Estimating Structural Parameters Using Vector Autoregressions∗

May 13, 2021

Christiane Baumeister
University of Notre Dame
cjsbaumeister@gmail.com

James D. Hamilton
University of California at San Diego
jhamilton@ucsd.edu

ABSTRACT

A number of prominent economic studies have tried to estimate behavioral elasticities from the ratios of elements of a single column of the impact matrix of a structural vector autoregression. This approach would be valid if there were only two variables in the VAR. But we demonstrate that in general it is not valid when there are more than two variables. We describe the optimal methods that applied researchers or policymakers should rely on to estimate elasticities or any other structural parameters using vector autoregressions.

JEL codes: C11, C32, Q43

Key words: structural vector autoregressions, Bayesian analysis, identification, elasticities, sign restrictions

∗This paper supersedes an earlier paper by the authors that was circulated under the title “Advances in Structural Vector Autoregressions with Imperfect Identifying Information.”
1 Introduction.

Vector autoregressions (VARs) offer a convenient way to summarize the dynamic correlations among a set of observed variables and are easily estimated by OLS regressions. In order to draw structural conclusions from those OLS regressions, we need to bring in some additional identifying information. Alternative approaches to using identifying information in VARs have been discussed by Amisano and Giannini (2012), Stock and Watson (2016), and Baumeister and Hamilton (2021), among many others. Bernanke (1986) noted that if we are only interested in the effects of one particular structural shock, it is not necessary to identify all of the structural parameters of the system as long as we can estimate the contemporaneous impact of the one structural shock of interest on all the variables in the system.

It has recently become a common practice among some applied researchers to try to draw conclusions about other structural parameters on the basis of the estimated effects of a single structural shock. For example, Kilian and Murphy (2014) proposed to estimate the elasticity of oil demand from the ratio of the change in oil consumption to the change in price in response to a shock to the supply of oil. Examples of other studies that have tried to estimate elasticities on the basis of the ratio of responses of different variables to a single structural shock include Kilian and Murphy (2012), Güntner (2014), Riggi and Venditti (2015), Kilian and Lütkepohl (2017), Ludvigson et al. (2017), Antolín-Díaz and Rubio-Ramírez (2018), Basher et al. (2018), Herrera and Rangaraju (2020), and Zhou (2020).

In this paper we show that if the VAR includes more than two variables, this popular approach is not estimating a demand elasticity, and indeed is not estimating any other identifiable feature of the structural equation characterizing demand. Instead it is estimating an equilibrium consequence of an oil supply shock that in general depends on parameters of all of the structural equations of the system. In Section 2 we demonstrate this using a simple 3-equation model of the global oil market. Section 3 discusses valid ways to use identifying information to estimate structural elasticities. Section 4 comments on some of the discussion related to this topic in the recent literature.

2 The relation between elasticities and the ratio of impacts.

Every economics textbook defines the price elasticity of demand as the response of buyers of the product to an increase in the price with other variables that influence demand held constant. We can illustrate this with a simple dynamic demand equation in which \( q_t \) is the log of the quantity of oil purchased, \( p_t \) is the log of the real price of oil, and \( y_t \) is the log of real income:

\[
q_t = \delta y_t + \beta p_t + b'x_{t-1} + u_t. \tag{1}
\]
In this equation, $\beta$ is the short-run price elasticity of demand, $\delta$ is the short-run income elasticity of demand, $u_t^d$ is a shock to demand, $x_{t-1} = (1, y_{t-1}', y_{t-2}', \ldots, y_{t-m}')'$ is a vector consisting of a constant term and $m$ lags of each of the three variables with $y_t = (q_t, y_t, p_t)'$, and $b_d$ characterizes the response of demand to lagged values of the variables. To see the relation between the elasticity $\beta$ and a magnitude inferred from the impacts of structural shocks, we embody (1) in a dynamic structural system that also describes the behavior of oil producers and the determinants of income:

$$q_t = \gamma y_t + \alpha p_t + b'_s x_{t-1} + u_t^s$$

(2)

$$y_t = \xi q_t + \psi p_t + b'_y x_{t-1} + u_t^y.$$  

(3)

