Handling Non-Invertibility: Theory and Applications*

Bill Dupor and Jing Han

March 8, 2011

Abstract

Existing research provides no systematic, limited information procedure for handling 6 non-invertibility, despite the well-known inference problem it causes as well as its presence in many types of dynamic systems. Non-invertibility means that structural shocks cannot be recovered from a history of observed variables. It can arise from q a form of delayed responses due to, among other things, time-to-plan, sticky infor-10 mation or news shocks. Structural VARs rule out non-invertibility by assumption. 11 Inference about structural responses can, in turn, be incorrect. We develop a practi-12 cal four-step procedure to partially, and sometimes fully, identify structural responses 13 whether or not non-invertibility is present. Our method combines structural VAR re-14 strictions, e.g. recursive identification, with "agnostic" identification, e.g. sign restric-15 tions and bounds on forecast error contributions. In two model-generated examples, 16 our procedure either fully or nearly fully identifies the structural responses whereas 17 SVARs do not. Also, we apply our procedure to real world data. We show that non-18 invertibility is unlikely in Fisher's (2006) study of technology shocks in the U.S. 19

²⁰ J.E.L. Classification: E3.

1

2

3

4

5

²¹ Keywords: State-space representation, vector-autoregression, non-invertibility.

^{*}Ohio State University and Chinese University of Hong Kong. We thank Karel Mertens, Jim Nason and Jesus Fernandez-Villaverde for helpful comments as well as participants at the 2010 NBER-NSF Time Series Conference and the Winter 2010 North American Econometric Society Meeting. We are especially grateful to Raymond Chan for helpful discussions on how to solve quadratic matrix equations.

²³ 1 Introduction

Suppose a police officer on foot patrol happens upon a dead man with a knife in his back. An autopsy firmly establishes that the time of death was 5:00 AM earlier that day. Detectives would like to know when he was stabbed. With no witnesses, the stabbing could have occurred at 4:59 AM with the victim dying quickly. Or, the stabbing could have occurred the previous evening with the victim dying slowly. There are other possibilities, and thus, the time of the crime is not identified.

A time series analyst often faces a similar problem. Suppose the analyst observes a 30 series of outcomes (e.g. real GDP), each of which is indexed by a known time. Suppose 31 the analyst does not observe the sequence of impulses (e.g. preference shocks) or their 32 associated times. A current change in an observable might be due to immediate response 33 to a contemporaneous impulse. Or, the current change might be a delayed response to 34 an impulse that occurred long ago. To the analyst, this is known as the *non-invertibility* 35 *identification problem*. It is distinct from the "simultaneous equation problem" that arises 36 because of multiple simultaneous unobserved shocks¹. 37

The police detective and the time series analyst have different standard operating pro-38 cedures for dealing with this identification problem. The police detective would look for 39 other evidence to inform when the shock (i.e. the stabbing) occurred, such as the stiffness 40 of the dead body. Faced with the same crime, on the other hand, the time series analyst 41 typically would assume that the stabbing occurred at 4:59, because this is the response 42 with the shortest delay from impulse to observable. In technical language, the analyst 43 has dealt with the non-invertibility problem by assuming the invertible representation, 44 i.e. the one with minimal delay, is the correct one. In non-technical terms, the analyst has 45 done shabby police work. 46

⁴⁷ In this paper, we develop a procedure for handling the identification problem with-

¹In most problems, one must cope with both equation simultaneity and non-invertibility. Handling both is a part of our paper.

out assuming that responses to structural shocks occur with minimal delay. Rather, we follow the police detective's method. We ask whether other evidence, including the comovement of the observable with other observables or the sign of impulse responses, are consistent or inconsistent with restrictions implied by economic theory. We wish to use as few clues given by economic theory as possible.

This paper addresses non-invertibility, also known as non-fundamentalness, in a limited information framework.² We treat non-invertibility in a similar manner to the one that researchers already use in SVARs to deal with the simultaneous equations identification problem. That is, compute all of the stochastic processes consistent with the data and then apply identifying restrictions from economic theory to exclude some (and potentially all but one) of these processes.

⁵⁹ Our procedure has four steps.

⁶⁰ **Step One:** *Estimate a reduced-form VARMA*(1,1) *on the observables.*

⁶¹ We begin by assuming the time series has a state-space representation. Under some ⁶² general assumptions discussed in later, the observables from a state-space representation ⁶³ can be written as a VARMA(1,1). Many dynamic economic models is consistent with this ⁶⁴ form. To be concrete, let Y_t represent a vector of k observable, stationary variables.

⁶⁵ **Step Two:** *Calculate all covariance equivalent representations.*

⁶⁶ With *k* observable variables, there are at most 2^k state-space forms that have the iden-⁶⁷ tical covariance functions, modulus the simultaneous equations problem. One of these ⁶⁸ state-space forms will be invertible, i.e. have minimal delay. However, there is no ratio-⁶⁹ nale for simply choosing this one over a non-invertible representation, without further ⁷⁰ identification restrictions in hand. As such, this step records and keep tracks of each one. ⁷¹ **Step Three:** *Define the structural shock of interest and impose an SVAR-type restriction on each* ⁷² *representation.*

This step mimics that of the SVAR approach. A structural shock is a primative of an
 economic model, such as an exogenous change in technology or monetary policy. The

²Throughout this paper, we use the term non-invertibility rather than the equivalent *non-fundamentalness*. Using the latter can generate confusion, since economists often refer to fundamental shocks as the economically meaningful shocks, such as changes in preferences or technology. Fundamental shocks in the time series sense are not necessarily fundamental in sense of economic theory.

restriction might concern the short run, e.g. output does not respond to current monetary
policy changes, or the long run, e.g. only technological change affects long-run labor
productivity. This step is needed because the simultaneous equations problem exists
apart from the non-invertibility issue.

Step Four: Impose agnostic restrictions on each representation, delivered from step three, to fur ther rule out potential structural responses.

⁸¹ Uhlig uses the phrase "agnostic restrictions" to describe identifying assumptions of ⁸² the kind implemented for example in Faust (1998), Scholl and Uhlig (2008) and Uhlig ⁸³ (2005).³ For example, a positive innovation to the structural shock might be required to: ⁸⁴ (i) have a non-negative long-run effect on a particular observable; (ii) imply a positive re-⁸⁵ sponse to an observable at the two-year horizon; (iii) explain the variation in one variable ⁸⁶ within a certain range. In contexts outside of non-invertibility, researchers have over the ⁸⁷ past several years found agnostic restrictions very useful.⁴

After step four, the researcher is left with one or multiple structural impulse responses to the structural shock of interest. When only one response remains, the impulse response is fully identified. When multiple remain, the impulse response is partially identified. In either case, the invertible form may or may not belong to the set. If the invertible form is consistent with the restrictions from step four, then it will be a valid structural response. Importantly, our procedure does not a priori choose this response.

The problem of non-invertibility has received great attention in economics and time series analysis. In an introductory chapter of his textbook, Hamilton (1994, pg. 64) discusses the issue and presents practical reasons for preferring the invertible representation.⁵ Sargent (1987) presents an early textbook discussion.⁶ Fernandez-Villaverde et al (2007, FRSW hereafter) explain that non-invertibility is induced by missing variables.

Economists have pointed out that non-invertibility can arise in many environments.

³Other work using agnostic identification include: Cardoso-Mendonca, Medrano and Sachsida (2008), Mountford and Uhlig (2009) and Owyang (2002).

⁴Fry and Pagan (2010) contains an extensive and critical survey of one type of agnostic restriction–the sign restriction.

⁵We discuss these reasons and how our method addresses them in section two.

⁶Other textbook presentations on the invertibility of MA processes include Brockwell and Davis (2009) and Lutkepohl (2010).

