* $\Delta_{ir} = d_{ir} - \tilde{\omega}_{i}\bar{d}_{\omega r}$ *defined in (33). When* $\Delta_{ir} = 0$ *for all* i *and* r*, then* $\mathbf{\Sigma}_{\Delta} = \mathbf{0}$ *and* $\mathbf{\Omega}_{P}$ *reduces to*

$$\mathbf{\Omega}_{P} = \sigma_{e}^{-2}\left\|\boldsymbol{\omega}\right\|^{4}\left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right)\mathbf{\Xi}\left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right),$$

which, using $\tilde{\boldsymbol{\omega}} = \|\boldsymbol{\omega}\|^{-2}\boldsymbol{\omega}$, $\boldsymbol{\Omega}_P$ *(for* $\Delta_{ir} = 0$*) further simplifies to*

$$\boldsymbol{\Omega}_P = \sigma_e^{-2}\left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right)\boldsymbol{\Xi}\left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}\right) = \boldsymbol{\Omega}_{MG},$$

and coincides with $\boldsymbol{\Omega}_{MG}$ *in (24). The advantage of using the pooled DL estimator is not its asymptotic efficiency but rather its higher small-sample accuracy. As our Monte Carlo results demonstrate, the pooled DL estimator tends to have lower root mean square error compared with the mean group DL estimator in sample sizes of practical interest. Moreover, the pooled DL estimator may be feasible in small samples even when the mean group DL estimator is not.*

Remark 6 *Our result that* $\widehat{\bar{\mathbf{d}}}_\omega^P$ *is not asymptotically more efficient than* $\widehat{\bar{\mathbf{d}}}_\omega^{MG}$ *might appear counter-intuitive. It arises from the orthogonality of the regressors in* $\mathbf{E}$ *in conjunction with the need to use a consistent pooling scheme given by* $\tilde{\boldsymbol{\omega}} = \|\boldsymbol{\omega}\|^{-2}\boldsymbol{\omega}$*, see (29). To clarify why this is the case, consider a simple illustrative example where a variable* y_t *is generated by*

$$y_t = x_{1t}\beta_1 + x_{2t}\beta_2 + u_t, \text{ for } t = 1, 2, ..., T,$$

with $u_t \sim IID\left(0, \sigma_u^2\right)$*. Then, under the usual regularity requirements, the LS estimator of* $\boldsymbol{\beta} = (\beta_1, \beta_2)'$*, is asymptotically distributed as*

$$\sqrt{T}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}\sigma_u^2\right),$$

where $\mathbf{X}$ *is* $T \times 2$ *matrix of observations on* $\mathbf{x}_t = (x_{1t}, x_{2t})'$*. Let* $\bar{\beta}_\omega = \boldsymbol{\omega}'\boldsymbol{\beta}$*, and consider the 'mean group' estimator of* $\bar{\beta}_\omega$*, given by* $\widehat{\bar{\beta}}_\omega^{MG} = \boldsymbol{\omega}'\hat{\boldsymbol{\beta}}$*. Then*

$$\sqrt{T}\left(\widehat{\bar{\beta}}_\omega^{MG} - \bar{\beta}_\omega\right) \to_d N\left(0, \boldsymbol{\omega}'\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}\boldsymbol{\omega}\sigma_u^2\right).$$

The 'pooled' estimator of $\bar{\beta}_\omega$ *is given by regression of* y_t *on* $\tilde{x}_t = \tilde{\boldsymbol{\omega}}'\mathbf{x}_t$*. In the ideal case when the slope coefficients are homogenous* ($\beta_1 = \beta_2 = \beta$) *and* $\boldsymbol{\omega} = (1, 1)'$*, the pooled estimator is asymptotically distributed as*

$$\sqrt{T}\left(\widehat{\bar{\beta}}_\omega^P - \bar{\beta}_\omega\right) \to_d N\left(0, \left(\frac{\tilde{\mathbf{x}}'\tilde{\mathbf{x}}}{T}\right)^{-1}\sigma_u^2\right),$$

where $\tilde{\mathbf{x}}$ *is* $T \times 1$ *vector of observations on* $\tilde{x}_t$*. Under these assumptions, we expect*

$$Avar\left(\widehat{\tilde{\beta}}_{\omega}^{P}\right) \equiv \left(\frac{\tilde{\mathbf{x}}'\tilde{\mathbf{x}}}{T}\right)^{-1}\sigma_u^2 \leq Avar\left(\hat{\beta}^{MG}\right) \equiv \boldsymbol{\omega}'\left(\frac{\mathbf{X}'\mathbf{X}}{T}\right)^{-1}\boldsymbol{\omega}\sigma_u^2. \tag{41}$$

When $T^{-1}\mathbf{X}'\mathbf{X} \to_p \mathbf{I}_2$*, then* $\left(T^{-1}\tilde{\mathbf{x}}'\tilde{\mathbf{x}}\right)^{-1} \to_p \left(\tilde{\boldsymbol{\omega}}'\tilde{\boldsymbol{\omega}}\right)^{-1} = \|\boldsymbol{\omega}\|^2$ *and* $\boldsymbol{\omega}'\left(T^{-1}\mathbf{X}'\mathbf{X}\right)^{-1}\boldsymbol{\omega} = \|\boldsymbol{\omega}\|^2$*. Hence,* $Avar\left(\hat{\beta}^{P}\right) = Avar\left(\hat{\beta}^{MG}\right)$*, and the pooled estimator has the same asymptotic variance as the mean group estimator when* $\beta_1 = \beta_2 = \beta$ *and* $\boldsymbol{\omega} = (1,1)'$*. This would not longer be true if* $T^{-1}\mathbf{X}'\mathbf{X} \to_p \mathbf{I}_2$ *did not hold.*

3.2 Estimators based on ARDL estimating equations

3.2.1 Mean Group ARDL estimator

Next, we propose augmenting the DL estimating equations (18) with p lags of the dependent variable. Using (16), the lagged dependent variables, $\bar{x}_{w,s-r}$ for $r = 1, 2, ..., p$, can be written as

$$\bar{x}_{w,s-r} = \sum_{i=0}^{m-1}\sum_{\ell=r}^{\infty} d_{i,s-\ell}e_{i,s-\ell} + \bar{\varepsilon}_{w,s-r}. \tag{42}$$

Augmenting (18) with $\{\bar{x}_{w,s-r}\}_{r=1}^{p}$, and using (42), we obtain

$$\bar{x}_{ws} = \sum_{r=1}^{p}\psi_{pr}\bar{x}_{w,s-r} + \sum_{i=0}^{m-1}\sum_{r=0}^{h}\beta_{pir}e_{i,s-r} + v_{phs}, \tag{43}$$

where under Assumption 3 the error term, v_{phs}, is distributed independently of $\{e_{i,s-r}, i = 0, 1, ..., m-1, r = 0, 1, ..., h\}$, and can be written as

$$v_{phs} = \sum_{i=0}^{m-1}\sum_{r=h+1}^{\infty}\beta_{pir}e_{i,s-r} + \bar{\varepsilon}_{w,s} - \sum_{r=1}^{p}\psi_{pr}\bar{\varepsilon}_{w,s-r}. \tag{44}$$

Given that the error term of the original, not-augmented DL representation (18) does not contain $\{e_{i,s-r}, i = 0, 1, ..., m-1, r = 0, 1, ..., h\}$, the population coefficients on the sequentially sampled

shocks in the augmented representation (43) must satisfy

$$\begin{aligned}
\beta_{pi0} &= d_{i0}, && (45)\\
\beta_{pi1} &= d_{i1} - \psi_{p1} d_{i0}, && (46)\\
\beta_{pi2} &= d_{i2} - \psi_{p1} d_{i1} - \psi_{p2} d_{i0}, && (47)\\
&\vdots &&\\
\beta_{pi\ell} &= d_{i\ell} - \sum_{r=1}^{p} \psi_{pr} d_{i,\ell-r}, && (48)
\end{aligned}$$

for $\ell = p, p+1, ..., h$, and for $i = 0, 1, ..., m-1$, regardless of the choice of the lag order p and regardless of the population values of the autoregressive coefficients $\{\psi_{pr}, r = 1, 2, ..., p\}$. We use the subscript p to highlight the dependence of the coefficients of (43) on the choice of the lag order p. It is useful to define polynomials

$$\psi(p, L) = 1 - \sum_{r=1}^{p} \psi_{pr} L^r,\ \beta_i(p, L) = \sum_{r=0}^{\infty} \beta_{pir} L^r, \text{ and } d_i(L) = \sum_{r=0}^{\infty} d_{ir} L^r,$$

for $i = 0, 1, ..., m-1$, where L is the low-frequency lag operator ($L\bar{x}_{ws} = \bar{x}_{w,s-1}$ and $Le_{is} = e_{i,s-1}$). Equations (45)-(48) imply that

$$\beta_i(p, L) = \psi(p, L) d_i(L). \tag{49}$$

Hence, regardless of the choice of the lag order, p, and regardless of the population values of $\{\psi_{pr}\}_{r=1}^{p}$, the impulse-response coefficients $\{d_{ir},\ i = 0, 1, ..., m-1,\ r = 0, 1, ..., h\}$ can be recovered from the coefficients of the ARDL model (43). These coefficients are recovered recursively as

$$\begin{aligned}
d_{i0} &= \beta_{pi0}, && (50)\\
d_{i1} &= \beta_{pi1} + \psi_{p1} d_{i0}, && (51)\\
d_{i2} &= \beta_{pi2} + \psi_{p1} d_{i1} + \psi_{p2} d_{i0}, && (52)\\
&\vdots &&\\
d_{i\ell} &= \beta_{pi\ell} + \sum_{r=1}^{p} \psi_{pr} d_{i,\ell-r}, && (53)
\end{aligned}$$

for $\ell = p, p+1, ...h$, and for $i = 0, 1, ..., m-1$. For future reference, let $\boldsymbol{\psi}_p = (\psi_{p1}, \psi_{p2}, ..., \psi_{pp})'$ and $\boldsymbol{\beta}_{pi} = (\beta_{pi0}, \beta_{pi1}, ..., \beta_{pih})'$, for $i = 1, 2, ..., m-1$. Collect the ARDL coefficients of (43) in a vector

$\boldsymbol{\theta}_p = \left(\boldsymbol{\psi}_p', \boldsymbol{\beta}_{p0}', \boldsymbol{\beta}_{p1}', ..., \boldsymbol{\beta}_{p,m-1}'\right)'$, and denote the mapping defined by (50)-(53) as $F\left(\boldsymbol{\theta}_p\right) = \mathbf{d}$.

Under Assumptions 2-3, $\bar{x}_{ws}$ and e_{is} are covariance stationary, the LS estimator of the ARDL coefficients in (43), denoted as $\hat{\boldsymbol{\theta}}_p$, is asymptotically normally distributed (see Lemma 4), and

$$\sqrt{T_m}\left(\hat{\boldsymbol{\theta}}_p - \boldsymbol{\theta}_p\right) \rightarrow_d N\left(\mathbf{0}, \boldsymbol{\Sigma}_{\theta p}\right), \tag{54}$$

as $T \rightarrow \infty$, where $T_m = T/m$. The mean group estimator based on the auxiliary ARDL representation (43), is then given by

$$\hat{\bar{d}}_{\omega pr}^{MGA} = \sum_{i=0}^{m-1} \omega_i \hat{d}_{pir}^{A}, \text{ for } r = 0, 1, 2, ...h, \tag{55}$$

where $\hat{d}_{pir}^{A}$ are the ARDL based estimates of d_{ir} given by the elements of $\hat{\mathbf{d}}_p^A = F\left(\hat{\boldsymbol{\theta}}_p\right)$. The following proposition establishes the asymptotic normality of the mean group ARDL estimator.

Proposition 3 *Suppose $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$ is given by (1), and Assumptions 2-3 hold. Consider the mean group ARDL estimator $\hat{\bar{\mathbf{d}}}_{\omega p}^{MGA} = \left(\hat{\bar{d}}_{\omega p0}^{MGA}, \hat{\bar{d}}_{\omega p1}^{MGA}, ..., \hat{\bar{d}}_{\omega ph}^{MGA}\right)'$, where $\hat{\bar{d}}_{\omega pr}^{MGA}$ for $r = 0, 1, ..., h$ are given by (55), and let $T_{mh} = T/m - h$. Then, for fixed m, p, and h, as $T \rightarrow \infty$,*

$$\sqrt{T_{mh}}\left(\hat{\bar{\mathbf{d}}}_{\omega p}^{MGA} - \bar{\mathbf{d}}_{\omega}\right) \rightarrow_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{MGA,p}\right), \tag{56}$$

where

$$\boldsymbol{\Omega}_{MGA,p} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right) \mathbf{J}_F\left(\boldsymbol{\theta}_p\right) \boldsymbol{\Sigma}_{\theta p} \mathbf{J}_F'\left(\boldsymbol{\theta}_p\right) \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}\right),$$

$\mathbf{J}_F\left(\boldsymbol{\theta}_p\right)$ is the Jacobian of the mapping $F\left(\boldsymbol{\theta}_p\right) = \mathbf{d}$ defined by (50)-(53), and $\boldsymbol{\Sigma}_{\theta p}$ is the asymptotic variance of $\hat{\boldsymbol{\theta}}_p$.