Here for example $\alpha$ is the short-run price elasticity of oil supply, $u_t^s$ is a shock to oil production, and $\psi$ is the contemporaneous effect of oil prices on economic activity. We can write this structural model in vector form as

$$A y_t = B x_{t-1} + u_t$$

(4)

$$A = \begin{bmatrix} 1 & -\gamma & -\alpha \\ -\xi & 1 & -\psi \\ 1 & -\delta & -\beta \end{bmatrix}$$

$$u_t = (u_t^s, u_t^y, u_t^d)'$$

$$B = \begin{bmatrix} b'_s \\ b'_y \\ b'_d \end{bmatrix}.$$  

In most applications, the structural shocks are viewed as serially uncorrelated and uncorrelated with each other,

$$E(u_t u_t') = \begin{cases} D & \text{for } t = s \\ 0 & \text{for } t \neq s \end{cases},$$

with $D$ diagonal.

The contemporaneous impacts of the structural shocks on the observed variables are given by the matrix

$$H = \frac{\partial y_t}{\partial u_t} = A^{-1} = |A|^{-1} \begin{bmatrix} -\beta - \delta \psi & \alpha \delta - \beta \gamma & \alpha + \gamma \psi \\ -\psi - \beta \xi & \alpha - \beta & \psi + \alpha \xi \\ \delta \xi - 1 & \delta - \gamma & 1 - \gamma \xi \end{bmatrix}. \tag{5}$$

What would we get if we tried to estimate the demand elasticity on the basis of the ratio of the change in $q_t$ to the change in $p_t$ in response to a shock to supply $u_t^s$? For this system that
ratio is given by\(^1\)
\[
\frac{h_{11}}{h_{31}} = \frac{-\beta - \delta \psi}{\delta \xi - 1}.
\] (6)
In general, expression (6) is not the demand elasticity \(\beta\). The reason is that if there is a shock to \(u_t^e\), not only will it change the price \(p_t\), but it will also change income. The size of the change in price is \(|A|^{-1}(\delta \xi - 1)\) and the size of the change in income is \(|A|^{-1}(-\psi - \beta \xi)\). From the demand curve, the change in price will lead to a change in quantity demanded of \(\beta\) times the change in price, namely \(\beta |A|^{-1}(\delta \xi - 1)\). Likewise the change in income will lead to a change in quantity demanded of \(\delta\) times the change in income, namely \(\delta |A|^{-1}(-\psi - \beta \xi)\). The observed change in quantity demanded in response to the shock in supply is the sum of these two terms,

\[
\begin{align*}
&\quad \beta |A|^{-1}(\delta \xi - 1) \\
&\quad \delta |A|^{-1}(-\psi - \beta \xi) = |A|^{-1}(-\beta - \delta \psi).
\end{align*}
\]

Dividing this by the magnitude of the change in price that results from the supply shock, \(|A|^{-1}(\delta \xi - 1)\), produces the result (6).

In the special case when demand does not respond to income (\(\delta = 0\)), expression (6) would simplify to the correct answer \(\beta\). But in general, expression (6) reflects a combination of the sensitivity of demand to price, the sensitivity of demand to income, and the effects of an oil supply shock on those two variables.

Expression (6) does not summarize the characteristics of demand but instead characterizes the equilibrium impact of the structural shock. This is a fundamental problem for any study that attempts to calculate structural elasticities from the ratios of the effects of a particular structural shock. In a different example, Kilian and Murphy (2014) considered a 4-equation model in which there were two different kinds of demand shocks. These authors calculated the short-run price elasticity of supply in two different ways, first as the ratio of the change in quantity to the change in price resulting from the first demand shock, and second as the ratio of the change in quantity to the change in price resulting from the second demand shock. Kilian and Murphy supposed that either of these magnitudes could be regarded as estimates of the supply elasticity. In practice they will be two different numbers,\(^2\) and neither corresponds

\(^1\)Many empirical studies normalize the standard deviation of structural shocks to be unity, writing the structural system (4) in the form \(A^*y_t = B^*x_{t-1} + u_t^e\) where \(A^* = D^{-1/2}A\) and \(E(u_t^*u_t^{e\prime}) = I_n\). The impacts of structural shocks in this normalization are given by \(H^* = HD^{1/2}\). Note that the ratio \(h_{11}^*/h_{31}^*\) in this normalization is identical to the ratio \(h_{11}/h_{31}\) analyzed here since \(h_{11}^*/h_{31}^* = (h_{11}/\sqrt{d})/(h_{31}/\sqrt{d}) = h_{11}/h_{31}\).