Model features that can induce non-invertibility in the structural responses include: per-100 manent income economies (Hansen and Sargent 1991, Hansen, Roberds and Sargent 1991 101 and FRSW 2007); learning-by-doing (Lippi and Reichlin 1993); anticipated fiscal policy 102 shocks (Leeper, Walker and Yang 2009); anticipated technology shocks (Blanchard et. al. 103 2009). Alessi, Barigozzi and Capasso (2011) surveys the prevalence of non-invertibility in 104 rational expectations models. Lippi and Reichlin (2003) discuss the possibility of misspec-105 ification due to non-invertibility in Blachard and Quah (1989). Sims (2009) is an exception 106 to the above studies. Using data simulated from a calibrated DSGE model, he finds that 107 non-invertibility, while present, introduces little bias in the impulse responses from a 108 structural VAR. 109

Despite these extensive discussions of the problem and its practical relevance, only three categories of solutions have been offered. These are: (i) adding more observables; (ii) using full information estimation of a correctly specified DSGE model; (iii) standard SVAR estimation augmented with something akin to our Step Three. Each differs from ours in separate and important ways.

First, one could expand the observables. Most directly, researchers can try to directly observe the structural shocks. If the shock and its arrival time are known, the identification problem disappears. Case studies applied to particular changes in tax policy are well-suited for this approach. Also, Romer and Romer (2004, 2010) have used the narrative approach to create time series measures of the values of actual monetary policy shocks and actual government spending shocks. However, in most cases, shocks are not directly observed.

Even when structural shocks are not observed, adding observables potentially eliminates non-invertibility. Alessi, Barigozzi and Capasso (2010) recommend using a large number of observables and then applying structural restrictions, e.g. a Choleski decomposition, to the estimated factor-augmented VAR. Forni, Giannone, Lippi and Reichlin (2009) advocate this approach by showing that moving from a structural VAR to a factoraugmented structural VAR changes the responses of output to permanent supply shocks.⁷ Second, FRSW (2007) draws upon their discussion of the danger in using SVARs.

⁷See also Giannone and Reichlin (2006).

SVARs always choose the invertible representation of a time series. When the actual 129 structural response is non-invertible, the SVAR leads to incorrect inference. Rather than 130 an SVAR, they recommend correctly specifying a full dynamic, stochastic general equi-131 librium (DSGE) model and using a full information technique. Our limited information 132 procedure is less likely to suffer from misspecification than using a fully specified model. 133 FRSW (2007) also provide a condition to use, case-by-case, to determine whether an 134 SVAR would generate incorrect inferences. To check this condition, one uses the estimates 135 or calibration of the DSGE model relevant for the particular time series. However, with 136 a correctly specified DSGE model in hand, one should use all of the information in the 137 DSGE model rather than the limited information SVAR on efficiency grounds. 138

In a somewhat-related way, Mertens and Ravn (2010) use DSGE models together with structural VARs in an inventive way, to address non-invertibility. They specify and calibrate a DSGE model with news shocks, and then use it to determine the placement of the non-invertibility in the system's moving-average structure, along with the magnitude of the roots associated with the non-invertibility. In their exercise, Mertens and Ravns preset the values of the roots associated with the non-invertibility.

Third, Lippi and Reichlin (1994) suggest a limited information approach. It is the closest antecedent of our work. They compute the structural impulse response using a VAR and a standard rotation restriction. The estimated structural response is by construction invertible, as discussed in FSRW. Recognizing that non-invertible solutions are also consistent with the observed data, they then do a visual inspection of roots from the estimated VAR in search of an MA structure.

Based on the inspection, they plot both non-invertible and invertible structural re-151 sponses implied by their VAR. This is similar to our step three. As they explain, their 152 method is only suitable for a two variable system. On the other hand, our procedure 153 works for a system with more variables because we estimate the MA component directly 154 (i.e. our step one). Also, our procedure allows us to exclude some of the potential struc-155 tural responses (i.e. our step four) in a systematic manner. Moreover, their procedure 156 can only analyze a single shock with non-invertibility, while our procedure is suitable for 157 cases with multiple non-invertible shocks. 158

Our procedure has three distinct benefits not shared by the other approaches: (i) it directly estimates the model's moving average component (i.e. Step One), which is the heart of identification issue; (ii) by using the quadratic matrix equation (i.e. Step Two), it quickly and intuitively finds the entire set of covariance equivalent stochastic processes; (iii) by using agnostic restrictions (i.e. Step Four), it stays within the limited information framework of structural VARs.

First, since the entire source of non-identification is the multiplicity of moving average components of an observed covariance function, it makes sense to estimate the moving average component directly. At the same time, an autoregressive part may also be present. As such, we use a VARMA model to capture both parts. Lippi and Reichlin (1994), in contrast, estimate a VAR and then do a visual inspection for MA roots. This limits the applicability of their procedure as discussed above.

In the past, researchers have avoided estimating moving average models with good reason. There is a relatively old (circa the 1970s) concern that implementing a VARMA is so difficult as to make their use infeasible. The erstwhile approach centered on nonlinear maximization of a likelihood function over a high dimensional parameter space. While possible in theory, it can be unreliable practically.

Numerous recent advances in VARMA estimation largely ameliorate this concern. Du-176 four and Pelletier (2008) for example extend to the vector case the innovation-substitution 177 method developed by Hannan and Rissanen (1982). The method involves feeding the 178 residuals from a long-lag AR as the innovations in the estimation of an ARMA model. 179 OTHER METHODS: Koreisha and Pukkila GLS (1990), Larimore CCA subspace (1983) 180 and Kapetanios iterative LS (2003), Hannan and Kavalieris 3SLS (1984). We use Dufour 181 and Pelletier's method in all of our examples. Kascha (2007) compares the above meth-182 ods using a well-known macro application and shows that the innovation-substitution 183 method dominates. 184

Second, we compute the entire set of structural representations using a simple formula
 (Potter 1966) that solves a quadratic matrix equation. We set out to develop a procedure
 is easy for practitioners to use. The Potter equation is easy to code and fast to run. It
 requires only a single matrix inversion and a single eigenvalue decomposition.

An alternative technique, Blaschke factorization, can in principle do the same job. It appears in many theoretical discussions about non-invertibility;⁸ however, to our knowledge, it has never been used in applications. Perhaps this is because it is much more involved from a practical standpoint. It begins with a single eigenvalue computation that is then followed by a large number of "root flipping" steps, where each root flipping requires the calculation of the null space of a particular matrix.

Third, our paper maintains the limited information spirit of BLAH BLAH BLAH.FINISH
 THIS.

To set the stage, the next section contains a bivariate process where non-invertibility is present. Section 3 presents the four-step procedure along with its theoretical substructure. Section 4 applies the procedure to two sets of model-generated data and section 5 applies the procedure to a real world application. Section 6 concludes.

201 2 Non-invertibility in A Bivariate Example

We illustrate the nature of non-invertibility using a two variable example.⁹ Suppose an economist observes y_{1t} and y_{2t} . For concreteness, call them the money growth rate and real output. Each variable has expectation zero and an own first-order autocorrelation equal to 0.01. At further lags, each has a zero autocorrelation. The two are uncorrelated with each other at every horizon. Also, suppose there are two shocks driving the system, which, for concreteness, are technology shocks and monetary policy shocks.

What VMA(1) processes are consistent with the above covariance structure? Indexing each process by *j*, these are

$$y_t = \Gamma_0^j \omega_t^j + \Gamma_1^j \omega_{t-1}^j$$

where Γ_0^j and Γ_1^j are square matrices of dimension two . The number of processes, or forms, modulus the simultaneous equations issue, is at most 2^k . Since k = 2, there are up to four forms. Figure 1 plots the impulse responses for three of these. We omit the

⁸These include Whiteman (1983), Hansen and Sargent (1991), Lippi and Reichlin (1994), Leeper, Walker and Yang (2009) and Alessi, Barigozzi and Capasso (2010).

⁹Examples using one variable are presented in Hamilton (1994) and Sargent (1987). While instructive, the scalar case cannot elucidate the important cross-covariagram implications of non-invertibility.

fourth to avoid clutter. Each row corresponds to a moving-average form and each column
 corresponds to a particular shock applied to a particular variable.