$\boldsymbol{\Omega}_{MGA,p}$ can be consistently estimated as

$$\hat{\boldsymbol{\Omega}}_{MGA,p} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right) \mathbf{J}_F\left(\hat{\boldsymbol{\theta}}_p\right) \hat{\boldsymbol{\Sigma}}_{\theta p} \mathbf{J}_F'\left(\hat{\boldsymbol{\theta}}_p\right) \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}\right), \tag{57}$$

where $\hat{\boldsymbol{\Sigma}}_{\theta p}$ is a consistent estimator of $\boldsymbol{\Sigma}_{\theta p}$. The Newey-West estimator of $\boldsymbol{\Sigma}_{\theta p}$ can be used in the presence of heteroskedastic and serially correlated ARDL residuals.

We have established that so long as h lags of sequentially sampled shocks are included in (43),

$\widehat{d}_{\omega r}^{MGA}$ for $r = 0, 1, ..., h$ will be consistent, regardless of p, even if p is not chosen to be a sufficiently large lag order for residuals to be serially uncorrelated.

The choice of the lag orders warrants further discussion. Consider the moving average representation (16) for $\bar{x}_{ws}$, where $\bar{\varepsilon}_{ws}$ is the temporal aggregate of $\varepsilon_t = \sum_{\ell=0}^{\infty} \mathbf{s}'_{n,1} \mathbf{\Phi}^{\ell} \boldsymbol{\eta}_{t-\ell}$. Assumptions 2-3 imply that the coefficients of $d_i(L) = \sum_{r=0}^{\infty} d_{ir} L^r$ are exponentially decaying in r, and the autocovariances of $\bar{\varepsilon}_{ws}$ are also exponentially decaying. To illustrate this, consider coefficients $\{b_\ell\}$ defined by (8). Using submultiplicative properties of matrix norms, we can bound b_ℓ as

$$|b_\ell| = \left\|\mathbf{s}'_{n,1}\right\| \left\|\mathbf{\Phi}^{\ell}\right\| \|\mathbf{a}\|, \text{ for } \ell = 0, 1, 2, ...,$$

for any choice of matrix norm $\|.\|$. We have $\left\|\mathbf{s}'_{n,1}\right\| = 1$. Assumptions 2-3 ensure the existence of constants $K > 0$ and $0 \leq \rho < 1$ such that $\left\|\mathbf{\Phi}^{\ell}\right\| \|\mathbf{a}\| < K\rho^{\ell}$, $0 \leq \rho < 1$. Hence $|b_\ell| < K\rho^{\ell}$, and using (13)-(14) we obtain that d_{ir} is exponentially decaying in r, namely there exist constants $K > 0$ and $0 \leq \rho < 1$ such that $|d_{ir}| < K\rho^{mr}$. Consider next

$$E(\varepsilon_t \varepsilon_{t-\ell}) = \sum_{j=0}^{\infty} \mathbf{s}'_{n1} \mathbf{\Phi}^{j+\ell} \mathbf{\Sigma}_{\eta} \mathbf{\Phi}^{j} \mathbf{s}_{n1}.$$

Taking any norm, such as the Euclidean norm, and using Assumptions 2-3, the autocovariances of ε_t decay exponentially, namely $|E(\varepsilon_t \varepsilon_{t-\ell})| < K\rho^{\ell}$, for some $0 \leq \rho < 1$. Noting $\bar{\varepsilon}_{ws} = \sum_{i=0}^{m-1} w_i \varepsilon_{is}$, we have

$$E(\bar{\varepsilon}_{ws} \bar{\varepsilon}_{w,s-r}) = \sum_{i=0}^{m-1} \sum_{i'=0}^{m-1} w_i w_{i'} E\left(\varepsilon_{is} \varepsilon_{i',s-r}\right).$$

Since $|E(\varepsilon_t \varepsilon_{t-\ell})| < K\rho^{\ell}$, $0 \leq \rho < 1$, $\{w_i\}$ are bounded, and m is finite, there must exist a finite positive constant K such that $|E(\bar{\varepsilon}_{ws} \bar{\varepsilon}_{w,s-r})| < K\rho^{mr}$. Hence, we have established that the autocovariances of $\bar{\varepsilon}_{ws}$ are exponentially decaying, and the coefficients of $d_i(L)$ are exponentially decaying as well. It now follows from (16), that $\bar{x}_{ws}$ has exponentially decaying autocovariances.

Since the process $\bar{\varepsilon}_{ws}$ is covariance stationary with exponentially decaying autocovariances, by Wold's decomposition theorem (Wold, 1938), it has MA(∞) representation,

$$\bar{\varepsilon}_{ws} = \alpha(L) v_s, \tag{58}$$

where $v_s = \bar{\varepsilon}_{ws} - \hat{E}\left(\bar{\varepsilon}_{ws} | \bar{\varepsilon}_{w,s-1}, \bar{\varepsilon}_{w,s-2}, ...\right)$, and $\hat{E}$ denotes the linear prediction operator. Polynomial $\alpha(L)$ and the innovations v_s depend on w but to simplify the notation, we suppress this subscript. If $\alpha(L)$ is invertible, then multiplying (16) by $\psi(L) = \alpha^{-1}(L)$, we obtain the ARDL representation

$$\psi(L)\bar{x}_{w,s} = \sum_{i=0}^{m-1} \beta_i(L) e_{is} + v_s, \tag{59}$$

where $\beta_i(L)$ is given by

$$\beta_i(L) = \psi(L) d_i(L).$$

This differs from (49) because $\beta_i(L)$ and $\psi(L)$ are of infinite orders in general, becasue we have assumed invertibility of $\alpha(L)$, and becasue the error term v_s in (59) is serially uncorrelated. In a special case of model (1) when the polynomials $\psi(L)$ and $\beta_i(L)$ happen to be finite-order polynomials, of orders p and q, respectively, these lag orders will suffice for recovering the first $h+1$ coefficients of $d_i(L)$, even if $h >> q$. However, in general these lag orders are not finite. If the coefficients of $\alpha^{-1}(L)$ decay exponentially, then it is possible to deal with infinite lag orders by allowing $p = p(T)$ (for lagged dependent variable) and $q = q(T)$ (for sequentially sampled shocks) to both increase with the sample size, at an appropriate rate that is not too fast and not too slow, such as the rate $T^{1/3}$, as in e.g. Said and Dickey (1984) or Gonçalves and Kilian (2007) in the context of approximating ARMA(p,q) or AR(∞) models by autoregressions. None of these additional assumptions are required for Proposition 3.

3.2.2 Pooled ARDL estimator

An alternative is to augment the pooled DL estimating equation (31) by p lags of the dependent variable. Using the identity

$$\sum_{i=0}^{m-1} d_{ir} e_{i,s-r} = \bar{d}_{\omega r} \tilde{e}_{s-r} + \sum_{i=0}^{m-1} \Delta_{ir} e_{i,s-r},$$

in the moving average representation (16), we can write the lags of $\bar{x}_{ws}$ as

$$\begin{aligned}\bar{x}_{w,s-r} &= \sum_{i=0}^{m-1}\sum_{\ell=r}^{\infty} d_{i,\ell-r}e_{i,s-\ell} + \bar{\varepsilon}_{w,s-r} \\ &= \sum_{\ell=r}^{\infty} \bar{d}_{\omega,\ell-r}\tilde{e}_{s-r} + \sum_{i=0}^{m-1}\sum_{\ell=r}^{\infty}\Delta_{i,\ell-r}e_{i,s-\ell} + \bar{\varepsilon}_{w,s-r}. \end{aligned} \tag{60}$$

Augmenting (31) by p lags of $\bar{x}_{ws}$, and using (60), we obtain the pooled version of the ARDL estimating equation,

$$\bar{x}_{ws} = \sum_{r=1}^{p}\psi_{\omega pr}\bar{x}_{w,s-r} + \sum_{r=0}^{h}\beta_{\omega pr}\tilde{e}_{s-r} + v_{\omega phs}, \tag{61}$$

where the error term $v_{\omega phs}$ is given by

$$v_{\omega phs} = \sum_{i=1}^{m-1}\sum_{r=0}^{\infty}\Delta_{\omega pir}e_{i,s-r} + \sum_{r=h+1}^{\infty}\beta_{\omega pr}\tilde{e}_{s-r} + \bar{\varepsilon}_{ws} - \sum_{r=1}^{p}\psi_{\omega pr}\bar{\varepsilon}_{w,s-r},$$

the coefficients $\beta_{\omega pr}$ are given by

$$\beta_{\omega p0} = \bar{d}_{\omega 0}, \tag{62}$$

$$\beta_{\omega p1} = \bar{d}_{\omega 1} - \psi_{\omega p1}\bar{d}_{\omega 0}, \tag{63}$$

$$\beta_{\omega p2} = \bar{d}_{\omega 2} - \psi_{\omega p1}\bar{d}_{\omega 1} - \psi_{\omega p2}\bar{d}_{\omega 0}, \tag{64}$$

$$\vdots$$

$$\beta_{\omega p\ell} = \bar{d}_{\omega\ell} - \sum_{r=1}^{p}\psi_{\omega pr}\bar{d}_{\omega,\ell-r}, \text{ for } \ell = p, p+1, ..., \tag{65}$$

and the coefficients $\Delta_{\omega pir}$ are similarly given by

$$\Delta_{\omega pi0} = \Delta_{i0}, \tag{66}$$

$$\Delta_{\omega pi1} = \Delta_{i1} - \psi_{\omega p1}\Delta_{i0}, \tag{67}$$

$$\Delta_{\omega pi2} = \Delta_{i2} - \psi_{\omega p1}\Delta_{i1} - \psi_{\omega p2}\Delta_{i0}, \tag{68}$$

$$\vdots$$

$$\Delta_{\omega pi\ell} = \Delta_{i\ell} - \sum_{r=1}^{p}\psi_{\omega pr}\Delta_{i,\ell-r}, \text{ for } \ell = p, p+1, ... \tag{69}$$

Following the same arguments as before, the IRF coefficients $\bar{\mathbf{d}}_{\omega} = \left(\bar{d}_{\omega 0}, \bar{d}_{\omega 1}, ..., \bar{d}_{\omega h}\right)'$ can be

recovered from (61). Let $\boldsymbol{\theta}_{\omega p} = \left(\boldsymbol{\psi}'_{\omega p}, \boldsymbol{\beta}'_{\omega p}\right)'$, where $\boldsymbol{\psi}_{\omega p} = \left(\psi_{\omega p1}, \psi_{\omega p2}, ..., \psi_{\omega pp}\right)'$, and $\boldsymbol{\beta}_{\omega p} = \left(\beta_{\omega p0}, \beta_{\omega p1}, ..., \beta_{\omega ph}\right)'$, and denote the mapping from $\boldsymbol{\theta}_{\omega p}$ to $\mathbf{d}_\omega$, implicitly defined by equations (62)-(65) as $F_\omega\left(\boldsymbol{\theta}_{\omega p}\right) = \bar{\mathbf{d}}_\omega$. Let $\hat{\boldsymbol{\theta}}_{\omega p} = \left(\hat{\boldsymbol{\psi}}'_{\omega p}, \hat{\boldsymbol{\beta}}'_{\omega p}\right)'$ be the LS estimator of $\boldsymbol{\theta}_{\omega p} = \left(\boldsymbol{\psi}'_{\omega p}, \boldsymbol{\beta}'_{\omega p}\right)'$ using (61). The asymptotic distribution of $\widehat{\bar{\mathbf{d}}}^{PA}_{\omega p} = F_\omega\left(\hat{\boldsymbol{\theta}}_{\omega p}\right)$ is given by the following proposition.

Proposition 4 *Suppose $\mathbf{z}_t = (x_t, \mathbf{y}'_t)'$ is given by (1), and Assumptions 2-3 hold. Consider the pooled ARDL estimator $\widehat{\bar{\mathbf{d}}}^{PA}_{\omega p} = F_\omega\left(\hat{\boldsymbol{\theta}}_{\omega p}\right)$, where $\hat{\boldsymbol{\theta}}_{\omega p} = \left(\hat{\boldsymbol{\psi}}'_{\omega p}, \hat{\boldsymbol{\beta}}'_{\omega p}\right)'$ is the LS estimator of $\boldsymbol{\theta}_{\omega p} = \left(\boldsymbol{\psi}'_{\omega p}, \boldsymbol{\beta}'_{\omega p}\right)'$ in (61), and the mapping $F_\omega\left(\boldsymbol{\theta}_{\omega p}\right) = \mathbf{d}_\omega$ is implicitly defined by (62)-(65). Let $T_m = T/m - h$. Then, for fixed m, p, and h, as $T \to \infty$,*

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}^{PA}_{\omega p} - \bar{\mathbf{d}}_\omega\right) \to_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{PA,p}\right), \tag{70}$$

where

$$\boldsymbol{\Omega}_{PA,p} = \mathbf{J}_{F\omega}\left(\boldsymbol{\theta}_{\omega p}\right) \boldsymbol{\Sigma}_{\omega\theta p} \mathbf{J}'_{F\omega}\left(\boldsymbol{\theta}_{\omega p}\right), \tag{71}$$

$\mathbf{J}_{F\omega}(.)$ is the Jacobian of F_ω and $\boldsymbol{\Sigma}_{\theta\omega p}$ is the asymptotic variance of $\hat{\boldsymbol{\theta}}_{\omega p}$.