\(^2\)For example, running the code for Kilian and Murphy (2014) that is publicly posted in the Journal of Applied Econometrics data archive generates 5 million draws for the vector of possible parameters. Their algorithm discards draws based on a long list of criteria other than the supply elasticity, from which 24,926 survive. The code then calculates one supply elasticity from the ratio of the change in quantity to the change in price in response to what the authors interpret as an aggregate demand shock. The median value for this magnitude across the 24,926 draws is 0.1184. The code next calculates a second supply elasticity from the ratio of the change in quantity to the change in price in response to what the authors interpret as a speculative
to the usual understanding of what we mean by the supply elasticity, which is the parameter \( \alpha \) in the structural equation (2).

## 3 Correct estimation of elasticities.

What is the correct way to estimate a parameter like the demand elasticity \( \beta \) in the system (4)? We cannot estimate the demand parameters from OLS estimation of (1) because the error \( u^d_t \) is correlated with the explanatory variables \( p_t \) and \( y_t \). For example, equations (2) and (3) imply that an increase in demand leads in equilibrium to an increase in price, inducing a positive correlation between \( u^d_t \) and \( p_t \).

### 3.1 Frequentist inference.

The traditional solution to this problem is to find instruments that are correlated with \( p_t \) and \( y_t \) but uncorrelated with \( u^d_t \). As an example of suitable instruments, the model implies that the supply shock \( u^s_t \) and income shock \( u^y_t \) affect \( p_t \) and \( y_t \) but are uncorrelated with \( u^d_t \). Thus if we had enough information about the other structural equations to be able to find consistent estimates \( \hat{u}^s_t \) and \( \hat{u}^y_t \) of these other shocks, we could use these as instruments to estimate (1).

To examine in more detail these IV estimates, recall that the structural system (4) implies the reduced-form VAR

\[
y_t = \Pi x_{t-1} + \varepsilon_t
\]

where

\[
\Pi = A^{-1}B
\]

and

\[
\varepsilon_t = A^{-1}u_t.
\]

The variance matrix of \( \varepsilon_t \) is given by \( E(\varepsilon_t \varepsilon'_t) = \Omega = A^{-1}D(A^{-1})' \). The reduced-form parameters are readily estimated by OLS:

\[
\hat{\Pi} = \left( \sum_{t=1}^{T} y_t x_{t-1} \right) \left( \sum_{t=1}^{T} x_{t-1} x_{t-1}' \right)^{-1}
\]

\[
\hat{\Omega} = T^{-1} \sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t
\]

\[
\hat{\varepsilon}_t = y_t - \hat{\Pi} x_{t-1} = (\varepsilon^q_t, \varepsilon^y_t, \varepsilon^p_t)'.
\]

The IV estimates of \( \delta \) and \( \beta \) using \( \hat{u}^s_t \) and \( \hat{u}^y_t \) as instruments can then be written as

\[
\begin{bmatrix}
\hat{\delta}_{IV} \\
\hat{\beta}_{IV}
\end{bmatrix} = 
\begin{bmatrix}
\sum_{t=1}^{T} \hat{u}^s_t \varepsilon^q_t \\
\sum_{t=1}^{T} \hat{u}^y_t \varepsilon^y_t \\
\sum_{t=1}^{T} \hat{u}^y_t \varepsilon^p_t
\end{bmatrix}^{-1}
\begin{bmatrix}
\sum_{t=1}^{T} \hat{u}^s_t \varepsilon^q_t \\
\sum_{t=1}^{T} \hat{u}^y_t \varepsilon^y_t \\
\sum_{t=1}^{T} \hat{u}^y_t \varepsilon^p_t
\end{bmatrix}
\]