To deal with the simultaneity of shocks, we have imposed a short-run restriction that output does not respond contemporaneously to the monetary shock. In the figure, the period zero response of output to the monetary shock is zero in each panel of the second column of the figure. Suppose this short-run restriction holds in the underlying structural model.

Suppose that the true structural model, or economy, that delivers the observed covari-218 ance matrices is in the first row of the figure. This economy corresponds to one of the 219 non-invertible forms. The economy has three key features: a money growth shock is not 220 neutral (see panel (b)), monetary policy responds counter-cyclically to technology shocks 221 (see panel (c)), and there is a large "news component" to money growth shocks (see panel 222 (d)). The news interpretation of panel (d) is appropriate because, although the money 223 growth shock arrives at time zero, the most substantial increase in the money supply 224 happens at time one. An economist that observes y_t , but does not observe either shock, 225 may try to identify the shocks using a structural VAR, which automatically chooses the 226 invertible form. Suppose the economist knows that the above short-run restriction is true 227 for this economy. If the economist runs an SVAR using the restriction, she will estimate 228 the second row of Figure 1. This is the invertible form. This economist would come away 229 incorrectly believing that money shocks are neutral (see panel (f)) and monetary growth 230 does not respond to technology shocks (see panel (g)). 231

What is going on? There is a 'covariance accounting' requirement that is satisfied for the various forms. Each form has sets of moving average coefficients that line up in a way that the corresponding second moments across forms are identical. In the next section, we provide a simple equation to construct all forms that satisfy the covariance requirement.

Armed with only the short-run restriction, the structural model is not identified. Even worse, an SVAR with only the short-run restriction will estimate the wrong model. The estimated model says money is neutral with respect to output when in reality it is not! How can one deal with this under-identification? Our solution is to bring more *a*

240 priori knowledge about the economic environment to the table. The goal should be to



Figure 1: Three covariance-equivalent stochastic processes

Notes: The fourth and final form, another non-invertible process, is not pictured above.

²⁴¹ bring restrictions that are agnostic, in the sense of Uhlig (2005), as possible to reduce the ²⁴² set of valid forms. An alternative approach, advocated by FRSW (2007) and discussed ²⁴³ in our introduction, is to bring a lot to the table, in the form of a fully-specified dynamic ²⁴⁴ general equilibrium model. As we explained in the introduction, the dynamic general ²⁴⁵ equilibrium approach goes against the spirit of the limited information technique and ²⁴⁶ moreover eliminates the need for limited information anyways.

²⁴⁷ 3 Theory and A Four-Step Procedure

²⁴⁸ A generic covariance-stationary stochastic process is given by:

$$s_{t+1} = Qs_t + Ue_{t+1}$$
 (1)
 $r_{t+1} = Ws_t + Ze_{t+1}$

where e_{t+1} is k by 1 and N(0, I). We refer to (Q, U, W, Z) as a *state-space form* (with associated shock process e_t) for the stochastic process $\{s_t, r_t\}$. Here, Q, U, W, Z are real-valued. Only r_t is observed by the economist. Also, we make the following assumptions.

Assumption 1: The left inverse of W, which we denote \overline{W} , exists.

Assumption 2: All eigenvalues of Q and $WQ\overline{W}$ are inside the unit circle.

Assumption 3: The matrix *Z* is invertible.

Assumption 1 requires that there are least as many observables as states. To identify the underlying system, economists need to have enough information, i.e enough observable variables. This assumption is not as restrictive as it may seem. If the economy is actually driven by a few common factors, e.g. the dynamic factors as those identified by Stock and Watson (2002) or used by Bernanke, Boivin and Eliasz (2005), most multivariate time series models have more observables than states.

Assumption 2 ensures the stationarity of observables. In our exercise, we rule out cases with non-stationary variables. However, it is straightforward to covert non-stationary variables to stationary ones by detrending them or choosing correct cointegration vectors. Our procedure then is ready to go.

Assumption 3 requires there are at least as many observables as structure shocks of concern. This assumption is for technical purposes and not restrictive, since we can add include measurement errors as structural shocks. FRSW (2007) also make this assumption.

In lieu of additional information, the time series analyst knows or can estimate the covariance generating function of the observables. Let this covariance structure be denoted $C_i = E(r_t r'_{t-i})$ for all *i*.

272

To understand the theory that follows as we as our procedure, it is useful to compute

²⁷³ these covariances as functions of the underlying structural form:

$$C_{0} = WQ\bar{W}C_{0} (WQ\bar{W})' + ZZ' + WUU'W'$$
$$-WQ\bar{W}C_{0} (WQ\bar{W})'$$
$$C_{1} = WQ\bar{W}C_{0} + WUZ' - WQ\bar{W}ZZ'$$
$$C_{i} = (WQ\bar{W})^{i-1}C_{1} \text{ for all } i > 1$$

In the theorem that follows, we find the number of matrix triples $\{A_j, B_j, D_j\}$ corresponding to covariance equivalent forms and also show how to conveniently compute each of them.

Moving from the structural form to an observationally equivalent one changes the amount of delay in the system, as seen in Section 2. Intuitively, this can be seen in the state space system by examining the MA representation of the original structural system. This MA representation is:

$$r_{t+1} = Ze_{t+1} + W\sum_{i=0}^{\infty} Q^i Ue_{t-i}$$

Because the original and observational equivalent state-space forms differ in terms of U 277 and Z, the corresponding impulse responses will differ in magnitude of a shock's in-278 stantaneous effect, i.e. e_{t+1} , versus its lagged effect, $e_t, e_{t-1,...}$. Moreover, as seen in the 279 bivariate example of section 2, changing the delay in the response of one variable to a 280 shock has implications for all of the other impulse responses because of the known co-281 variance structure of the observables. The theorem below formalize the relation between 282 the structural form and its covariance-equivalent cousins. Furthermore, it lays out the 283 theoretical foundation for the practical procedure we use to tackle non-invertibilities. 284

Theorem 1: If r_t is a length k stochastic process with the structural state-space form (1) and assumptions 1 through 3 are satisfied, then there exists at most 2^k infinite-order covariance equivalent moving average representations for $\{r_t\}$, indexed by j, where the ²⁸⁸ innovations process ε_t^j satisfies $E\left(\varepsilon_t^j \varepsilon_t^{j\prime}\right) = I_k$. Representation j is given by

$$r_{t+1} = (I - AL)^{-1} [D_j + B_j L] \varepsilon_{t+1'}^j$$
(2)

The coefficient matrices, A, B_i and \tilde{C}_i , i = 0, 1 are:

$$\begin{array}{rcl}
A &= & C_2 C_1^{-1} \\
\tilde{C}_1 &= & C_1 - A C_0 \\
\tilde{C}_0 &= & C_0 - A C_0 A' - A \tilde{C}'_1 - \tilde{C}_1 A' \\
B_j &= & \tilde{C}_1 (D'_j)^{-1}
\end{array}$$
(3)

where C_i is the *i*th order autocovariance of the observable vector. The matrix, D_j , satisfies: (i)

$$(D_j D'_j) (\tilde{C}'_1)^{-1} (D_j D'_j) - \tilde{C}_0 (\tilde{C}'_1)^{-1} (D_j D'_j) + \tilde{C}_1 = 0,$$
(4)

(ii) $D_j = D_j^c K_j$, where D_j^c is the lower triangular matrix generated by the Cholesky decomposition of $D_j D'_j$. The orthonormal matrix, K_j is determined by the relation between the cholesky decomposition and the identifying restriction. When we use the short-run restriction, $K_j \equiv I$. If we use the long-run restriction, K_j differs from each other.

(iii) one of the D_j s is invertible and the corresponding MA form matches the Wold representation for r_t .

This theorem tells us: (i) a time series can have multiple representations; (ii) all of these forms can be backed out from a single reduced-form estimation. This multiplicity of covariance equivalent forms is one source of an identification problem with VARs.¹⁰ Equation (4) provides a way to find all of these forms. Hence, it allows us to dramatically reduce the dimension of the identification problem.