$\boldsymbol{\Omega}_{PA,p}$ can be consistently estimated as

$$\boldsymbol{\Omega}_{PA,p} = \mathbf{J}_{F\omega}\left(\hat{\boldsymbol{\theta}}_{\omega p}\right) \hat{\boldsymbol{\Sigma}}_{\theta\omega p} \mathbf{J}'_{F\omega}\left(\hat{\boldsymbol{\theta}}_{\omega p}\right), \tag{72}$$

where $\hat{\boldsymbol{\Sigma}}_{\theta\omega p}$ is a consistent estimator of $\boldsymbol{\Sigma}_{\theta\omega p}$, such as the Newey-West estimator.

Remark 7 *Estimators of impulse-response coefficients $\bar{\mathbf{d}}_\omega$ based on the ARDL representations will be asymptotically more efficient compared with the DL estimating equations, so long as at least one of the AR terms, $\psi_{pr}/\psi_{\omega pr}$, is nonzero. As in the case of the DL mean group and pooled estimators, the mean group ARDL estimator will be at least as efficient asymptotically as the pooled version. However, in the finite samples of practical interest, the pooled ARDL estimator could outperform the mean group version because of its greater parsimony.*

3.3 Mean Group and Pooled Estimators based on Vector Autoregressive Distributed Lag (VARDL) estimating equations

It is also possible to augment ARDL estimating equations with additional variables in $\mathbf{z}_t$. However, there can be a significant cost in finite samples from estimating higher-dimensional models. Unlike

in standard VAR models, in our case omitting a possibly relevant variable does not result in the loss of consistency of the resulting impulse response estimator. It only involves the loss of asymptotic efficiency. The benefit of working with more parsimonious models is the estimator may be more accurate for sample sizes of practical relevance, as we will illustrate in Section 4.

Consider replacing $\bar{x}_{ws}$ in (61) with a $k \times 1$ vector $\bar{\mathbf{z}}_{wks} = (\bar{x}_{w,s}, \bar{y}_{w1,s}, \bar{y}_{w2,s}..., \bar{y}_{w,k-1,s})'$, containing a subset of variables in $\bar{\mathbf{z}}_{ws}$, for some $k > 1$. Using the high-frequency VAR model given by (1), we obtain a moving average representation for $\mathbf{z}_{kt} = \mathbf{S}_k'\mathbf{z}_t = (x_t, y_{1t}, y_{2t}, ..., y_{k-1,t})'$, where $\mathbf{S}_k'$ is an $n \times k$ selection matrix for the first k variables in $\mathbf{z}_t$, and

$$\mathbf{z}_{kt} = \mathbf{S}_k'\mathbf{z}_t = \mathbf{S}_k' \sum_{\ell=0}^{\infty} \mathbf{\Phi}^\ell \mathbf{a} e_{t-\ell} + \mathbf{S}_k' \sum_{\ell=0}^{\infty} \mathbf{\Phi}^\ell \boldsymbol{\eta}_{t-\ell} = \sum_{h=0}^{\infty} \mathbf{b}_{k\ell} e_{t-\ell} + \boldsymbol{\varepsilon}_{kt},$$

where $\mathbf{b}_{k\ell} = \mathbf{S}_k' \sum_{\ell=0}^{\infty} \mathbf{\Phi}^\ell \mathbf{a}$, and $\boldsymbol{\varepsilon}_{kt} = \mathbf{S}_k' \sum_{\ell=0}^{\infty} \mathbf{\Phi}^\ell \boldsymbol{\eta}_{t-\ell}$. Temporally aggregating $\mathbf{z}_t$ to the low frequency, we obtain the corresponding representation for the observed variables,

$$\bar{\mathbf{z}}_{wks} = \sum_{i=0}^{m-1} \sum_{r=0}^{\infty} \mathbf{d}_{kir} e_{i,s-r} + \bar{\boldsymbol{\varepsilon}}_{wks} = \sum_{i=0}^{m-1} \mathbf{d}_{kir}(L) e_{is} + \bar{\boldsymbol{\varepsilon}}_{wks}, \tag{73}$$

which is an extension of (16) to the multivariate case. Thus, the extension of the ARDL estimating equations to the multivariate case is straightforward. It follows the same steps as in the case of ARDL approach.

We continue to focus on estimating the low-frequency IRF coefficients $\bar{\mathbf{d}}_\omega = \left(\bar{d}_{\omega 0}, \bar{d}_{\omega 1}, ..., \bar{d}_{\omega r}\right)'$ for the target variable of interest, x. In terms of coefficients of (74), $\bar{d}_{\omega r}$ is given by

$$\bar{d}_{\omega r} = \sum_{i=0}^{m-1} \omega_i \mathbf{s}_{k,1}' \mathbf{d}_{kir}, \text{ for } r = 0, 1, 2, ...h, \tag{74}$$

where $\mathbf{s}_{k,1}$ is a $k \times 1$ selection vector for the first element. (74) can be equivalently written as

$$\bar{\mathbf{d}}_\omega = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right) \left(\mathbf{I}_{h+1} \otimes \mathbf{s}_{k,1}'\right) \mathbf{d}_{(k)}, \tag{75}$$

where $\mathbf{d}_{(k)} = (\mathbf{d}_{k\circ 0}', \mathbf{d}_{k\circ 1}', ..., \mathbf{d}_{k\circ h}')'$ and $\mathbf{d}_{k\circ r} = \left(\mathbf{d}_{k,0,r}', \mathbf{d}_{k,1,r}', ..., \mathbf{d}_{k,m-1,r}'\right)'$, for $r = 0, 1, ..., h$. The vector $\mathbf{d}_{(k)}$ can be estimated using multivariate versions of the DL or ARDL estimating equations. We focus on the latter. Replacing $\bar{x}_{ws}$ and its lags in (43) with $\bar{\mathbf{z}}_{wks}$ and its lags, we obtain the

VARDL estimating equations

$$\bar{\mathbf{z}}_{wks} = \sum_{r=1}^{p} \boldsymbol{\Psi}_{kpr} \bar{\mathbf{z}}_{w,k,s-r} + \sum_{i=0}^{m-1} \sum_{r=0}^{h} \boldsymbol{\beta}_{kpir} e_{i,s-r} + \mathbf{v}_{kphs}, \tag{76}$$

where, regardless of the population values of the autoregressive coefficient matrices $\{\boldsymbol{\Psi}_{kp1}, \boldsymbol{\Psi}_{kp2}, ..., \boldsymbol{\Psi}_{kpp}\}$ and regardless of the choice of p, we have

$$\begin{aligned}
\mathbf{d}_{ki0} &= \boldsymbol{\beta}_{kpi0}, & (77)\\
\mathbf{d}_{ki1} &= \boldsymbol{\beta}_{kpi1} + \boldsymbol{\Psi}_{kp1}\mathbf{d}_{ki0}, & (78)\\
\mathbf{d}_{ki2} &= \boldsymbol{\beta}_{kpi2} + \boldsymbol{\Psi}_{kp1}d_{i1} + \boldsymbol{\Psi}_{kp2}\mathbf{d}_{ki0}, & (79)\\
&\vdots & \\
\mathbf{d}_{ki\ell} &= \boldsymbol{\beta}_{kpi\ell} + \sum_{r=1}^{p} \boldsymbol{\Psi}_{kpr}\mathbf{d}_{k,i,\ell-r}, \text{ for } \ell = p, p+1, ...h. & (80)
\end{aligned}$$

Collect the coefficients of (76) in $\boldsymbol{\theta}_{kp} = vec\left(\boldsymbol{\Psi}_{kp1}, ..., \boldsymbol{\Psi}_{kpp}, \boldsymbol{\beta}_{k,p,0,0}, ..., \boldsymbol{\beta}_{k,p,m-1,h}\right)$ and denote the corresponding LS estimates by hats. In addition, define mapping $F_k\left(\boldsymbol{\theta}_{kp}\right) = \mathbf{d}_{(k)}$ given by (77)-(80). Then the mean group VARDL estimator of $\bar{d}_{\omega r}$ is given by

$$\widehat{\bar{\mathbf{d}}}_{\omega kp}^{MGV} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right)\left(\mathbf{I}_{h+1} \otimes \mathbf{s}_{k,1}'\right)\widehat{\mathbf{d}}_{(k),p}^{MGV}, \tag{81}$$

where $\widehat{\mathbf{d}}_{(k),p}^{MGV} = F_k\left(\hat{\boldsymbol{\theta}}_{kp}\right)$. Similarly, the pooled VARDL estimator is based on a vector version of (61),

$$\bar{\mathbf{z}}_{wks} = \sum_{r=1}^{p} \boldsymbol{\Psi}_{\omega kpr} \bar{\mathbf{z}}_{w,k,s-r} + \sum_{r=0}^{h} \boldsymbol{\beta}_{\omega kpr} \tilde{e}_{s-r} + \mathbf{v}_{\omega kphs}. \tag{82}$$

Let $\boldsymbol{\theta}_{\omega kp} = vec\left(\boldsymbol{\Psi}_{\omega kp1}, ..., \boldsymbol{\Psi}_{\omega kpp}, \boldsymbol{\beta}_{\omega kp0}, ..., \boldsymbol{\beta}_{\omega kpp}\right)$, and define mapping $F_{\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right) = \bar{\mathbf{d}}_{\omega k} =$

$\left(\bar{\mathbf{d}}'_{\omega k,0}, \bar{\mathbf{d}}'_{\omega k,1}, ..., \bar{\mathbf{d}}'_{\omega k,h}\right)'$ given by

$$\begin{aligned}
\bar{\mathbf{d}}_{\omega k0} &= \boldsymbol{\beta}_{\omega kp0}, && (83)\\
\bar{\mathbf{d}}_{\omega k1} &= \boldsymbol{\beta}_{\omega p1} + \boldsymbol{\Psi}_{\omega kp1}\bar{\mathbf{d}}_{\omega k0}, && (84)\\
\bar{\mathbf{d}}_{\omega k2} &= \boldsymbol{\beta}_{\omega p2} + \boldsymbol{\Psi}_{\omega kp1}\bar{\mathbf{d}}_{\omega k1} + \boldsymbol{\Psi}_{\omega kp2}\bar{\mathbf{d}}_{\omega k0}, && (85)\\
&\vdots &&\\
\bar{\mathbf{d}}_{\omega k\ell} &= \boldsymbol{\beta}_{\omega p\ell} + \sum_{r=1}^{p}\boldsymbol{\Psi}_{\omega kpr}\bar{\mathbf{d}}_{\omega,k,\ell-r}, \text{ for } \ell = p, p+1, ..., h, && (86)
\end{aligned}$$

Then the pooled VARDL estimator is given by

$$\widehat{\bar{\mathbf{d}}}^{PV}_{\omega kp} = \left(\mathbf{I}_{h+1}\otimes \mathbf{s}'_{k,1}\right)\widehat{\bar{\mathbf{d}}}_{\omega kp} = \left(\mathbf{I}_{h+1}\otimes \mathbf{s}'_{k,1}\right)F_{\omega k}\left(\hat{\boldsymbol{\theta}}_{\omega kp}\right). \quad (87)$$

The following proposition establishes the asymptotic normality of the mean group and pooled estimators based on VARDL estimating equations.

Proposition 5 *Suppose $\mathbf{z}_t = (x_t, \mathbf{y}'_t)'$ is given by (1), and Assumptions 2-3 hold. Consider the mean group VARDL estimator $\widehat{\bar{\mathbf{d}}}^{MGV}_{\omega kp}$ given by (81) and the pooled VARDL estimator $\widehat{\bar{\mathbf{d}}}^{PV}_{\omega kp}$ given by (87). Let $T_{mh} = T/m - h$. Then, for fixed k, m, p and h, as $T \to \infty$,*

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}^{MGV}_{\omega kp} - \bar{\mathbf{d}}_{\omega}\right) \to_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{MGV,kp}\right), \quad (88)$$

and

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}^{PV}_{\omega kp} - \bar{\mathbf{d}}_{\omega}\right) \to_d N\left(\mathbf{0}, \boldsymbol{\Omega}_{PV,kp}\right), \quad (89)$$

where

$$\boldsymbol{\Omega}_{MGV,kp} = \left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}'\right)\left(\mathbf{I}_{h+1}\otimes \mathbf{s}'_{k,1}\right)\mathbf{J}_{Fk}\left(\boldsymbol{\theta}_{kp}\right)\boldsymbol{\Sigma}_{\theta kp}\mathbf{J}'_{Fk}\left(\boldsymbol{\theta}_{kp}\right)\left(\mathbf{I}_{h+1}\otimes \mathbf{s}_{k,1}\right)\left(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}\right), \quad (90)$$

$$\boldsymbol{\Omega}_{PV,kp} = \mathbf{J}_{F\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right)\boldsymbol{\Sigma}_{\omega k\theta p}\mathbf{J}'_{F\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right), \quad (91)$$

$\mathbf{J}_{Fk}\left(\boldsymbol{\theta}_{kp}\right)$ is the Jacobian of the mapping $F_k\left(\boldsymbol{\theta}_{kp}\right) = \mathbf{d}_{(k)}$ defined by (77)-(80), $\mathbf{J}_{F\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right)$ is the Jacobian of the mapping $F_{\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right) = \bar{\mathbf{d}}_{\omega k}$ defined by (83)-(86), $\boldsymbol{\Sigma}_{\theta kp}$ is the asymptotic variance of $\hat{\boldsymbol{\theta}}_{kp}$, and $\boldsymbol{\Sigma}_{\omega k\theta p}$ is the asymptotic variance of $\hat{\boldsymbol{\theta}}_{\omega kp}$.