The demand shock. The median value of this magnitude is 0.6184. Their code rejects the draw if either of the two supply elasticities is larger than 0.0258. Only 16 of the original 5 million draws survive all the restrictions. The results in their paper are based on properties of these 16 final retained draws.
And where would we obtain consistent estimates of the structural shocks \( \hat{u}_t^s \) and \( \hat{u}_t^y \)? One simple example is if we knew from other information the true values of the structural parameters \( \gamma, \alpha, \xi, \) and \( \psi \). In this case we would have

\[
\begin{bmatrix}
\hat{u}_t^s \\
\hat{u}_t^y
\end{bmatrix} = \Gamma \hat{\xi}_t
\]

\[
\Gamma = \begin{bmatrix}
1 & -\gamma & -\alpha \\
-\xi & 1 & -\psi
\end{bmatrix}.
\]

This allows (10) to be expressed as

\[
\Gamma \hat{\Omega} \hat{\eta}_{IV} = 0
\]

for \( \eta = (1, -\delta, -\beta)' \); for details, see Appendix A.

More generally, if we make the distributional assumption that the structural shocks are \( N(0, D) \), the log likelihood for the structural model is given by

\[
L(A, D, B) = -\left(\frac{Tn}{2}\right) \log(2\pi) - \left(\frac{T}{2}\right) \log |A^{-1}D(A^{-1})'| - \left(\frac{1}{2}\right) \sum_{t=1}^{T} \left( y_t - A^{-1}Bx_{t-1}' \right) \left( A^{-1}D(A^{-1})' \right)^{-1} \left( y_t - A^{-1}Bx_{t-1} \right) \cdot (12)
\]

If we impose enough restrictions to result in the structural parameters being just-identified, then maximizing (12) subject to those restrictions results in estimates characterized by

\[
\hat{D}_{MLE} = \hat{A}_{MLE} \hat{\Omega}_{MLE}';
\]

see Hamilton (1994, equation 11.6.33). But note that

\[
A_{(3 \times 3)} = \begin{bmatrix}
\Gamma_{(2 \times 3)} \\
\eta'_{(1 \times 3)}
\end{bmatrix}.
\]

Since \( D \) is diagonal, the (1,3) and (2,3) elements of (13) state that

\[
\hat{\Gamma}_{MLE} \hat{\Omega}_{MLE} = 0.
\]

Comparing (14) with (11), it is clear that maximum likelihood estimation of the structural model (4) subject to the identifying restrictions is just a generalization of the familiar idea of estimation by instrumental variables.\(^3\) Moreover, we know that when restrictions are just-identifying or over-identifying, the MLE of the demand elasticity \( \beta \) has a smaller asymptotic

\(^3\)To our knowledge, Shapiro and Watson (1988) were the first to point out the IV interpretation of maximum likelihood estimation of structural VARs.
variance than any other consistent estimate of \( \beta \); see Rothenberg (1973, Chapter 2). Hence for frequentist inference of a fully identified structural model, the question of how we should estimate \( \beta \) has a very clear answer: we should estimate \( \beta \) along with all the other structural parameters directly by maximizing the likelihood function (12).

### 3.2 Bayesian inference.

Suppose next that some of our information about structural parameters is imperfect. We can characterize inexact information about the structural parameters in the form of a Bayesian probability distribution \( p(A, D, B) \), where higher values for \( p(A, D, B) \) are associated with parameter values that previous research or economic theory suggests are more plausible, and is zero for values that we can rule out a priori. Bayes’ Law then allows us to calculate the posterior distribution \( p(A, D, B|Y) \) for \( Y = (y_{T,1}', y_{T-1,1}',...,y_{m+1}'')' \) the set of observed data. Given a loss function that characterizes the penalty for reporting an estimate that differs from the true value, statistical decision theory calls for minimizing the expected loss as calculated using the posterior distribution. For example, for a quadratic loss function, we should use the posterior mean as our estimate of \( \beta \); see for example Baumeister and Hamilton (2018).

Even if the researcher is not a Bayesian, if there is information about \( A, B, \) or \( D \) from a previous sample, incorporating that information results in the smallest possible asymptotic variance for \( \hat{\beta} \) from a frequentist perspective, as again shown by Rothenberg (1973, Chapter 3). Intuitively, if we replace the assumption that \( \Gamma \) is known a priori with certainty with a probability distribution \( p(\Gamma) \) over plausible values, the Bayesian posterior inference will weight (11) by the prior plausibility of \( \Gamma \) and takes an integral over all the possible values of \( \Gamma \).