As an aside, note that this identification difficulty is not specific to structural VARs. The difficulty can also apply if a full information method, such as maximum likelihood, is used instead. This is because the covariance equivalence of the various forms implies multiple peaks in the likelihood function.

¹⁰The other source is the well-known simultaneous equations problem.

From Theorem 1, we develop our four-step procedure. In sections 4 and 5, we use model-generated data and real-world data to demonstrate the procedure.

³⁰⁹ **Step One**: *Estimate a reduced-form VARMA*(1,1) *model on the observables*.

³¹⁰ Under Assumptions 1 through 3, the structural model has a unique invertible VARMA(1,1) ³¹¹ form. It can be consistently estimated with traditional methods.

Step Two: *Calculate all covariance equivalent representations.*

Under the same assumptions, the true model can have multiple non-invertible VARMA(1,1) 313 forms in addition to the one invertible form. Each corresponds to a solution of a quadratic 314 matrix equation. All can be found simultaneously using the Potter (1964) equation. This 315 computation is simpler than the existing Blaschke method, as discussed in the introduc-316 tion. Although the number of forms at this step can theoretically large,¹¹, this issue is 317 mitigated in practice. As seen in the following two sections, (i) impulse response from a 318 subset or subsets of forms is often 'clustered,' making them quantitatively indistinguish-319 able; (ii) solutions with imaginary components are thrown out. 320

Step Three: Define the structural shock of interest and impose an SVAR-type restriction on each
 representation.

³²³ When the dimension of the observable variables is k, there are at most 2^k solutions for ³²⁴ fully specified rotation matrices. There is always at least one solution—the invertible one. ³²⁵ **Step Four**: *Impose agnostic restrictions on each representation, delivered from step three, to rule* ³²⁶ *out further structural representations.*

Usually there are multiple solutions after step three. More restrictions other than those on the pattern on the rotation matrix help reduce the set of covariance-equivalent forms. If only one solution remain, the structural model is fully identified, otherwise, the model is only partially identified.

331 4 Two Model-Based Implementations of Our Procedure

In this section, we use two model-generated examples to illustrate how our procedure identifies the structural model when the structural VAR cannot. The first example is

¹¹For example, with eight observables there are potentially 256 covariance equivalent forms.

³³⁴ adopted from the permanent income economy in FRSW (2007). The second example
³³⁵ is from a model of anticipated tax shocks (i.e. "news" regarding the future tax rate) in
³³⁶ Leeper, Walker and Yang (2009).

³³⁷ 4.1 Savings and Permanent Income in FRSW (2007)

The permanent income model is a workhorse of modern economics. FRSW (2007) show how applying structural VAR analysis to data from a permanent income model leads to an incorrect conclusion about the consumption response to an income shock. The incorrect conclusion occurs because the procedure fails to handle an inherent non-invertibility. We show how our procedure leads to the correct conclusion.

³⁴³ The economic model has two equations.

$$c_{t+1} = \beta c_t + \sigma_w (1 - R^{-1}) w_{t+1}, \tag{5}$$

$$z_{t+1} = y_{t+1} - c_{t+1} = -c_t + \sigma_w R^{-1} w_t,$$
(6)

Equation (5) is the intertemporal Euler equation and equation (6) defines saving. In the model, c_t is the unobserved state, while $z_t = y_t - c_t$ is saving, the only observable in the model. This process is non-invertible, since $Q - UZ^{-1}W = \beta + R - 1 > 1$, when β is close enough to one. The ARMA(1,1) representation of the observable is given by:

$$z_{t+1} = \beta z_t + \sigma_w R^{-1} w_{t+1} - \sigma_w [1 - R^{-1} + \beta R^{-1}] w_t,$$
(7)

³⁴⁸ which is non-invertible. The innovation representation is:

$$\hat{c}_{t+1} = \beta \hat{c}_t + \sigma_w (\frac{\beta - \beta^2 + 1}{R} - \beta) \epsilon_{t+1}$$
(8)

$$z_{t+1} = -\hat{c}_t + \sigma_w(\frac{\beta - 1 + R}{R})\epsilon_{t+1}.$$
(9)

Straightforwardly, the ARMA(1,1) model corresponding to the innovation representation
 is:

$$z_{t+1} = \beta z_t + \sigma_w \left(\frac{\beta - 1 + R}{R}\right) \epsilon_{t+1} - \frac{\sigma_w}{R} \epsilon_t.$$
(10)

The innovation representation is invertible, since $\hat{Q} - \hat{U}\hat{Z}^{-1}\hat{W}' = \frac{1}{R+\beta-1} \in (0,1)$. However, since the implied state variable is not the true state variable, i.e, $\hat{c}_t = E(c_t|z^t) \neq c_t$, where z^t refers to the history of the observable saving, z_t ; therefore, FRSW (2007) warn that inference based on the (estimated) innovation representation is not reliable.

Suppose the economist observes a time series for savings, z_t , however, she is uninformed regarding consumption and income. She would apply our procedure as follows:

Step One: *Estimate a reduced-form ARMA*(1,1) *on the observable.*

Step Two: Calculate all covariance-equivalent representations.

With only one observable variable, there are only two covariance equivalent MA(1) representations.

Step Three: Define the structural shock of interest and impose an SVAR-type restriction on each
 representation.

Define a positive savings shock as a disturbance that increases savings in the period of the shock. Different researchers may have different interpretations as to what exogenous factors drive savings changes, such as shocks to permanent income, transitory income or preferences. With a scalar observable and a scalar shock, there is no simultaneous equations problem. As such, an SVAR-type restriction is unnecessary.

Before imposing Step Four, we plot the impulse responses that come out of Step Three. These areappear in Figure 2 both the saving level rate and the consumption. The solid and dashed lines are, respectively, the non-invertible and invertible responses. Both of these impulse response functions give the same population moments as those from (7) or (10). The non-invertible response is the true response and the invertible representation is spurious. As FRSW (2007) explain, a structural VAR always selects the invertible representation; therefore, it would lead to the incorrect impulse responses.

Step Four: Impose an agnostic restriction on each representation, delivered from Step Three, to
 rule out further potential structural responses..

Rather than a priori select the invertible form, we impose an agnostic restriction based on economic theory. We will impose the standard idea that people save now in order to consume more later. Formally, we require that: *if savings is non-zero in at least one period*, ³⁸⁰ then it must switch signs at least once.

Examining Figure 2, only the invertible response satisfies the agnostic restriction. After Step Four, we have a single structural impulse response, which is the true response from the economic model. It is exactly the structural model's impulse response.

Figure 2: Covariance-equivalent impulse responses to a positive savings shock



Notes: From the permanent income model with r = 0.2. Impulse responses to a one unit shock from step three and before application of step four.

In a wide class of models, an individual increases current savings in order to finance greater future consumption. The use of agnostic restrictions is, in our view, very powerful exactly because it implies transparency regarding the source of identification.

³⁸⁷ 4.2 An Anticipated Fiscal Shock in Leeper, Walker and Yang (2009)

The second model-generated example is based on Leeper, Walker and Yang (2009, LWY hereafter). This example has an anticipated fiscal shock: changes in the tax rate are an nounced two quarters before their implementation.

Consider a neoclassical model with fixed labor supply and full capital depreciation. The capital stock k_t is the single endogenous state variable. In equilibrium, it satisfies

$$(1 - \alpha L)(1 - \theta L^{-1})k_t = -\frac{\tau}{1 - \tau} E_t (\tau_{t+1}) + a_t - \theta E_t (a_{t+1})$$

³⁹¹ where every variable is the log deviation from its steady-state value. The variables τ_t and ³⁹² a_t are the tax rate and technology level. LWY assume there is a random component to the tax rate, which is announced two periods before the tax implementation. This news is denoted by $\epsilon_{\tau,t}$. The equilibrium law of motion for capital, consumption c_t and output y_t are:

$$k_{t+1} = \alpha k_t + a_{t+1} - \frac{\tau}{1-\tau} (1-\theta) [\theta \epsilon_{\tau,t+1} + \epsilon_{\tau,t}], \qquad (11)$$

$$c_{t+1} = \alpha k_t + a_{t+1} + \frac{\tau}{1-\tau} \theta[\theta \epsilon_{\tau,t+1} + \epsilon_{\tau,t}], \qquad (12)$$

$$y_{t+1} = \alpha k_t + a_{t+1}. (13)$$

³⁹⁶ LWY show that non-invertibility affects not only the identification of fiscal shocks, ³⁹⁷ but also the identification of the technology shock. They assume that the tax rate has ³⁹⁸ both the above anticipated random component as well as a contemporaneous response to ³⁹⁹ technology. The tax rate is $\tau_t = \psi a_t + \epsilon_{\tau,t-2}$.