Asymptotic variance-covariance matrices $\mathbf{\Omega}_{MGV,kp}$ and $\mathbf{\Omega}_{PV,kp}$ can be estimated using

$$\hat{\mathbf{\Omega}}_{MGV,kp} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right)\left(\mathbf{I}_{h+1} \otimes \mathbf{s}'_{k,1}\right)\mathbf{J}_{Fk}\left(\hat{\boldsymbol{\theta}}_{kp}\right)\hat{\mathbf{\Sigma}}_{\theta kp}\mathbf{J}'_{Fk}\left(\hat{\boldsymbol{\theta}}_{kp}\right)\left(\mathbf{I}_{h+1} \otimes \mathbf{s}_{k,1}\right)\left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}\right), \quad (92)$$

and

$$\hat{\mathbf{\Omega}}_{PV,kp} = \mathbf{J}_{F\omega k}\left(\hat{\boldsymbol{\theta}}_{\omega kp}\right)\hat{\mathbf{\Sigma}}_{\omega k\theta p}\mathbf{J}'_{F\omega k}\left(\hat{\boldsymbol{\theta}}_{\omega kp}\right), \quad (93)$$

where $\hat{\mathbf{\Sigma}}_{\theta kp}$ and $\hat{\mathbf{\Sigma}}_{\omega k\theta p}$ are the Newey-West estimators of $\mathbf{\Sigma}_{\theta kp}$ and $\mathbf{\Sigma}_{\omega k\theta p}$, respectively.

4 Monte Carlo Experiments

Since the pooled estimators proposed in this paper are more parsimonious, they may have an advantage for the samples of interest in practice, despite being asymptotically less efficient compared to the mean group estimators. This section sheds light on the small sample accuracy of the six estimators discussed in Section 3. We first outline the data generating process (Subsection 4.1), then describe the implementation of the estimators (Subsection 4.2), and summarize the main findings (Subsection 4.3). The findings in this section confirm our intuition and the theoretical arguments provided in Section 3. In particular, we find that parsimony can pay off in realistic samples sizes. Pooled versions of estimators tend to be more accurate than their mean group versions in our experiments. Augmenting the estimating equations by lags of the dependent variable can be quite helpful, too.

4.1 Data Generating Process

Let $n = 3$, and generate $\mathbf{z}_t = (x_t, q_{1t}, q_{2t})'$ from the VAR(1) process:

$$\mathbf{z}_t = \left(\mathbf{I}_3 - \mathbf{\Phi}\right)\boldsymbol{\mu}_z + \mathbf{\Phi}\mathbf{z}_{t-1} + \mathbf{A}\boldsymbol{\varepsilon}_t,\ \boldsymbol{\varepsilon}_t IIDN(0, I) \quad (94)$$

for $t = -B+1, -B+2..., 0, 1, ..., T$, where $\boldsymbol{\mu}_z = (1, 1, 1)'$, and the starting values are $\boldsymbol{\xi}_{-B} = \boldsymbol{\mu}_z$. We set $B = 100$ and discard the first B time periods, leaving periods $1, 2, ..., T$ for estimation. Individual elements of $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})'$ are generated using GARCH$(1, 1)$ specifications given by

$\varepsilon_{st} = \sigma_{st}\epsilon_{st}$ for $s = 1, 2, 3$, where $\epsilon_{st} \sim IIDN(0, 1)$, and

$$\sigma_{st}^2 = (1 - \psi_{s1} - \psi_{s2}) + \psi_{s1}\varepsilon_{s,t-1}^2 + \psi_{s2}\sigma_{s,t-1}^2,$$

for $t = -B, -B+1, ..., 0, 1, 2, ..., T$, with initial values $\sigma_{s,-B-1}^2 = 1$. We set $\psi_{s1} = 0.2$, and $\psi_{s2} = 0.6$.

The matrix of the autoregressive coefficients, $\mathbf{\Phi}$, and the matrix $\mathbf{A}$ are set to

$$\mathbf{\Phi} = \begin{pmatrix} 0.8 & -0.1 & 0 \\ 0.4 & 0.6 & -0.2 \\ 0 & 0.2 & 0.4 \end{pmatrix}^{1/m}, \text{ and } \mathbf{A} = \varkappa \begin{pmatrix} 1 & -0.5 & 0 \\ 0 & 1 & -0.5 \\ -0.5 & 0.5 & 1 \end{pmatrix}. \tag{95}$$

The absolute values of the largest eigenvalue of $\mathbf{\Phi}$ is $0.63^{1/m}$. We assume that ε_{1t} is observed for $t = 1, 2, ..., T$. In addition, the variables in $\mathbf{z}_t$ are only observed as temporal aggregates, $\bar{\mathbf{z}}_{w,s} = \sum_{i=0}^{m-1} w_i \mathbf{z}_{t_s - i}$, for $s = 1, 2, ..., T_m$. We set $m = 21$, which corresponds to a setting in which there are monthly observations on the outcome variable of interest and daily data on the shock of interest (assuming 21 working days in a month). Furthermore, we consider aggregation weights $\mathbf{w} = (1, 1, ..., 1)'$, representing the sums of the high-frequency periods, as one would use to temporally aggregate industrial production or retail sales, for example.[3] While our estimators do not require the specification of the aggregation weights, these weights are required for the evaluation of the estimators in the simulation study.

We consider $T_m = \{240, 300, 360, 480, 600\}$, corresponding to 20, 25, 30, 40 and 50 years of monthly data. We utilize $R = 2,000$ Monte Carlo draws for each sample size. Our objective is the estimation of $IRF_\omega(\bar{x}_w, r) = d_{\omega,r}$. We set $\boldsymbol{\omega} = \boldsymbol{\tau}_m/m = (1/m, 1/m, ..., 1/m)'$, where τ_m is $m \times 1$ vector of ones. We set $\varkappa = 10.3356$ in (95) to ensure $d_{\omega_a,0} = 1$.

4.2 Estimators

We use the six estimators of the $d_{\omega,r}$ proposed in Section 3: the mean group DL estimator given by (22); the pooled DL estimator given by (35); the mean group ARDL estimator given by (55); the pooled ARDL estimator based on (61); the mean group VARDL estimator given by (81) with

[3] Note that the relative small sample accuracy of individual estimators would be the same if we considered simple temporal averages given by $\mathbf{w} = (1/m, 1/m, ..., 1/m)'$, since temporal averaging and summation differ only with respect to a scaling factor.

variable vector $\bar{\mathbf{z}}_{w2s} = (\bar{x}_{ws}, \bar{q}_{1,ws})'$; and the pooled VARDL estimator given by (87) using the same variable vector. The lag orders (p, q) are set equal to the integer part of $T_m^{1/3}$. The largest impulse response horizon is set to $h = 12$.

4.3 Findings

Table 1 reports results for the bias of $\hat{d}_{\omega,r}$, the RMSE, the coverage rates of nominal 95 percent confidence intervals for $d_{\omega,r}$, and the mean length of these intervals (all $\times 100$), averaged across horizons $r = 0, 1, 2..., 12$. The left panel reports the mean group estimators and the right panel reports their pooled versions. The mean group DL estimator reported in the first column is not feasible for the smallest value of $T_m = 240$ (and the chosen horizon $h = 12$), due to the lack of degrees of freedom. This estimator does not exhibit much bias for the remaining values of T_m, and its RMSE declines in T_m, as expected according to Proposition 1. However, the coverage rates are very low, starting at only 22 to 24 percent for $T_m = 300$, and only slowly improving with an increase in T_m, reaching 66 to 83 percent coverage for $T_m = 600$, which is still significantly below 95 percent. Inference is therefore highly inaccurate and the 95 percent confidence intervals are severely underestimating the true extent of the sampling uncertainty.

The other estimators do not have this problem. Confidence interval coverage rates are quite close to 95 percent for all sample sizes. None of these estimators suffers from serious bias, and all achieve a lower RMSE than the mean group DL estimator. All three pooled estimators have lower RMSE than the corresponding mean group estimators. The best estimator in terms of the RMSE is the pooled ARDL estimator. Augmenting the ARDL specification with one additional variable $(\bar{q}_{1,ws})$ results in a slight increase of the RMSE by about 1 to 5 percent, depending on the sample size. This result underscores that the inclusion of a relevant variable does not necessarily make the estimator more precise. Overall, our Monte Carlo study shows that the proposed estimators, especially the more parsimonious ones, perform well, with the exception of the mean group DL estimator.

TABLE 1: Bias (×100), RMSE (×100), coverage rates (×100), and mean length of 95% confidence intervals (×100) for $\widehat{\bar{d}}_{\omega,r}$, averaged over horizons $r = 0$ to 12.

	Mean Group MF estimators			Pooled MF estimators		
	DL	ARDL	VARDL	DL	ARDL	VARDL
T_m	A. Bias (×100)					
240	n.a.	-0.46	-0.50	0.07	-0.33	-0.31
300	0.25	-0.18	-0.25	-0.05	-0.31	-0.30
360	-0.08	-0.14	-0.19	-0.09	-0.29	-0.32
480	-0.27	-0.33	-0.29	-0.06	-0.27	-0.25
600	-0.05	-0.26	-0.25	-0.05	-0.24	-0.22
	B. RMSE (×100)					
240	n.a.	17.01	18.69	13.24	10.44	10.95
300	49.70	12.66	13.33	11.58	9.16	9.62
360	21.18	11.57	12.03	10.50	8.86	9.13
480	13.08	8.97	9.22	8.99	7.58	7.73
600	10.21	7.96	8.11	7.96	6.93	6.99
	C. 95% Confidence Interval Coverage Rates (×100)					
240	n.a.	94.01	94.38	92.22	94.04	94.04
300	23.86	94.59	94.45	93.07	94.65	94.54
360	57.72	94.12	94.53	93.40	94.15	94.12
480	76.83	94.57	94.89	93.75	94.19	94.53
600	83.38	94.62	94.80	93.96	94.61	94.61
	D. 95% Confidence Intervals Length (×100)					
240	n.a.	63.91	71.14	47.19	39.23	41.60
300	28.31	48.15	51.64	42.39	34.94	36.87
360	34.15	44.43	46.71	38.72	33.49	34.61
480	31.16	34.35	35.59	33.64	28.85	29.72
600	28.26	30.72	31.35	30.05	26.71	27.06

Notes: The mean group DL estimator is not feasible for $T_m = 240$, $m = 21$, and $h = 12$. See Subsection 4.1 for the description of the simulation design and Subsection 4.2 for the description of the implementation of the individual estimators. All results are based on $R = 2000$ simulation draws.

5 Empirical Application: Passthrough of Daily Gasoline Price Shocks to Monthly Inflation

Our empirical illustration is motivated by the question of how occasional spikes or surges in the daily New York Harbor wholesale price of gasoline driven by exogenous events in the Middle East impact headline CPI inflation, core CPI inflation (defined as CPI inflation excluding energy and food prices), and gasoline price inflation. This is a recurrent situation faced by policymakers. From a policymaker's point of view, it is important to understand in real time how a sequence of wholesale gasoline prices triggered by a crisis in the Middle East, at some point during the current month is expected to change inflation in the current month and in futures months. The reason for focusing on daily gasoline as opposed to daily oil prices is that changes in the price of gasoline and the price of crude oil are not proportionate in general, reflecting the evolution of the cost share of crude oil in gasoline prices over time (see Kilian and Zhou, 2022a).

We focus on the daily wholesale price of gasoline because that price is a benchmark price for U.S. retail gasoline markets, is not subject to revisions and is readily available in real time. We focus on sequences of daily gasoline price shocks, as even temporary spikes in gasoline prices tend to be longer lived in practice than a one-time daily shock.

Our maintained assumption is that the daily gasoline price is predetermined with respect to daily and monthly inflation. Evidence in support of this assumption is provided in Kilian and Vega (2011). Given that this price is well approximated by a random walk, its growth rate may be viewed as the daily gasoline price shock. Our assessment of the causal impact of a given sequence of wholesale gasoline price shocks on inflation can be made in real time and does not require inflation data to be available for the current month. In practice, it may be used to adjust existing inflation forecasts that predate the surge in gasoline prices.

5.1 Some Hypothetical Thought Experiments

We start by building intuition by considering three hypothetical gasoline price shock scenarios that differ in the timing of the daily gasoline price shocks within the month (see Figure 1A). The implied change in the price of gasoline is shown in Figure 1B. In Scenario 1, 10 consecutive shocks of magnitude 1 are followed by zero shocks in the remainder of the month, resulting in a steadily

increasing gasoline price during the first ten days of the month and a flat price of 10 for the remainder of the month. Scenario 2 makes the opposite assumption of no increase in the gasoline price during the first ten days, followed by a steady increase in the remainder of the month to a value of 10. Scenario 3 postulates a steady increase over all 21 trading days, reaching the same gasoline price of 10 at the end of the month as in the other two scenarios.

The top panel in Figure 2 illustrates the implications of each scenario for the annualized inflation rates expressed in percent for the pooled ARDL estimator. Regardless of the choice of scenario, the pattern of the response of headline inflation mirrors that of consumer gasoline price inflation except that the scale is much lower, reflecting the expenditure share of gasoline in overall consumer spending of about 4% on average. The responses of core inflation are generally an order of magnitude smaller, indicating very limited passthrough to inflation in other consumer prices.