Thus when our information about structural parameters is inexact, the question of how to estimate the demand elasticity \( \beta \) again has a clear answer: we should represent the prior information in the form of a probability distribution, use Bayes’ Law to calculate the posterior distribution of parameters, and use as an estimate of \( \beta \) the value that minimizes the expected posterior loss.

### 3.3 Inference when there is only information about a single structural shock.

Returning to the original case discussed by Bernanke (1986), what can we do if we only have information about the effects of a single structural shock, and have no information at all, even imperfect or inexact information, about anything else in the system? This may arise for example in applications of the proxy-variable methods proposed by Stock and Watson (2012) and Mertens and Ravn (2013) if we only have available a proxy for one of the structural shocks. The effects of the single shock are easy to estimate using the instrument in local projections (Plagborg-Møller and Wolf (2021)). In this case, the researcher simply has to acknowledge
that there is no basis for reporting an estimate of parameters like behavioral elasticities. The impacts of the single identified shock are the only thing that we are able to report.

4 Discussion.

In this section we relate these results to some other claims in the recent literature.

4.1 Role of parameterization.

Many researchers normalize structural shocks to have unit variance, namely \( u_t^* = D^{-1/2}u_t \), and consider the effects of the structural shocks directly using a parameterization based on (7) and (9):

\[
y_t = \Pi x_{t-1} + H^* u_t^*
\]

\[
H^* = \frac{A^{-1}D^{1/2}}{A^{-1}D^{1/2}}
\]

\[
H^* = \frac{\partial y_t}{\partial u^*_t}.
\]

The matrix \( H^* \) summarizes the contemporaneous effects of one-standard-deviation structural shocks on the observed vector \( y_t \). \( Uhlig (2017) \) has argued that the \( H^* \) parameterization is to be preferred since from the perspective of policy, what we often care about are the equilibrium effects of possible interventions. None of the points we have been making depend in any way on whether the structural model is parameterized as (4) or (15). Whether one is interested in \( A \) or \( A^{-1} \), the principles are the same and the method of estimation is the same. From the perspective of maximum likelihood estimation, any restriction on \( A \) and \( D \) can be translated into an analogous restriction on \( H^* \) in (16). Maximum likelihood estimates are invariant with respect to parameterization, meaning that the inference about \( H^* \) based on the parameterization (4) will be identical to the inference about \( H^* \) based on the parameterization (15), with the equation

\[
\hat{H}^*_{MLE} = \hat{A}^{-1}_{MLE} \hat{D}^{1/2}_{MLE}
\]

holding exactly. For Bayesian inference, prior information about \( A \) and \( D \) can be translated into exactly equivalent prior information about \( H^* \) using the change-in-variables formula. Bayesian posterior inference about \( H^* \) will again be exactly the same regardless of the parameterization.

The issue is not whether elements of \( A \) or elements of \( H^* \) are the objects of interest. As we noted in the introduction, \( Kilian and Murphy (2012) \), \( Güntner (2014) \), \( Riggi and Venditti (2015) \), \( Kilian and Lütkepohl (2017) \), \( Ludvigson et al. (2017) \), \( Antolín-Díaz and Rubio-Ramírez (2018) \), \( Basher et al. (2018) \), \( Herrera and Rangaraju (2020) \), and \( Zhou (2020) \) all reported estimates of both. The real question is: What are the structural objects about which the
researcher has prior information? We would argue that prior information often comes in the form of insights about $A$ rather than $H^*$. Most microfounded models take the form of a system like (4), in which individual equations represent the actions of different agents such as consumers, firms, or government, rather than in the form of postulated general equilibrium impacts of the actions of individual actors. Formulating a prior $p(A, B, D)$ typically involves looking at previous findings about elasticities (Baumeister and Hamilton (2019), Brinca et al. (2020), Aastveit et al. (2022), Valenti et al. (2020)), policy rules (Baumeister and Hamilton (2018), Nguyen (2019), Belongia and Ireland (2021)), behavioral equations from economic theory (Aruoba et al. (2021), Lukmanova and Rabitsch (2021)), and responses of agents to permanent changes (Baumeister and Hamilton (2015)). Typically these are most naturally represented as information about $A$, not $A^{-1}$, even though they all have implications for prior information about $A^{-1}$.