LWY demonstrate the non-invertibility problem using a structural VAR where τ_t and k_t observed. In this case, the shocks identified by the structural VAR are not the true shocks, but rather combinations of the technology and tax/news shocks.

Our four-step procedure can identify, at least partially, the structural shocks in the model. It is applied step-by-step below. We require having enough observable variables, hence, we augment the observable space with consumption, c_t and the shocks with u_t , a measurement error on consumption. The addition of consumption does not remove the non-invertibility.

⁴⁰⁸ The state-space representation is:

s_{t+1}	_		Q		s _t	_		<u>u</u>		e_{t+1}
$\begin{bmatrix} k_{t+1} \end{bmatrix}$	Í	- α	$-rac{ au(1- heta)}{1- au}$	0	$\begin{bmatrix} k_t \end{bmatrix}$	ĺ	1	$-rac{ au heta(1- heta)}{1- au}$	0	$\begin{bmatrix} a_{t+1} \end{bmatrix}$
$\epsilon_{ au,t+1}$	=	0	0	0	$\epsilon_{\tau,t}$	+	0	1	0	$\epsilon_{\tau,t+1}$ (14)
$\epsilon_{\tau,t}$		0	1	0	$\left[\epsilon_{ au,t-1} \right]$		0	0	0	$\begin{bmatrix} u_{t+1} \end{bmatrix}$
r_{t+1}			W		\sim	_		Z		e_{t+1}
$\begin{bmatrix} \tau_{t+1} \end{bmatrix}$		0	0	1	$\begin{bmatrix} k_t \end{bmatrix}$	Í	ψ	0	0	$\begin{bmatrix} a_{t+1} \end{bmatrix}$
k_{t+1}	=	α	$-\tfrac{\tau(1-\theta)}{1-\tau}$	0	$\epsilon_{ au,t}$	+	1	$-rac{ au heta(1- heta)}{1- au}$	0	$\epsilon_{ au,t+1}$
$\begin{bmatrix} c_{t+1} \end{bmatrix}$		α	$\frac{\tau\theta}{1-\tau}$	0	$\left[\epsilon_{ au,t-1} \right]$		1	$rac{ au heta^2}{1 - au}$	1	$\begin{bmatrix} u_{t+1} \end{bmatrix}$

Our analysis requires setting values for the parameters. We follow LWY for most parameters.¹² In addition, the fiscal shock has unit standard deviation and $\sigma_a = 0.1$, The standard deviation of the measurement error is 0.05.¹³

⁴¹² By checking the "poor man's invertibility condition" from FRSW (2007), we see that ⁴¹³ the system is non-invertible. This is because the matrix $Q - UZ^{-1}W$ has eigenvalues ⁴¹⁴ outside the unit circle for our parameterization. The three eigenvalues of $Q - UZ^{-1}W$ are ⁴¹⁵ .33, -8.98 and -0.45; therefore, there is one dimension of non-invertibility.

The structural VAR approach ignores the embedded non-invertibility. On the other hand, our procedure takes all possible non-invertibilities into consideration.

⁴¹⁸ **Step One:** *Estimate a reduced-form VARMA*(1,1) *on the observables.*

Denote the VARMA(1,1) representation of the structural model as $r_{t+1} = \overbrace{WQW}^{A} r_t + \overbrace{Z}^{D} e_{t+1} + \overbrace{(WU - WQWZ)}^{B} e_t$ with the following matrices:

$$A = \begin{bmatrix} 0 & \frac{(\tau-1)}{\tau} & \frac{(1-\tau)}{\tau} \\ 0 & \alpha & 0 \\ 0 & \alpha & 0 \end{bmatrix} , D = \begin{bmatrix} \psi \sigma_a & 0 & 0 \\ \sigma_a & \frac{\tau\theta(\theta-1)}{1-\tau} & 0 \\ \sigma_a & \frac{\tau\theta^2}{1-\theta} & \sigma_u \end{bmatrix} , B = \begin{bmatrix} 0 & \theta & \frac{(1-\tau)}{\tau} \sigma_u \\ 0 & \frac{\tau(1-\theta)}{\tau-1} & 0 \\ 0 & \frac{\tau\theta}{1-\theta} & 0 \end{bmatrix}$$

419

The traditional structural VAR approach can only give the innovation representation, $r_{t+1} = Ar_t + \hat{D}\hat{e}_{t+1} + \hat{B}\hat{e}_t$, of the true model. The AR coefficient matrix, *A* is consistently identified, but \hat{D} and \hat{B} are biased. In our numerical example, the true VARMA(1,1) representation is:

	0 -3 3		.12	0	0		0	27	15	
A =	0.360	, D =	.12	.065	0	, B =	0	.24	0	.
	0.360		.12	024	.05		0	.89	0	

¹²We choose $\alpha = .36$, $\beta = .99$, $\tau = .25$.

¹³The size of technology shock is set up to allow the contribution of technology shocks and tax shocks on the variance of consumption is equalized in the long run. This parameterization is purely for analytical simplicity, and it does not affect the result qualitatively

The estimated innovation representation, on the other hand, is:¹⁴

$$A = \begin{bmatrix} 0 & -3 & 3 \\ 0 & .36 & 0 \\ 0 & .36 & 0 \end{bmatrix}, \quad \hat{D} = \begin{bmatrix} .29 & 0 & 0 \\ .21 & .14 & 0 \\ -.01 & -.01 & .15 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0 & -.12 & -.08 \\ 0 & .13 & .03 \\ 0 & -.04 & 0.01 \end{bmatrix}$$

In the true VARMA(1,1) representation, there are eigenvalues of BD^{-1} outside the unit circle, while every eigenvalue of $\hat{B}\hat{D}^{-1}$ in the innovation representation is inside the unit circle.¹⁵

423 **Step Two:** Calculate all covariance equivalent representations

This step finds all the representations with the same autocovariance structure. Each covariance equivalent form has an associated triple $\{A_j, D_j, B_j\}$. It is easy to verify that A_{26} $A_j = A$ and every pair of $\{D_j, B_j\}$ satisfies the following equations:

$$D_{j}D'_{j} + B_{j}B'_{j} =$$

$$\begin{bmatrix} \psi^{2}\sigma_{a}^{2} + \theta^{2} + \left(\frac{\sigma_{u}}{\kappa}\right)^{2} & \psi\sigma_{a}^{2} + \kappa\theta(1-\theta) & \psi\sigma_{a}^{2} - \kappa\theta^{2} \\ \psi\sigma_{a}^{2} + \kappa\theta(1-\theta) & \sigma_{a}^{2} + \kappa^{2}(1+\theta^{2})(1-\theta)^{2} & \sigma_{a}^{2} - \kappa^{2}\theta(1-\theta)(1+\theta^{2}) \\ \psi\sigma_{a}^{2} - \kappa\theta^{2} & \sigma_{a}^{2} - \kappa^{2}\theta(1-\theta)(1+\theta^{2}) & \sigma_{a}^{2} + \kappa^{2}\theta^{2}(1+\theta^{2}) + \sigma_{u}^{2} \end{bmatrix}$$

$$B_{j}D'_{j} = \begin{bmatrix} 0 & \kappa\theta^{2}(1-\theta) & -\kappa\theta^{3} - \frac{\sigma_{u}^{2}}{\kappa} \\ 0 & \kappa^{2}\theta(1-\theta)^{2} & -\kappa^{2}\theta^{2}(1-\theta) \\ 0 & -\kappa^{2}\theta^{2}(1-\theta) & \kappa^{2}\theta^{3} \end{bmatrix},$$
(15)

where $\kappa = \tau / (1 - \tau)$. The equation system (16) can be equivalently converted into a quadratic matrix equation in $D_j D'_j$. The solution of this quadratic matrix equation is given in Potter (1964). Since $D_j D'_j$ is a 3 × 3 matrix for each *j*, there are at most $2^3 = 8$ different lower triangular matrices solving the quadratic matrix equation. Under this current parameterization, $D_j D'_j$ has one pair of complex eigenvalues. As such, there are only four sets of real-valued structural responses.