Not surprisingly, under Scenario 1 gasoline price shocks have the highest impact effect on all three inflation measures, and Scenario 3 has the lowest impact effect. Scenario 2 generates the largest peak effect on all inflation measures by a slight margin. Whereas the peak effect on headline inflation occurs in month 1, for core inflation it occurs in month 2. Under Scenarios 1 and 2, the timing of the peak effect is similar, except the peak effect on core inflation under Scenario 2 is delayed until month 3.

The peak response of headline inflation is noticeably lower under Scenario 3 than under the other scenarios and more persistent. The most striking differences across scenarios arise for the response of core inflation. Notably, the pattern of the core inflation responses under Scenario 2 is quite different from that under the other scenarios. The peak effect under Scenario 3 is lower than in the other two scenarios, and the persistence of the response function is higher, reflecting the fact that the inflationary shock occurred only late in the initial month.

The three lower panels in Figure 2 show the pooled ARDL point estimate for each scenario along with 68% and 95% delta method confidence intervals constructed using the asymptotic variance expression derived in Section 3. The extent of the sampling uncertainty is noticeably lower under Scenario 3 than under Scenario 1 and somewhat higher under Scenario 2. This evidence illustrates that the timing of the daily gasoline price shocks matters.

5.2 Four Episodes of Daily Gasoline Price Shocks

Figure 3 shows sequence of daily gasoline price shocks associated with four episodes of stress in U.S. gasoline markets driven by global events. It also shows the cumulative change in daily gasoline prices associated with these shocks and the evolution of the New York Harbor wholesale price of gasoline over this month.

We focus on the invasion of Kuwait by Iraq (August 1-31, 1990), the Iranian missile attack on Saudi oil fields (September 16-30, 2019), the drop in demand associated with the pandemic and increase in supply associated with the Saudi Price War (March 1-31, 2020), and the 12-Day War between Israel and Iran and its aftermath (June 11-30, 2025).[4]

Three of these episodes involved a rise in U.S. gasoline prices and one a decline. As Figure 3 shows, in the case of the Iranian missile attack, the gasoline price increase was almost completely reversed by the end of the month. After the invasion of Kuwait and the 12-Day War, we see a partial reversal of the initial increase. In the case of the pandemic and Saudi Price War, the price stabilized by the end of the month well below its peak level. This illustrates rich heterogeneity in the patterns of daily gasoline price shocks in practice.

Figure 4 reports the responses of monthly U.S. consumer gasoline price inflation, core inflation and headline inflation along with 68% and 95% confidence intervals. There are some striking differences. For example, the Iranian missile attack on Saudi Arabia in 2019 had a much more muted impact on gasoline and headline inflation than the invasion of Kuwait in 1990. There is no evidence of these responses being statistically significant either at the 95% level.

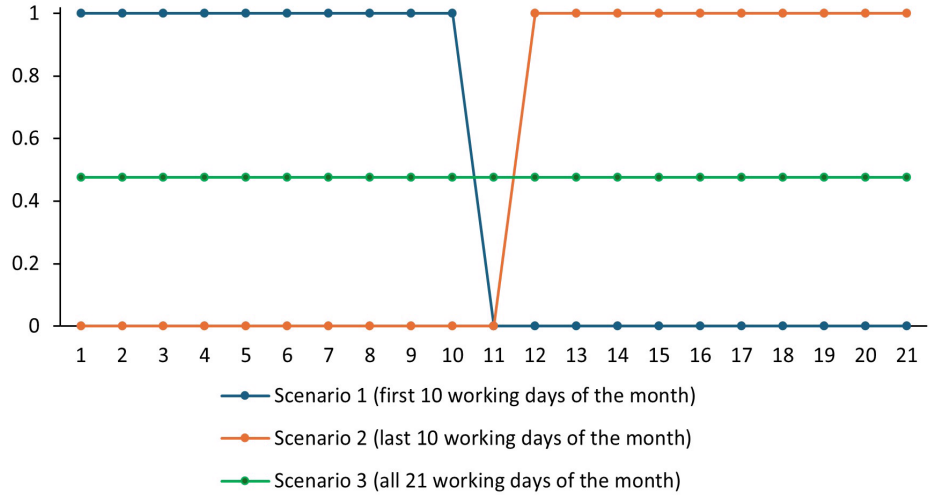
The only event that generated a large and statistically significant response at the 95% level in both headline and core inflation occurred in March 2020. The 12-Day War produced a statistically significant increase in gasoline price and headline inflation in months 0 and 1, but only a barely statistically significant increase in core inflation in month 0 at the 95% confidence level. It is also noteworthy that the responses of core inflation to the invasion of Kuwait and the missile attack on Saudi Arabia took longer to materialize than following the 2020 and 2025 episodes, once again underscoring the importance of the timing of the daily gasoline price shocks.

These examples illustrate that the estimators proposed in Section 3 may be used to assess the causal effect of sequences of gasoline price shocks on monthly inflation in real time based on

[4] The attack on Iran started on June 13 2025. We include the two days leading up to the attack since the wholesale gasoline price started to increase on June 11 and 12 already.

estimates of regressions on historical data, providing an additional tool for policymakers and applied researchers.

A. Daily shock sequences



B. Cumulated values of shocks

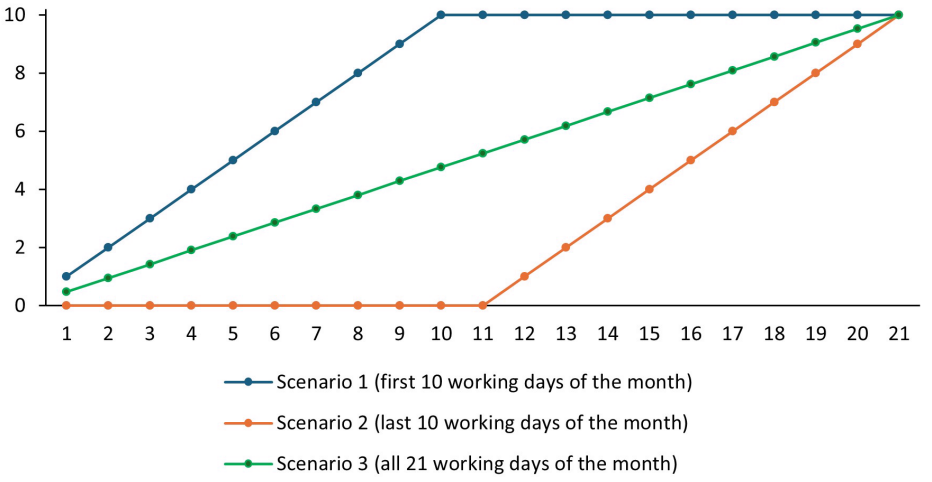


FIGURE 1: Daily shock sequence scenarios for log-differenced NY Harbor wholesale gasoline prices (left chart) and their cumulated values (right chart). Y-axis shows approximate percent units (log-differences $\times 100$). In each scenario, NY Harbor wholesale gasoline prices increase by 10 percent. X-axis shows working days of the month.

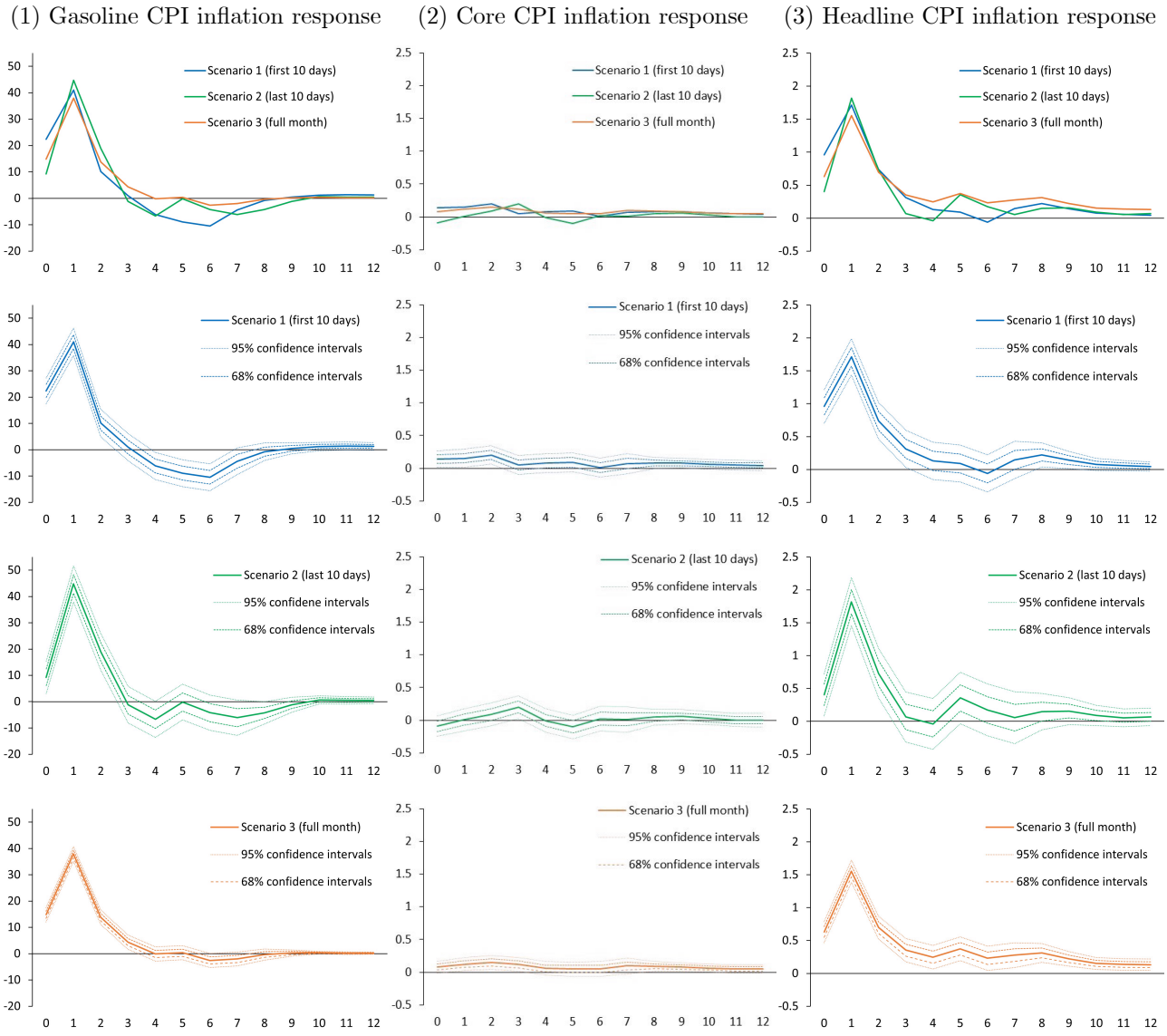


FIGURE 2: Pass-through from daily shocks to NY Harbor wholesale gasoline prices to monthly gasoline, core and headline CPI inflation under scenarios 1-3. Under each scenario, wholesale gasoline prices increase by 10 percent. Y-axis shows log-differences $\times 1,200$. X-axis shows monthly horizons. Dashed and dotted lines show 68 and 95 percent asymptotic confidence intervals.