Notwithstanding, researchers may also have some useful information about the equilibrium impacts of structural shocks. For example, extremely large impacts of policy changes on broad macroeconomic variables may be regarded as unlikely, or we may claim to know a priori the signs of certain elements of $H$. There is no problem in incorporating information about $H$ as a supplement to information about $A$. Suppose that for the system given by (1)-(3) (and the necessary implication of those three equations in the form of expression (5)) we had prior information about both the price elasticity of supply $p_1(\alpha)$ and the contemporary effect of a supply shock on income $p_2(|A|^{-1}(-\psi - \beta \xi))$. Then we could use the product $p(A) \propto p_1(\alpha)p_2(|A|^{-1}(-\psi - \beta \xi))$ as a composite prior for $A$. As discussed by Baumeister and Hamilton (2018), there is no problem with including multiple sources of information about the same parameter, just as there is no problem with using multiple earlier samples that all contain information about a common parameter to form a Bayesian prior in standard settings. The applications in Baumeister and Hamilton (2018, 2019), Grisse (2020), Valenti et al. (2020), Lukmanova and Rabitsch (2021), and Aruoba et al. (2021) all incorporate prior information about both $A$ and $A^{-1}$.

4.2 Role of inventories.

Kilian (2020) has argued that in order to talk about the elasticity of oil demand it is necessary to decompose the quantity of oil purchased by refiners into the portion that they use to produce refined products (which Kilian calls “oil in use”) and the portion that goes into inventory (which Kilian calls “speculative demand”). Whether one wants to estimate a price elasticity separately for these two components of demand or is just interested in their sum as

\[ p(A) \propto p_1(\alpha)p_2(|A|^{-1}(-\psi - \beta \xi)) \]

as a composite prior for $A$. As discussed by Baumeister and Hamilton (2018), there is no problem with including multiple sources of information about the same parameter, just as there is no problem with using multiple earlier samples that all contain information about a common parameter to form a Bayesian prior in standard settings. The applications in Baumeister and Hamilton (2018, 2019), Grisse (2020), Valenti et al. (2020), Lukmanova and Rabitsch (2021), and Aruoba et al. (2021) all incorporate prior information about both $A$ and $A^{-1}$.

Kilian and Zhou (2020) asserted that “Baumeister and Hamilton’s approach is not designed to handle the restrictions on $[A^{-1}]$ typical of conventional oil market models, except in the special case of a recursively identified model.” The same claim was repeated in Kilian and Lütkepohl (2017, page 454). The statement is false. The method described by Baumeister and Hamilton (2018, 2019) for incorporating information about both $A$ and $A^{-1}$ is completely general.
assumed in equation (1) would seem to depend on the application. In any case, this issue is completely irrelevant to the points we are making in this paper. Whatever structural system one has in mind and whatever detail of individual behavioral equations it includes, an IV or full-information approach is necessary to estimate the structural parameters in any of the behavioral equations. One can not estimate any of those structural parameters from the ratios of equilibrium responses to a single structural shock.

Baumeister and Hamilton (2019) illustrated the correct way to do this. If $U_t$ denotes oil in use, $\Delta I_t$ the change in inventories, and $Q_t$ the quantity of oil purchased by refiners, then the magnitudes are related by the identity $Q_t = U_t + \Delta I_t$. It is convenient to keep the system in log-linear form. For $q_t = \log Q_t - \log Q_{t-1}$ the rate of growth of total oil use, Baumeister and Hamilton (2019, page 1888) achieved this by using the approximation

$$\frac{U_t}{U_{t-1}} = \frac{Q_t - \Delta I_t}{U_{t-1}} \simeq \frac{Q_t - \Delta I_t}{Q_{t-1}}$$

from which

$$\Delta \log U_t \simeq (U_t/U_{t-1}) - 1 \simeq \Delta \log Q_t - \Delta I_t/Q_{t-1} = q_t - \Delta I_t/Q_{t-1}.$$

Baumeister and Hamilton then represented the structural equation for use demand in their equation (29) as

$$q_t - \frac{\Delta I_t}{Q_t} = \delta y_t + \beta p_t + b^d x_{t-1} + u_t^d$$

in which $\delta$ is the short-run income elasticity and $\beta$ the short-run price elasticity of use demand. Baumeister and Hamilton drew on information from other studies to form Bayesian priors for $\delta$ and $\beta$ along with other structural parameters to find the posterior distribution of all the structural parameters, and hence the posterior distribution of any other object of interest such as $H$ or $H^*$.