433 **Step Three:** *Define the structural shock of interest and impose an SVAR-type restriction on each*

¹⁴Here we only show the result after imposing a short run restriction.

¹⁵The true model has two eigenvalues outside the unit circle, which are complex conjugates of each other.

434 representation.

A positive technology shock is defined as a shock which increases consumption and 435 does not reduce the tax rate. Consumption increases because of the positive effect of 436 technology shocks on production capacity. Obviously, a positive tax shock increases the 437 tax rate as well; however, the way it affects capital and consumption is not clear. One 438 possible way to separate the positive tax shock from the positive technology shock is by 439 assuming that an anticipated tax rate change cannot changes the current tax rate. Since 440 we know that measurement error only affects the measurement of consumption, it should 441 not affect the tax rate or capital on impact. Based on the definitions, we can impose the 442 following short-run restriction: a valid *D* matrix should be lower triangular. 443

Figure 3 plots the impulse responses to a positive tax shock (upper panel) and those to 444 a positive technology shock (lower panel) in all the four possible cases after imposing the 445 short-run restriction. One of them overlaps with the VAR-based inference, which is the 446 (invertible) innovation representation of the model. In response to a positive tax shocks, 447 capital and output falls in all four possible cases and tax rate increases in all of them. The 448 only difference is the magnitude of responses. When studying the responses to a positive 449 technology shock, capital falls in two cases but rises in other two. Output falls in the 450 innovation representation but rises in all the other three cases. The fall in output seems to 451 contradict traditional wisdom, however, there are evidences in existing research to show 452 technology shocks are contractionary. At this stage, we cannot rule out any of the four 453 cases without further justification. 454

455

Step Four: Impose agnostic restrictions on each representation, delivered from step three, to fur ther rule out structural responses.

Two agnostic restrictions are imposed. Both are based on the short-term forecast error variance decomposition. In order to identify the true impulse responses, we employ multiple criteria based on reasonable economic intuition. First, measurement errors should not be important factors to explain volatilities in any of the variables, especially in the longer term. Therefore, we setup a quantitative threshold of 30% for the average contribution of measurement errors on all observable variables (*criterion one*). Second, technol-



Figure 3: Response To Tax and Technology Shocks (after step three)

Notes: upper panel responses to a positive tax shock; lower panels responses to a positive technology shock. *PS i*: the *i*th solution based on the Potter equation.

	Model One	Model Two	Model Three	Model Four
The average contra	ibutions on diffe	erent horizons o	f identified measure	ement errors on variables
tax rate	0	34.82	0	14.78
capital	0	39.32	0	0.51
consumption	7.84	39.45	7.84	70.51
-				
The ave	rage contributio	ons of technolog	y on tax rate at diff	erent horizons
	1.42	35.05	1.42	53.24
The contrib	oution of techno	logy shocks on a	capital and consum	vtion when $h = 1$
capital	0	37.55	79.11	71.01
consumption	0	48.01	83.23	0.09
_				

Table 1: Identification Based on Short-Term Variance Decomposition

ogy shocks should not be the dominant factor to explain the volatilities in the tax rate,
especially in longer time horizons. Quantitatively, we set up the threshold value to be
50% when the time horizon is longer than two quarters (*criterion two*). The result of
this variance decomposition exercise is shown in Table 1.

Based on criterion one, case 2 and case 4 are ruled out, since these two cases attribute 468 too much variation to measurement errors. In this model, case 4 corresponds to the inno-469 vation representation, in other words, the model identified with traditional SVAR meth-470 ods. This specification can be ruled out based on our second criterion as well, since tech-471 nology shocks should not be the main driving force for tax rates. The reason why we can 472 use variance decompositions to identify the correct model is that covariance-equivalent 473 representations other than true models are likely to mix different shock together. There-474 fore, the variance decomposition is distorted in those representations. Leeper et al (2009) 475 makes a similar point from a different perspective. They view this as a failure of iden-476 tification with traditional SVAR methods. Our procedure goes one step further: some 477 mis-identification will give wildly implausible variance decomposition. Therefore, we 478 can rule out such mis-identified models. Such identification scheme share the same spirit 479

as the identification methods proposed by Faust (1997) and Uhlig (2005). As long as economic theory gives us enough restrictions on the model, e.g, the variance decomposition,
the sign of impulse responses or the sign of magnitude of a particular coefficient, we can
always apply them to rule out mis-identified models.

However, we still cannot achieve full identification here. As shown in Table 1, we cannot choose between case one and case three based on the first two criteria we proposed. Until this step, we achieve partial identification. Figure 4 compares the impulse responses implied by the remaining solutions to those of the true model and by the innovation representation. Both solutions recover the true responses to a positive tax shock in the structural model. One of them (the *"identified model"*) recovers the true responses to technology shocks as well.

In this example, we cannot uniquely pin down the true model. The reason is that the 491 first solution based on our procedure only mis-specifies the timing or invertibility of the 492 technology shock, but it does disentangle tax shocks and technology shocks effectively. To 493 further refine the result, we require more restrictions. For instance, if we have a strong be-494 lief that the transmission of technology shocks is fast enough, then the technology shock 495 should explain the bulk of changes in capital and consumption in the short term. Hence, 496 we might add a third agnostic restriction: the contribution of technology shocks to the 497 one step forecast error variances in consumption and capital should be higher than 30%. 498 With this extra restriction, we uniquely pin down the model as shown in Table 1. In the 499 true model, capital and output fall in response to an anticipated tax shock. Consumption 500 rises on impact but falls in following period. The initial rise is due to the substitution 501 effect induced by higher tax rate in the future while the following decrease is because of 502 the drop in production capacity. 503

When the model is identified correctly, capital, output and consumption all rise in response to a positive technology shock, while the innovation representation shows capital and output falls in response to it. Adding this third criterion, the true model is uniquely identified. From our perspective, criteria three is too strong to be used. Here, our procedure is not a slam dunk.



Figure 4: Response To Tax and Technology Shocks (after step four)

Notes: upper panel responses to a positive tax shock; lower panels responses to a positive technology shock

⁵⁰⁹ 5 Application with Real Data: Technology Shocks in the ⁵¹⁰ U.S.

Fisher (2006) uses a three-variable model to study the effect of technology shocks on the U.S. economy in the second half of the twentieth century. In his exercise, the investmentspecific shock, which is captured by surprise changes in the relative price of investment, is important to explain the variation in output and working hours in U.S.