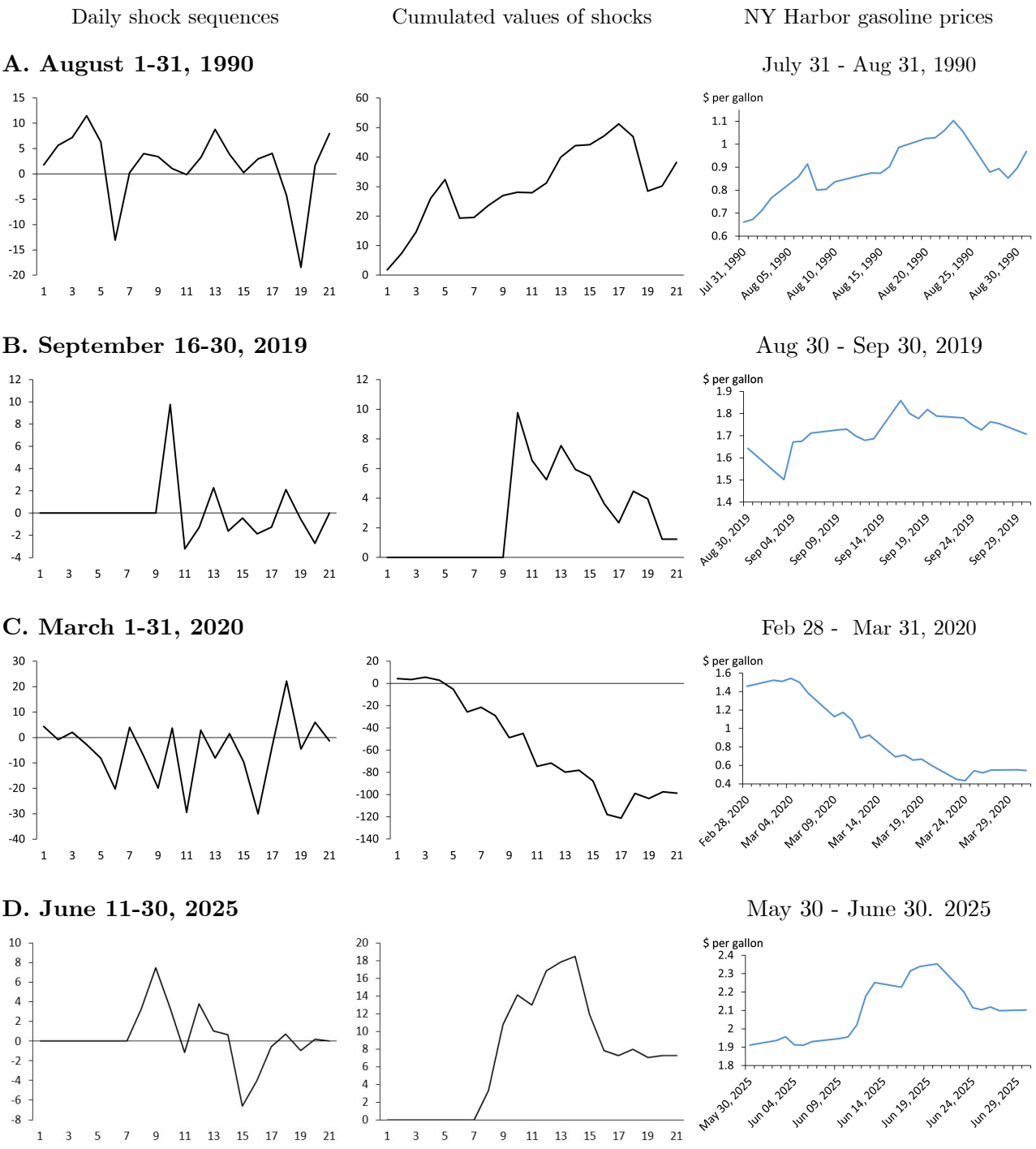
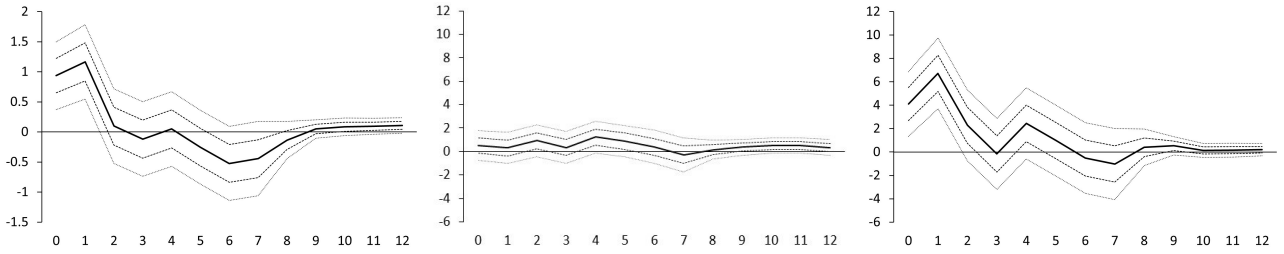


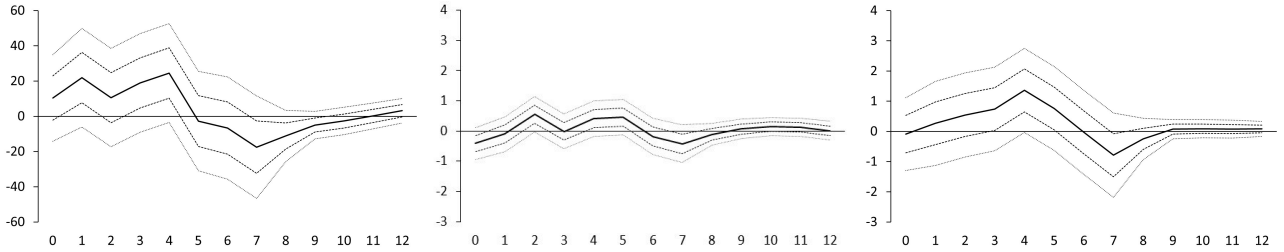
FIGURE 3: August 1990 (Iraqi invasion of Kuwait), September 2019 (Iranian missile attacks on Saudi Arabia), March 2020 (Covid and Saudi price war), and June 2025 (12-Day War between Iran and Israel) shock scenarios. Y-axis shows approximate percent units (log-differences $\times 100$). X-axis shows working days of the month. Daily shock sequences are given by first differences of logs of NY Harbor gasoline prices.

(1) Gasoline CPI inflation responses (2) Core CPI inflation responses (3) Headline CPI inflation responses

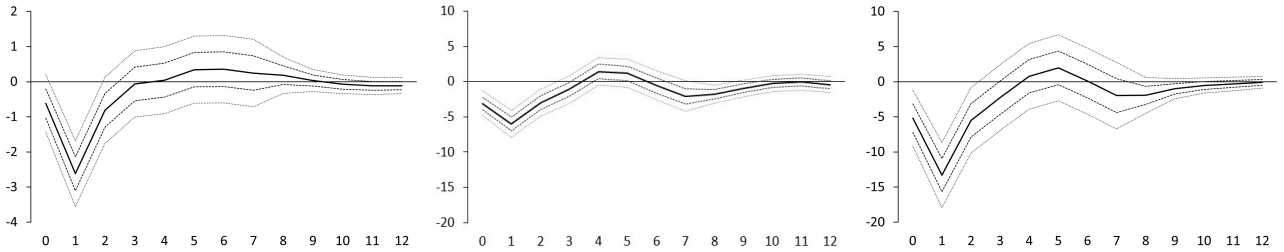
A. August 1990 (Invasion of Kuwait)



B. September 2019 (Iranian missile attack on Saudi Arabia)



C. March 2020 (Covid and Saudi price war)



D. June 2025 (12 Day War between Israel and Iran)

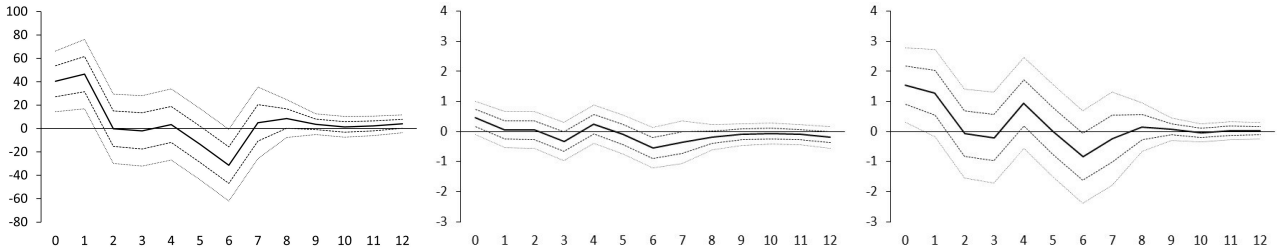


FIGURE 4: Pass-through from daily shocks to NY Harbor wholesale gasoline prices to monthly gasoline, core and headline CPI inflation for selected episodes. Y-axis shows log-differences $\times 1,200$. X-axis shows monthly horizons. Dashed and dotted lines show 68 and 95 percent asymptotic confidence intervals.

6 Conclusion

We provided several novel solutions to the problem of estimating the responses of low-frequency macroeconomic aggregates to directly observed high-frequency shocks. We also derived the asymptotic distribution of these estimators. Our simulation analysis shows that pooled ARDL and pooled VARDL estimators are most reliable in realistically small samples. While we focused on the relationship between daily shocks and monthly outcomes, our analysis could be applied to other frequencies.

Our analysis allows for stock as well as flow variables. One key advantage of our approach compared to mixed-frequency VAR models is that it does not require the user to fully specify the VAR model generating the data. Another advantage is that reliable estimation remains feasible even for realistically small sample sizes. A third advantage is that the proposed estimators accommodate one-time daily shocks as well as sequences of daily shocks. A limitation of our analysis is that we assume that the high-frequency shock of interest is observable.

Going forward, extensions of our work to sparse daily shock data would be of particular interest. There has been a surge in the use of proxy VAR models of the effect of exogenous shocks to price expectations. In this literature, the effects of policy announcements on price expectations are estimated using high-frequency surprises in asset prices around policy announcements as an external instrument. Such news shocks are infrequent. They typically are observed at most once or twice in a given month. As shown in Kilian (2024), utilizing such daily shocks in monthly proxy VAR models involves a nontrivial temporal aggregation problem (see also Lee and Sekhposyan (2024)). This problem has typically been ignored in the literature (see, e.g., Känzig, 2021). Very similar problems also arise in the literature on identifying high-frequency monetary policy shocks, typically defined as the change in the futures price of over a 30-minute window around FOMC announcments (see, e.g. Gertler and Karadi, 2015, Bauer and Swanson, 2023). This makes it of interest to develop generalizations of the approach in our paper that can deal with the sparseness of the time series of such news shocks, which is not allowed for in our current analysis. We delegate this task to future research.

References

Bauer, M. D. and E. T. Swanson (2023). A Reassessment of Monetary Policy Surprises and High-Frequency Identification. *NBER Macroeconomics Annual 37*, 87–155.

Breitung, J. and N. R. Swanson (2002). Temporal aggregation and spurious instantaneous causality in multiple time series models. *Journal of Time Series Analysis 23*(6), 651–665.

Chudik, A. and G. Georgiadis (2022). Estimation of Impulse Response Functions When Shocks Are Observed at a Higher Frequency Than Outcome Variables. *Journal of Business and Economic Statistics 40*, 965–979.

Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press.

Diab, S. and M. B. Karaki (2023). Do increases in gasoline prices cause higher food prices? *Energy Economics 127, 107066*.

Foroni, C., E. Ghysels, and M. Marcellino (2013). Mixed-Frequency Vector Autoregressive Models. *Advances Econometrics 32*, 247–272.

Foroni, C. and M. Marcellino (2014). Mixed-Frequency Structural Models: Identification, Estimation, and Policy Analysis. *Journal of Applied Econometrics 29*(7), 1118–1144.

Foroni, C. and M. Marcellino (2016). Mixed Frequency Structural Vector Auto-Regressive Models. *Journal of the Royal Statistical Society A 179*(2), 403–425.

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

Ghysels, E. (2016). Macroeconomics and the Reality of Mixed Frequency Data. *Journal of Econometrics 193*(2), 294–314.

Gonçalves, S. and L. Kilian (2007). Asymptotic and Bootstrap Inference for AR(∞) Processes with Conditional Heteroskedasticity. *Econometric Reviews 26*(6), 609–641.

Gründler, D. and J. Scharler (2025). Does uncertainty amplify the inflation pass-through of gasoline price shocks? *Energy Economics 144, 108348*.

Karaki, M. B. and J. E. Chaar (2026). Fuel price shocks and inflation dynamics in Lebanon. *Energy Economics 153, 109029*.

Kilian, L. (2024). How to Construct Monthly VAR Proxies Based on Daily Surprises in Futures Markets. *Journal of Economic Dynamics and Control 168, 104966*.

Kilian, L. and H. Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge Books. Cambridge University Press.

Kilian, L. and X. Vega (2011). Do Energy Prices Respond to U.S. Macroeconomic News? A Test of the Hypothesis of Predetermined Energy Prices. *Review of Economics and Statistics 93*(2), 660–671.

Kilian, L. and X. Zhou (2022a). Oil Prices, Gasoline Prices and Inflation Expectations. *Journal of Applied Econometrics 37*(5), 867–881.

Kilian, L. and X. Zhou (2022b). The Impact of Rising Oil Prices on US. Inflation and Inflation Expectations in 2020-23. *Energy Economics 113, 106228*.

Kilian, L. and X. Zhou (2023). A Broader Perspective on the Inflationary Effects of Energy Price Shocks. *Energy Economics 125, 106893*.

Kilian, L. and X. Zhou (2025). Oil Price Shocks and Inflation. In G. Ascari and R. Trezzi (Eds.), *Handbook of Inflation*, pp. 420–438. Cheltenham, UK: Edward Elgar.

Känzig, D. (2021). The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements. *American Economic Review 111*, 1092–1125.

Lee, Y. J. and T. Sekhposyan (2024). The Relevance of Temporal Aggregation for the Propagation of Macroeconomic Shocks. Working Paper, Texas A&M University.

Lütkepohl, H. (1987). *Forecasting aggregated vector ARMA processes*. Springer, Berlin, Heidelberg.

Marcellino, M. (1999). Some Consequences of Temporal Aggregation in Empirical Analysis. *Journal of Business and Economic Statistics 17*, 129–136.

Rossana, R. J. and J. J. Seater (1995). Temporal Aggregation and Economic Time Series. *Journal of Business and Economic Statistics 13*(4), 441–451.

Said, E. and D. A. Dickey (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika 71*, 599–607.

Swanson, N. R. and C. W. J. Granger (2012). Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions. *Journal of the American Statistical Association 92*(437), 357–367.

Vatsa, P. and G. Pino (2024). Do Petrol Prices Affect Inflation and Inflation Expectations? Evidence from New Zealand. *Energy Economics 139, 107939*.

White, H. (1984). *Asymptotic Theory for Econometricians*. Academic Press, Inc, Orlando, Florida.

Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*. Uppsala: Almquist and Wiksel.