5 Conclusion.

The ratio of elements of a single column of the impact matrix $H$ in a structural VAR cannot be used to estimate a parameter such as the price elasticity of oil demand. Those ratios instead summarize the equilibrium effects of a particular shock and in general are functions

\cite{Kilian2020}
evidently objects to the approximation in (17), claiming in his footnote 13 that our equation is only valid if $\Delta I_t = 0t$. The accuracy of the approximation depends not on the absolute size of $\Delta I_t$ but on its size relative to $U_t$. Between February 1990 and April 2021, U.S. refiners purchased an average of 454.4 million barrels of crude oil each month (data from http://www.eia.gov/dnav/pet/pet_pnp_wiup_dcu_nus_4.htm), while the average absolute value of the 4-week change in U.S. inventories of crude oil and refined products was 14.2 million barrels (http://www.eia.gov/dnav/pet/pet_stoc_wstk_dcu_nus_w.htm). The ratio between these is 14.2/454.4 = 0.03. The approximation in (17) is excellent.
of all the structural parameters of the system. There is a single unique function of $H$ that corresponds to the elasticity. In the example presented in Section 2, this is $h^{33}$, the $(3,3)$ element of $H^{-1}$. And there is an unambiguously optimal way to estimate this elasticity. For frequentist inference, this is the full-information maximum likelihood estimate of $h^{33}$. This will be the same estimate regardless of how the structural model is parameterized. For Bayesian inference, this is given by the value of $\tilde{h}^{33}$ that minimizes the expected posterior loss. This again will be the same estimate regardless of how the structural model is parameterized.

Whether the goal is applied research or policy guidance, there is a clear answer to the question of how to use structural vector autoregressions to estimate any magnitude of interest.
References


Appendix A: Derivation of equation (11).

Expression (10) can be rewritten

\[
\begin{bmatrix}
\sum_{t=1}^{T} \hat{u}_{t}^e \hat{y}_{t}^e & \sum_{t=1}^{T} \hat{u}_{t}^p \hat{y}_{t}^p \\
\sum_{t=1}^{T} \hat{y}_{t}^e \hat{z}_{t}^e & \sum_{t=1}^{T} \hat{y}_{t}^p \hat{z}_{t}^p
\end{bmatrix}
\begin{bmatrix}
\hat{\delta}_{IV} \\
\hat{\beta}_{IV}
\end{bmatrix}
= \begin{bmatrix}
\sum_{t=1}^{T} \hat{u}_{t}^e \hat{z}_{t}^e \\
\sum_{t=1}^{T} \hat{u}_{t}^p \hat{z}_{t}^p
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 - \gamma & -\alpha \\
-\xi & 1 - \psi
\end{bmatrix}
\begin{bmatrix}
\sum_{t=1}^{T} \hat{y}_{t}^e \hat{z}_{t}^e \\
\sum_{t=1}^{T} \hat{y}_{t}^p \hat{z}_{t}^p
\end{bmatrix}
\begin{bmatrix}
\hat{\delta}_{IV} \\
\hat{\beta}_{IV}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
1 - \gamma & -\alpha \\
-\xi & 1 - \psi
\end{bmatrix}
\begin{bmatrix}
\sum_{t=1}^{T} \hat{y}_{t}^e \hat{z}_{t}^e \\
\sum_{t=1}^{T} \hat{y}_{t}^p \hat{z}_{t}^p
\end{bmatrix}
\begin{bmatrix}
1 \\
-\hat{\delta}_{IV}
\end{bmatrix}
\begin{bmatrix}
0 \\
-\hat{\beta}_{IV}
\end{bmatrix}
\]

(18)

from which (11) follows from the definitions of \( \Gamma \), \( \eta \), and \( \hat{\Omega} \).