Recently, studies on the effect of "news shocks", which is the anticipated component 515 in technology shocks, have drawn more and more attentions of economists, since the sem-516 inal work by Beaudry and Portier (2006). They show that technology shocks identified by 517 traditional long run restrictions can be well replicated by another shock originated in the 518 stock index but are orthogonal to contemporaneous technology changes. They argue that 519 this piece of evidence shows technology shocks are anticipated ("news shocks") and they 520 further show this news shock is important to explain business fluctuations. Jaimovich and 521 Rebelo (2009) show that certain real frictions, including habit persistence in consumption, 522 investment adjustment costs and costly capacity utilization, are important to the propa-523 gation of news shocks in a real business cycle model. Christiano et al (2009) estimate a 524 dynamic general equilibrium model featuring nominal and real frictions for the U.S. econ-525 omy and show that news shocks are important sources of business fluctuations. However, 526 Sims (2009) uses traditional SVAR methods to identify news shocks in a large scale VAR 527 model and finds that news shocks fail to generate co-movement in macro variables, so 528 news shocks cannot be a valid candidate for the main driving force of business cycles. 529

To shed light on the effect of anticipated technology shocks or news shocks on the 530 economy, we estimate a small scale VARMA model similar to Fisher (2006). There are 531 three variables in the model: the growth rate of real equipment price, the growth rate of 532 labor productivity and the log index of average working hours. The rationale behind this 533 exercise is as follows: if there is a significant anticipated component in either the invest-534 ment specific technology shock or the neutral technology shock, the implied time series 535 becomes non-invertible. With our four-step procedure, we should be able to identify the 536 true model with enough reasonable restrictions, no matter it is non-invertible or not. The 537

⁵³⁸ application of the four-step procedure is given as follows:

539

Step one: *Estimate a reduced-form VARMA*(1,1) *on the observables*

First, we estimate a VARMA(1,1) model on the data. In practice, there are at least two ad-541 vantages of this VARMA(1,1) setup over the traditional long VAR models: (i) the model 542 requires less parameters, which relieves the concern on too many estimated parameters 543 to some extent; (ii) the VARMA(1,1) setting is more consistent with the DSGE models 544 studied in macroeconomics.¹⁶ The VARMA model is estimated in a two-step manner. 545 The first step is estimating a long VAR model to obtain a residual series. In the second 546 step, we estimate a VARMA(1,1) model by adding the residual series from the first step 547 as a regressor and check for convergence.¹⁷ After obtaining the estimated VARMA(1,1) 548 model, we get variance matrix of error terms, $\hat{\Omega}$, which is the estimate of $D_i D'_i$, and the 549 MA coefficient matrix, N, which is the estimate of $B_j D_j^{-1}$. These moment estimates are 550 used in the second step. 551

552

Step two: Calculate all covariance equivalent representations

Second, we compute all covariance equivalent representations. As we show in section three, all the covariance equivalent representations are solutions of the Potter equation defined by the moments of observable variables, and the true model should be one of them. In the current application, the Potter equation is given by:

$$D_j D'_j + B_j B'_j = \hat{\Omega} + N \hat{\Omega} N'$$

$$B_j D'_j = N \hat{\Omega}.$$
(16)

Step three: Define the structural shocks of interest and impose an SVAR-type restrictions on each representation.

- ⁵⁶⁰ Following Fisher (2006) and Altig et al (2009), a positive investment specific shock is de-
- fined as the only shock which lowers the real equipment price in the long run, while a

¹⁶See for example Kehoe (2007).

¹⁷The efficiency of estimation could be improved by employing a 3SLS procedure or iterated 2SLS procedure. Kascha (2007) gives a good survey on estimation methods of the VARMA models.

⁵⁶² positive neutral technology shock is define as the other shock which increases labor pro-⁵⁶³ ductivity in the long run apart from the positive investment specific shock. Based on the ⁵⁶⁴ definitions, two long run restrictions are imposed on the estimated model to identify the ⁵⁶⁵ two technology shocks. There are eight structural representations satisfying the Potter ⁵⁶⁶ equation as well as the two long run restrictions.

Figure 5 shows the impulse responses of all eight cases along with the point estimate 567 and the confidence interval based on the innovation representation. The latter is the coun-568 terpart of the traditional VAR identification in our VARMA(1,1) setup. In the invertible 569 case, the estimated effect of identified shocks are in line with existing research: in re-570 sponse to a positive investment shock, hours and output increase prominently, however, 571 labor productivity falls for a long period after the shock. Output and labor hours increase 572 less significantly in the case with a positive neutral technology shock. In non-invertible 573 cases, the responses to the investment shocks are similar to those in the invertible case. 574 In response to the neutral technology shock, hours rise faster and stronger in some non-575 invertible cases, but the response of output on impact becomes weaker. In those cases, 576 labor productivity increases gradually, instead of jumping up as shown in the invertible 577 case. If technology is only disseminated slowly in the economy, we should observe the 578 slow buildup of labor productivity in response to technology shocks as shown here. The 579 strong response of hours in can be readily explained by strong intertemporal substitution 580 effect as in Jaimovich and Rebelo (2009). Up to this step, economic theory cannot distin-581 guish between the invertible and the invertible models. Therefore, we need additional 582 selection criteria to pin down the true model, which is the purpose of the fourth step in 583 our procedure. 584

585

Step four: Impose agnostic restrictions on each representation, delivered from step three, to fur ther rule out structural responses.

In this step, we impose agnostic restrictions on variance decompositions: (i) the investment shock should explain the long run variance in the growth of real equipment price at least 10%; (ii) the neutral technology shock contributes the long run variance on the growth of labor productivities at least 10%; (iii) the third shock, with is a combination of



Figure 5: Response To Technology Shocks (All Cases)

Notes: solid blue line: the point estimate of impulse responses in the innovation representation; gray area: 90% confidence interval in the innovation representation; dashed black lines: impulse responses from the solutions of the Potter equation

other non-technology shocks and measurement errors, should not contribute more then
30% to the long run volatility in either the real equipment price or the labor productivity.
The result of the variance decomposition is summarized in table 2.

As shown in the table, we successfully rule out some cases. Based on the third criterion, we can rule out case models 1, 3, 5 and 7. In all the four cases, the contribution of other non-technology shocks on the growth of technology in the long run are unreasonably large. However, we cannot refine the outcome further, in other words, we only achieve a partial identification in this example.

Figure 6 plot the responses of models satisfying the agnostic restrictions based on vari ance decompositions along with the invertible case. In all the four valid cases, impulse
 responses are very similar to each other. Furthermore, the invertible case is among the

four cases we keep. The variance decomposition analysis also show similar result in all 603 the four cases. Therefore, we can reach the conclusion that the inference based on analy-604 sis on an invertible VAR model is valid and reliable. In other words, news or anticipated 605 components in technology shocks does not play important roles when studying the ef-606 fect of these two types of technology shocks. Between the two technology shocks, the 607 investment specific shock is more important to explain the dynamics in labor hours. In 608 additional, we notice that the remaining cases actually "cluster" based on our identifica-609 tion. It might indicate all the identification on technology shocks are correct, while the 610 identification of the third shock might differ. If our interest is only on technology shocks, 611 we probably can keep all of them. 612



Figure 6: Response To Technology Shocks (Identified)

Notes: solid blue line: the point estimate of impulse responses in the innovation representation; gray area: 90% confidence interval in the innovation representation; dashed black lines: impulse responses from the solutions of the Potter equation

	Model	Model	Model	Model	Model	Model	Model	Model
	One	Two	Three	Four	Five	Six	Seven	Eight
	con	tribution	of invest	ment sho	cks in the	e long run		
variable 1	97.32	96.46	97.32	96.07	98.24	98.88	98.31	99.22
variable 2	8.87	9.93	8.86	10.10	8.06	7.55	8.08	7.51
variable 3	51.66	51.35	51.66	51.51	51.68	51.69	51.62	51.50
	ŭ	ontributic	in of neut	tral shock	s in the l	ong run		
variable 1	.35	2.84	0.32	2.23	0.40	0.69	0.31	0.46
variable 2	5.69	74.50	6.14	70.08	5.74	76.77	6.14	72.24
variable 3	17.29	13.07	17.23	12.96	17.19	13.31	17.22	13.21
		contributi	on of oth	er shocks	in the lo	ng run		
variable 1	2.33	0.70	2.35	1.69	1.36	0.43	1.38	0.32
variable 2	85.44	15.57	85.00	19.82	86.21	16.68	85.77	20.24
variable 3	31.04	35.58	31.11	35.99	31.11	35.00	31.18	35.28

Table 2: Identification Based on Short-Term Variance Decomposition

variable 1: the growth rate of real equipment price; variable 2: the growth rate of labor productivity; variable 3: labor hours