A Appendix

This appendix contains proofs and the data sources for the empirical applications. Section A.1 states and establishes Lemmas needed for the proofs of the propositions presented in Section A.2. Section A.3 describes data sources and variable construction for the empirical application.

A.1 Lemmas: Statements and proofs

Lemma 1 *Consider matrix* $\mathbf{E}$ *defined below (21). Under Assumption 3, for fixed* m *and* h*, as* $T \to \infty$*,*

$$\left(\frac{\mathbf{E}'\mathbf{E}}{T_{mh}}\right)^{-1} \to_p \sigma_e^{-2}\mathbf{I}_{(h+1)m}. \tag{A.1}$$

Proof. $\mathbf{E}'\mathbf{E}/T_{mh}$ is an $m(h+1)\times m(h+1)$ matrix composed of elements $T_m^{-1}\sum_{s=h+1}^{T_m} a_s(i_1,i_2,r_1,r_2)$, for $i_1,i_2 = 0,1,2,...m-1$, and $r_1,r_2 = 0,1,2,...,h$, where $a_s(i_1,i_2,r_1,r_2) = e_{i_1,s-r_1}e_{i_2,s-r_2}$. For given integers i_1,i_2,r_1,r_2, the sequence $\{a_s(i_1,i_2,r_1,r_2)\}_s$ is an uncorrelated sequence under Assumption 3, with bounded variances, $Var[a_s(i_1,i_2,r_1,r_2)] < K$. Hence, $T_m^{-1}\sum_{s=h+1}^{T_m} a_s(i_1,i_2,r_1,r_2) \to_{L_2} E(e_{i_1,s-r_1}e_{i_2,s-r_2})$,[5] where $E(e_{i_1,s-r_1}e_{i_2,s-r_2}) = \sigma_e^2$ for the diagonal elements ($i_1 = i_2$ and $r_1 = r_2$), and $E(e_{i_1,s-r_1}e_{i_2,s-r_2}) = 0$ for the off-diagonal elements ($i_1 \neq i_2$ and/or $r_1 \neq r_2$). Convergence in L_2 norm implies convergence in probability, and we obtain $\mathbf{E}'\mathbf{E}/T_{mh} \to_{pp} \sigma_e^2\mathbf{I}_{(h+1)m}$. Given that $\sigma_e^2 > 0$ under Assumption 3, result (A.1) follows. ■

[5] See, e.g. Theorem 19.1 of Davidson (1994)

Lemma 2 *Consider matrix* $\mathbf{E}$ *defined below (21) and vector* $\boldsymbol{\vartheta}_h = (\vartheta_{h,h+1}, \vartheta_{h,h+2}, ..., \vartheta_{h,T_m})'$*, where* ϑ_{hs} *is defined by (19). Under Assumptions 2-3, for fixed* m *and* h*, as* $T \to \infty$*,*

$$\frac{\mathbf{E}'\boldsymbol{\vartheta}_h}{\sqrt{T_{mh}}} \to_d N\left(0, \sigma_e^2 \boldsymbol{\Xi}\right), \tag{A.2}$$

where matrix $\boldsymbol{\Xi}$ *is defined in (25).*

Proof. Result (A.2) is established in Chudik and Georgiadis (2022). See result (A.8) in Chudik and Georgiadis (2022). ■

Lemma 3 *Consider matrix* $\mathbf{E}$ *defined below (21). Under Assumption 3, for fixed* m *and* h*, as* $T \to \infty$*,*

$$T_{mh}^{-1/2} Vec\left(\mathbf{E}'\mathbf{E} - \sigma_e^2 \mathbf{I}_{(h+1)m}\right) \to_d N\left(0, \boldsymbol{\Sigma}_E\right), \tag{A.3}$$

where $T_{mh} = T/m - h$*.*

Proof. $\left(\mathbf{E}'\mathbf{E} - \sigma_e^2 \mathbf{I}_{(h+1)m}\right)$ is an $m(h+1) \times m(h+1)$ matrix composed of elements

$$\sum_{s=h+1}^{T_m} \left[e_{i_1,s-r_1} e_{i_2,s-r_2} - E\left(e_{i_1,s-r_1}, e_{i_2,s-r_2}\right)\right] = \sum_{s=h+1}^{T_m} \xi_{s,r_1,r_2,i_1,i_2},$$

for $r_1, r_2 = 0, 1, ..., h$ and $i_1, i_2 = 0, 1, ..., m-1$, where $\xi_{s,r_1,r_2,i_1,i_2} = e_{i_1,s-r_1} e_{i_2,s-r_2} - E\left(e_{i_1,s-r_1}, e_{i_2,s-r_2}\right)$, and e_{is} are the sequentially sampled shocks, $e_{is} = e_{t_s-i}$, and $t_s = ms$. Under Assumption 3, $e_t \sim IID\left(0, \sigma_e^2\right)$, $\sigma_e^2 > 0$, $E\left(e_t^4\right) = \kappa_4$, and $4 + \epsilon$ moments of e_t exist for some $\epsilon > 0$. For $r_1 = r_2 = r \in \{0, 1, ..., h\}$, $\left\{\xi_{s,r,r,i_1,i_2}\right\}_{s=h+1}^{T_m}$ is a sequence of i.i.d. random variables with mean zero and a finite variance. Hence, by the Lindeberg-Levy central limit theorem, $T_{mh}^{-1/2} \sum_{s=h+1}^{T_m} \xi_{s,r,r,i_1,i_2}$ is asymptotically normally distributed, for fixed m and h, as $T \to \infty$. For $r_1 \neq r_2$, sequence $\left\{\xi_{s,r_1,r_2,i_1,i_2}\right\}_{s=h+1}^{T_m}$ is a special case of a stationary ergodic sequence that satisfies the conditions of central limit Theorem 5.15 of White (1984), and $T_{mh}^{-1/2} \sum_{s=h+1}^{T_m} \xi_{s,r_1,r_2,i_1,i_2}$ is also asymptotically normally distributed, for fixed m and h, as $T \to \infty$. Elements of $\boldsymbol{\Sigma}_E$ are given by $T_{mh}^{-1} \sum_{s=h+1}^{T_m} \sum_{s'=h+1}^{T_m} E\left(\xi_{s,r_1,r_2,i_1,i_2} \xi_{s',r_1,r_2,i_1,i_2}\right)$ for $r_1, r_2 = 0, 1, ..., h$ and $i_1, i_2 = 0, 1, ..., m-1$. We do not provide expression for $\boldsymbol{\Sigma}_E$ since this expression is not essential for any of our results. ■

Lemma 4 *Suppose* $\mathbf{z}_t = (x_t, \mathbf{y}_t')'$ *is given by (1), Assumptions 2-3 hold, and let* $T_{mh} = T/m - h$*. Then, for fixed* k, p, m *and* h*, as* $T \to \infty$*, the LS estimators* $\hat{\boldsymbol{\theta}}_p$*,* $\hat{\boldsymbol{\theta}}_{\omega p}$*, and* $\hat{\boldsymbol{\theta}}_{kp}$ *and* $\hat{\boldsymbol{\theta}}_{\omega kp}$ *are*

asymptotically normally distributed and converge at the usual rate $\sqrt{T_{mh}}$.

Proof. Each of the four LS estimators can be written as $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = (\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{v}$, for a given choice of the regressor data matrix $\mathbf{Q}$ and the respective error vector $\mathbf{v}$, where we use $\boldsymbol{\theta}$ to denote the population value of $\hat{\boldsymbol{\theta}}$. For instance, in the case of the LS estimator $\hat{\boldsymbol{\theta}}_p$, the matrix $\mathbf{Q}$ is a $T_{mh} \times [p + m(h+1)]$ dimensional data matrix consisting of p lags of the dependent variable $\bar{x}_{ws}$ and $m(h+1)$ regressors $e_{i,s-r}$ for $i = 0, 1, ..., m-1$ and $r = 0, 1, ..., h$, and the elements of $\mathbf{v}$ are given by v_{phs}, see (44). For simplicity, denote the rows of $\mathbf{Q}$ as $\mathbf{q}_s$ and the elements of $\mathbf{v}$ as v_s, for $s = p+1, p+2, ..., T_{mh}$. For each of the four LS estimators, $\mathbf{q}_s$ is stationary and ergodic in variance under Assumption 2-3, and $T_{mh}^{-1}\mathbf{Q}'\mathbf{Q}$ converges in probability to a positive definite matrix. In addition, $\mathbf{q}_s v_s$ is stationary and ergodic with finite variance, and given the eigenvalue condition in Assumption 2, conditions of Theorem 5.15 of White (1984) are satisfied. Hence, $T_{mh}^{-1/2}\mathbf{Q}'\mathbf{v}$ is asymptotically normally distributed. This ensures asymptotic normality of $\hat{\boldsymbol{\theta}}_p$, $\hat{\boldsymbol{\theta}}_{\omega p}$, and $\hat{\boldsymbol{\theta}}_{kp}$ and $\hat{\boldsymbol{\theta}}_{\omega kp}$. ■

A.2 Proofs of propositions

Proof of Proposition 1. Using (22), the mean group DL estimator can be written as $\widehat{\bar{\mathbf{d}}}_{\omega}^{MG} = (\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}')\hat{\mathbf{d}}$, where $\hat{\mathbf{d}}$ is given by (21). Using (20) in (21), we obtain

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}_{\omega}^{MG} - \bar{\mathbf{d}}_{\omega}\right) = (\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}')\left(\frac{\mathbf{E}'\mathbf{E}}{T_{mh}}\right)^{-1}\frac{\mathbf{E}'\boldsymbol{\vartheta}_h}{\sqrt{T_{mh}}}, \tag{A.4}$$

where $\boldsymbol{\vartheta}_h = (\vartheta_{h,h+1}, \vartheta_{h,h+2}, ..., \vartheta_{h,T_m})'$. Using result (A.1) of Lemma 1 and result (A.2) of Lemma 2 in (A.4), result (23) readily follows. ■

Proof of Proposition 2. From (34), we can write $\bar{\mathbf{x}}_w$ as $\bar{\mathbf{x}}_w = \tilde{\mathbf{E}}\bar{\mathbf{d}}_{\omega} + \boldsymbol{\vartheta}_h^* = \tilde{\mathbf{E}}\bar{\mathbf{d}}_{\omega} + \boldsymbol{\vartheta}_h + \mathbf{E}\boldsymbol{\Delta}$, where $\boldsymbol{\Delta}$ is vector of stacked Δ_{ir}, namely $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_0', \boldsymbol{\Delta}_1', ..., \boldsymbol{\Delta}_h')'$, in which $\boldsymbol{\Delta}_r = (\Delta_{0r}, \Delta_{1r}, ..., \Delta_{m-1,r})'$ for $r = 0, 1, ..., h$. Substituting this expression for $\bar{\mathbf{x}}_w$ into the definition of $\widehat{\bar{\mathbf{d}}}_{\omega}^{P}$ given by (35), we obtain

$$\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}_{\omega}^{P} - \bar{\mathbf{d}}_{\omega}\right) = \left(\frac{\tilde{\mathbf{E}}'\tilde{\mathbf{E}}}{T_{mh}}\right)^{-1}\frac{\tilde{\mathbf{E}}'(\boldsymbol{\vartheta}_h + \mathbf{E}\boldsymbol{\Delta})}{\sqrt{T_{mh}}}, \tag{A.5}$$

where $\tilde{\mathbf{E}}$ can be written as $\tilde{\mathbf{E}} = \mathbf{E}(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}})$, see (30). We consider the individual terms on the

right side of (A.5) in turn. Regarding the first term, we have

$$\frac{\tilde{\mathbf{E}}'\tilde{\mathbf{E}}}{T_{mh}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \frac{\mathbf{E}'\mathbf{E}}{T_{mh}} \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right),$$

and

$$\frac{\tilde{\mathbf{E}}'\tilde{\mathbf{E}}}{T_{mh}} \to_p \sigma_e^2 \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \mathbf{I}_{(h+1)m} \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right) = \sigma_e^2 \left\|\tilde{\boldsymbol{\omega}}\right\|^2 \mathbf{I}_{h+1} = \sigma_e^2 \left\|\boldsymbol{\omega}\right\|^{-2} \mathbf{I}_{h+1}, \tag{A.6}$$

where $T_{mh}^{-1}\mathbf{E}'\mathbf{E} \to_p \sigma_e^2 \mathbf{I}_{(h+1)m}$ by (A.1), $\left\|\tilde{\boldsymbol{\omega}}\right\| = \sqrt{\tilde{\boldsymbol{\omega}}'\tilde{\boldsymbol{\omega}}}$ is the Euclidean norm of $\tilde{\boldsymbol{\omega}}$, and

$$\left\|\tilde{\boldsymbol{\omega}}\right\|^2 = \tilde{\boldsymbol{\omega}}'\tilde{\boldsymbol{\omega}} = \sum_{i=0}^{m-1} \tilde{\omega}_i^2 = \left(\sum_{i=0}^{m-1} \omega_i^2\right)^{-1} = \left\|\boldsymbol{\omega}\right\|^{-2}, \tag{A.7}$$

follows from (29). Hence,

$$\left(\frac{\tilde{\mathbf{E}}'\tilde{\mathbf{E}}}{T_{mh}}\right)^{-1} \to_p \sigma_e^{-2} \left\|\boldsymbol{\omega}\right\|^2 \mathbf{I}_{h+1}. \tag{A.8}$$

Consider next

$$\frac{\tilde{\mathbf{E}}'\boldsymbol{\vartheta}_h}{\sqrt{T_{mh}}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \frac{\mathbf{E}'\boldsymbol{\vartheta}_h}{\sqrt{T_{mh}}} \to_d N\left(0, \sigma_e^2 \tilde{\boldsymbol{\Xi}}\right), \tag{A.9}$$

where we used (A.2) of Lemma 3, and $\tilde{\boldsymbol{\Xi}}$ is given by

$$\tilde{\boldsymbol{\Xi}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \boldsymbol{\Xi} \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right).$$

Finally, consider

$$\frac{\tilde{\mathbf{E}}'\mathbf{E}\boldsymbol{\Delta}}{\sqrt{T_{mh}}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \frac{\mathbf{E}'\mathbf{E}\boldsymbol{\Delta}}{\sqrt{T_{mh}}}, \tag{A.10}$$

where $\boldsymbol{\Delta} = \left(\boldsymbol{\Delta}_0', \boldsymbol{\Delta}_1', ..., \boldsymbol{\Delta}_h'\right)'$, and $\boldsymbol{\Delta}_r$, for $r = 0, 1, ..., h$, is given by $\boldsymbol{\Delta}_r = \mathbf{d}_r - \tilde{\boldsymbol{\omega}}\left(\boldsymbol{\omega}'\mathbf{d}_r\right)$. Hence,

$$\tilde{\boldsymbol{\omega}}'\boldsymbol{\Delta}_r = \tilde{\boldsymbol{\omega}}'\mathbf{d}_r - \tilde{\boldsymbol{\omega}}'\tilde{\boldsymbol{\omega}}\boldsymbol{\omega}'\mathbf{d}_r = \tilde{\boldsymbol{\omega}}'\mathbf{d}_r - \left\|\tilde{\boldsymbol{\omega}}\right\|^2 \boldsymbol{\omega}'\mathbf{d}_r = \tilde{\boldsymbol{\omega}}'\mathbf{d}_r - \left\|\boldsymbol{\omega}\right\|^{-2} \boldsymbol{\omega}'\mathbf{d}_r = \tilde{\boldsymbol{\omega}}'\mathbf{d}_r - \tilde{\boldsymbol{\omega}}'\mathbf{d}_r = \mathbf{0},$$

for $r = 0, 1, ..., h$, where the third equality follows from (A.7), and the fourth equality follows by noting $\tilde{\boldsymbol{\omega}} = \left\|\boldsymbol{\omega}\right\|^{-2} \boldsymbol{\omega}$ (see (29)). The identity $\tilde{\boldsymbol{\omega}}'\boldsymbol{\Delta}_r = \mathbf{0}$, for $r = 0, 1, ..., h$, implies $\left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \boldsymbol{\Delta} = \mathbf{0}$, so we obtain

$$\sigma_e^2 \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \mathbf{I}_{(h+1)m} \boldsymbol{\Delta} = \mathbf{0}. \tag{A.11}$$

Subtracting (A.11) from (A.10), we obtain

$$\frac{\tilde{\mathbf{E}}'\mathbf{E}\boldsymbol{\Delta}}{\sqrt{T_{mh}}} = \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right) \left(\frac{\mathbf{E}'\mathbf{E} - \sigma_e^2 \mathbf{I}_{(h+1)m}}{\sqrt{T_{mh}}}\right) \boldsymbol{\Delta}. \tag{A.12}$$

Given that $Vec\left(\mathbf{ABC}\right) = \left(\mathbf{C}' \otimes \mathbf{A}\right) Vec\left(\mathbf{B}\right)$ for conformable matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, vectorization of (A.12) yields

$$Vec\left(\frac{\tilde{\mathbf{E}}'\mathbf{E}\boldsymbol{\Delta}}{\sqrt{T_{mh}}}\right) = \left[\boldsymbol{\Delta}' \otimes \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right)\right] Vec\left(\frac{\mathbf{E}'\mathbf{E} - \sigma_e^2 \mathbf{I}_{(h+1)m}}{\sqrt{T_{mh}}}\right). \tag{A.13}$$

By Lemma 3,

$$Vec\left(\frac{\mathbf{E}'\mathbf{E} - \sigma_e^2 \mathbf{I}_{(h+1)m}}{\sqrt{T_{mh}}}\right) \rightarrow_d N\left(0, \boldsymbol{\Sigma}_E\right). \tag{A.14}$$

Using (A.14) in (A.13), it follows

$$\frac{\tilde{\mathbf{E}}'\mathbf{E}\boldsymbol{\Delta}}{\sqrt{T_{mh}}} \rightarrow_d N\left(0, \boldsymbol{\Sigma}_\Delta\right), \tag{A.15}$$

where $\boldsymbol{\Sigma}_\Delta = \left[\boldsymbol{\Delta}' \otimes \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}'\right)\right] \boldsymbol{\Sigma}_E \left[\boldsymbol{\Delta} \otimes \left(\mathbf{I}_{h+1} \otimes \tilde{\boldsymbol{\omega}}\right)\right]$. Using results (A.8), (A.9) and (A.15) in (A.5), we obtain (36). ■

Proof of Proposition 3. Proposition 3 is a direct application of the delta method. Under Assumptions 2-3, Lemma 4 establishes that $\hat{\boldsymbol{\theta}}_p$ is asymptotically normally distributed. In addition, the mapping between the estimates of the ARDL model and its multipliers, $F : \boldsymbol{\theta}_p \rightarrow \mathbf{d}$, defined by (50)-(53) is continuously differentiable and its Jacobian $\mathbf{J}_F\left(\boldsymbol{\theta}_p\right)$ is nonsingular. Invoking the delta method, it follows that $\hat{\mathbf{d}}_p^A = F\left(\hat{\boldsymbol{\theta}}_p\right)$ is asymptotically normal and $\sqrt{T_{mh}}\left(\hat{\mathbf{d}}^A - \mathbf{d}\right) \rightarrow_d N\left(\mathbf{0}, \mathbf{J}_F\left(\boldsymbol{\theta}_p\right) \boldsymbol{\Sigma}_{\theta p} \mathbf{J}_F'\left(\boldsymbol{\theta}_p\right)\right)$, where $\boldsymbol{\Sigma}_{\theta p}$ is the asymptotic variance of $\hat{\boldsymbol{\theta}}_p$. Using this result in $\hat{\mathbf{d}}_\omega^{MGA} = \left(\mathbf{I}_{h+1} \otimes \boldsymbol{\omega}'\right) \hat{\mathbf{d}}^A$, we obtain (56). ■

Proof of Proposition 4. Similarly to the proof of Proposition 3, result (70) can be established using the delta method. Mapping $F_\omega : \boldsymbol{\theta}_{\omega p} \rightarrow \bar{\mathbf{d}}_\omega$, given by (62)-(65), is continuously differentiable and its Jacobian $\mathbf{J}_{F\omega}\left(\boldsymbol{\theta}_{\omega p}\right)$ is nonsingular. Under Assumption 2-3, $\hat{\boldsymbol{\theta}}_{\omega p}$ is asymptotically normal (see Lemma 4), and $\sqrt{T_{mh}}\left(\hat{\boldsymbol{\theta}}_{\omega p} - \boldsymbol{\theta}_\omega\right) \rightarrow_d N\left(\mathbf{0}, \boldsymbol{\Sigma}_{\theta\omega p}\right)$. Using the delta method, it follows that $\sqrt{T_{mh}}\left(\hat{\bar{\mathbf{d}}}_{\omega p}^{PA} - \bar{\mathbf{d}}_\omega\right) \rightarrow_d N\left(\mathbf{0}, \mathbf{J}_{F\omega}\left(\boldsymbol{\theta}_{\omega p}\right) \boldsymbol{\Sigma}_{\omega\theta p} \mathbf{J}_{F\omega}'\left(\boldsymbol{\theta}_{\omega p}\right)\right)$, as required. ■

Proof of Proposition 5. Consider mappings $F_k\left(\boldsymbol{\theta}_{kp}\right) = \mathbf{d}_{(k)}$ defined by (77)-(80) and $F_{\omega k}\left(\boldsymbol{\theta}_{\omega kp}\right) = \bar{\mathbf{d}}_{\omega k}$ defined by (83)-(86). Both mappings are continuously differentiable with non-singular Jaco-

bians matrices $\mathbf{J}_{Fk}(\boldsymbol{\theta}_{kp})$ and $\mathbf{J}_{F\omega k}(\boldsymbol{\theta}_{\omega kp})$, respectively. Similarly to the proofs of Propositions 3 and 4, under Assumptions 2-3, the LS estimators $\hat{\boldsymbol{\theta}}_{kp}$ and $\hat{\boldsymbol{\theta}}_{\omega kp}$ are both asymptotically normally distributed and converge at the rate $T_{mh}^{-1/2}$ (see Lemma 4). Using the delta method, we obtain $\sqrt{T_{mh}}\left(\widehat{\mathbf{d}}_{(k),p}^{MGV}-\mathbf{d}_{(k)}\right)\rightarrow_d N\left(\mathbf{0},\mathbf{J}_{Fk}(\boldsymbol{\theta}_{kp})\boldsymbol{\Sigma}_{\theta kp}\mathbf{J}'_{Fk}(\boldsymbol{\theta}_{kp})\right)$ and result (88) now follows from noting $\widehat{\bar{\mathbf{d}}}_{\omega kp}^{MGV}=(\mathbf{I}_{h+1}\otimes\boldsymbol{\omega}')\left(\mathbf{I}_{h+1}\otimes\mathbf{s}'_{k,1}\right)\widehat{\mathbf{d}}_{(k),p}^{MGV}$. Similarly, the delta method implies the asymptotic normality of the pooled estimator, $\sqrt{T_{mh}}\left(\widehat{\bar{\mathbf{d}}}_{\omega kp}^{PV}-\bar{\mathbf{d}}_{\omega}\right)\rightarrow_d N\left(\mathbf{0},\mathbf{J}_{F\omega k}(\boldsymbol{\theta}_{\omega kp})\boldsymbol{\Sigma}_{\omega k\theta p}\mathbf{J}'_{F\omega k}(\boldsymbol{\theta}_{\omega kp})\right)$.

■

A.3 Data description for inflation pass-through application

- G_s: Daily gasoline price data from NY Harbor, regular, available from: https://www.eia.gov/dnav/pet/hist/eer_epmru_pf4_y35ny_dpgD.htm (downloaded on September 24, 2025). Series name in the database: "New York Harbor Conventional Gasoline Regular Spot Price FOB (Dollars per Gallon)". Data availability: June 02, 1986 - September 22, 2025, working days only. ($T_D = 9877$) Units: $ per gallon.

- P_t: Seasonally adjusted headline CPI inflation from FRED, series "CPILFESL," Consumer Price Index for All Urban Consumers: All Items in U.S. City Average, Index 1982-1984=100, Monthly, Seasonally Adjusted. Downloaded on September 24, 2025, from: https://fred.stlouisfed.org/series/CPIAUCSL. Data availability: 1947M1 - 2025M8 ($T_M = 944$). Units: Index 1982-1984=100.

- P_{ct}: Seasonally adjusted CPI inflation ex food and energy from FRED (core CPI), series "CPILFESL," Consumer Price Index for All Urban Consumers: All Items Less Food and Energy in U.S. City Average, Index 1982-1984=100, Monthly, Seasonally Adjusted. Downloaded on September 24, 2025, from: https://fred.stlouisfed.org/series/CPILFESL. Data availability: 1957M1 - 2025M8 ($T_M = 824$). Units: Index 1982-1984=100.

- P_{gt}: Seasonally adjusted gasoline consumer price index from FRED. Series "CUSR0000SETB01," Consumer Price Index for All Urban Consumers: Gasoline (All Types) in U.S. City Average. Downloaded on September 24, 2025, from: https://fred.stlouisfed.org/series/CUUR0000SETB01. Data availability: 1935M13- 2025M8 but only the subperiod 1966M1 - 2025M8 has no gaps ($T_M = 705$). Units: Index 1982-1984=100.

The common sample for all variables is 1986M6 - 2025M8 ($T_M = 471$).

Figure B1 below reports the histogram of the number of daily gasoline observations per month.

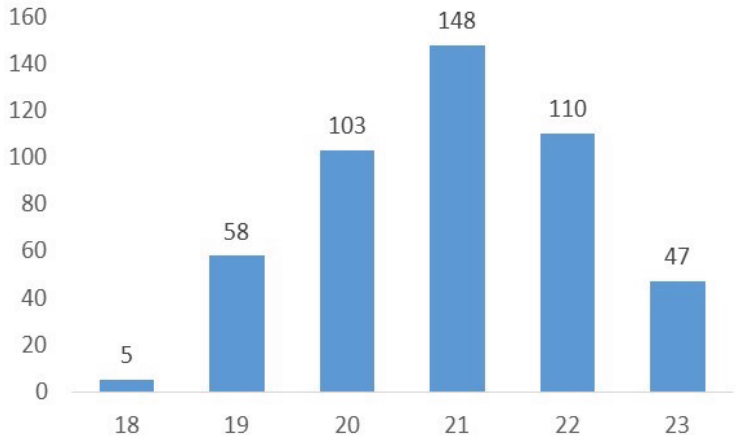


FIGURE B1: Histograms of the number of trading days per month. Sample: 1986M6 - 2025M8.