613 6 Conclusion

Traditional limited information econometric methods, including the widely applied struc-614 tural VAR approach, cannot handle non-invertiblility embedded in many business cy-615 cle models. However, researchers need not abandon the limited information approach, 616 which is the power and soul of the structural VAR. We show that non-invertible time se-617 ries can be recovered with its invertible counterpart. That is, there is always an invertible 618 innovation representation corresponding to a non-invertible model. The invertible in-619 novation representation shares the same population moment with the structural model. 620 Therefore, we can recover all the valid models through those consistently estimated mo-621 ments, regardless of invertibility. 622

Based on the theory developed in this paper, we propose a four step procedure to 623 handle non-invertibility in practice. This four steps are: (i) estimate a reduced form 624 VARMA(1,1); (ii) compute all VARMA(1,1) models with the same autocovariance struc-625 ture using Potter's (1964) algorithm; (iii) use the outcomes from step two and an SVAR-626 type restriction to find a finite number of valid structural impulse responses; (iv) use ag-627 nostic restriction implied by economic theory to identify, at least partially, the true model. 628 We then apply this procedure to two model-generated examples. In both the perma-629 nent income model of FRSW (20007) and the anticipated fiscal shock model in LWY, our 630 procedure recovers the true model. We further apply our method to cases with real data. 631 We find that result in Fisher (2006)'s study on technology shocks holds even when we 632 consider possible non-invertibilities in the model. It indicates that anticipated compo-633 nent technology shocks or "news shocks" do not spoil the inference of the transmission 634 mechanism of technology shocks. 635

636 Appendix

637 A Proofs

Proof of Theorem 1: First, we prove equations (2) -(4) are necessary for any MA representation
tation to be covariance equivalent to the structural form. That is, every MA representation
of the structural form satisfies these conditions. The structural form has an MA representation
tation in the same format as (2).

Let \overline{W} be the left inverse of W, which exists by Assumption 1. The MA representation of s_{t+1} is:

$$s_{t+1} = (I - QL)^{-1} U e_{t+1} = \sum_{i=0}^{\infty} Q^i U e_{t+1-i}.$$
 (A. 1)

⁶⁴⁴ Substituting (A. 1) in the observer equation from the state-space form is:

$$r_{t+1} = W \sum_{i=0}^{\infty} Q^i U e_{t-i} + Z e_{t+1},$$
(A. 2)

Premultiplying both side by $\overline{W}L$ and rearranging,

$$\sum_{i=0}^{\infty} Q^i U e_{t-1-i} = \bar{W}(r_t - Z e_t)$$

⁶⁴⁵ Hence, (A. 2) can be rewritten as:

$$r_{t+1} = W[Ue_t + Q\bar{W}(r_t - Ze_t)] + Ze_{t+1}$$

$$= WQ\bar{W}r_t + Ze_{t+1} + (WU - WQ\bar{W}Z)e_t,$$
(A. 3)

⁶⁴⁶ The MA representation of (A. 3) is given by:

$$r_{t+1} = [I - WQ\bar{W}L]^{-1}[Z + W(U - Q\bar{W}Z)L]e_t,$$
(A. 4)

647

In the next step, we prove that $WQ\overline{W} = A$ and $W(U - Q\overline{W}Z) = \tilde{C}_1(Z')^{-1}$.

Next, we show that the MA representation (A. 4) satisfies (3) and (4). Define C_i to be the *i*th order autocovariance matrix of r_t .

Suppose y_t is a general VARMA(p,q), $y_t = M(L)y_t + N(L)w$ with $w_t \sim \mathcal{N}(0, I)$. The autocovariance-generating function of y_t is

$$G_{y}(z) = \left[I - M(z)\right]^{-1} N(z) N(z^{-1})' \left[I - M(z^{-1})'\right]^{-1}$$

⁶⁵¹ Therefore, we have:

$$C_{0} = E(r_{t}r_{t}')$$

= $WQ\bar{W}C_{0}(WQ\bar{W})' + ZZ' + WUU'W'$
 $-WQ\bar{W}ZZ'(WQ\bar{W})'$ (A. 5)

$$C_1 = E(r_t r_{t-1})$$

= $WQ\bar{W}C_0 + WUZ' - WQ\bar{W}ZZ'$ (A. 6)

$$C_{i} = E\left(r_{t}r_{t-i}'\right)$$

= $\left(WQ\bar{W}\right)^{i-1}C_{1} \text{ for } i \geq 2$ (A. 7)

Simplifying notation, let $A = WQ\overline{W}$, B = WU - AZ and D = Z. Then,

$$A = WQ\bar{W} = C_2C_1^{-1}$$

Based on the definitions, \tilde{C}_0 and \tilde{C}_1 satisfy:

$$\tilde{C}_1 = C_1 - AC_0 = BD'$$

$$\tilde{C}_0 = C_0 - AC_0A' - A\tilde{C}_1' - \tilde{C}_1A' = DD' + BB'$$

Therefore, we have:

$$B = W(U - Q\bar{W}Z) = \tilde{C}_1 Z'^{-1}$$

We further substitute *B* in the equation for \tilde{C}_0 ,

$$\tilde{C}_0 = ZZ' + \tilde{C}_1 (ZZ')^{-1} \tilde{C}_1'$$

⁶⁵³ Premultiplying both sides with $\tilde{C}_1^{-1}ZZ'$,

$$(ZZ')(\tilde{C}'_1)^{-1}(ZZ') - \tilde{C}_0(\tilde{C}'_1)^{-1}(ZZ') + \tilde{C}_1 = 0,$$
(A. 8)

Thus, ZZ' satisfies (4). Also, since ZZ' is a symmetric positive semi-definite matrix, its Cholesky decomposition generates a lower triangular matrix Z^c such that $Z^cZ^{c'} = ZZ'$. Based on Uhlig (2005), there always exists an orthonormal matrix K such that $K = (Z^c)^{-1}Z$.

658

In our final step, we show that (2)-(4) are also sufficient for a valid covariance equivalent representation: every process satisfying (2)-(4) is covariance equivalent the structural form.

It is obvious that the proposed representations have the same first moments as the structural form. Hence, if the second moments of the proposed processes are also the same as those implied by the structural form, then the proposed forms are covariance equivalent..

⁶⁶⁶ Based on the construction, the general form of each candidate is:

$$\hat{r}_{t+1} = A\hat{r}_t + Z_j \varepsilon_{t+1}^j + \tilde{C}_1 (Z_j')^{-1} \varepsilon_t^j$$
(A. 9)

where $\varepsilon_t^j \sim \text{is } \mathcal{N}(0, I)$ and A, Z_j and \tilde{C}_1 are determined by (3) and (4). The autocovariance of \hat{r}_t is:

$$\begin{aligned} \hat{C}_{0} &= E\left(\hat{r}_{t+1}\hat{r}_{t+1}'\right) \\ &= A\hat{C}_{0}A' + AZ_{j}(Z_{j})^{-1}\tilde{C}_{1} + \left(AZ_{j}(Z_{j}')^{-1}\tilde{C}_{1}'\right)' + Z_{j}Z_{j}' + \tilde{C}_{1}(Z_{j}Z_{j})^{-1}\tilde{C}_{1}' \\ \hat{C}_{1} &= E\left(\hat{r}_{t}\hat{r}_{t-1}'\right) = A\hat{C}_{0} + \tilde{C}_{1}(Z_{j}')^{-1}Z_{j} \\ \hat{C}_{i} &= E\left(\hat{r}_{t}\hat{r}_{t-i}'\right) = (A)^{i-1}\hat{C}_{1} \text{ for } i \geq 2 \end{aligned}$$

⁶⁶⁹ Since $Z_j Z'_j$ is a solution to (A. 8),

$$\tilde{C}_0 = (Z_j Z'_j) + \tilde{C}_1 (Z_j Z'_j)^{-1} \tilde{C}'_1$$
(A. 10)

⁶⁷⁰ Therefore, the equation for \tilde{C}_0 becomes:

$$\hat{C}_0 = A\hat{C}_0 A' + A\tilde{C}_1 + \tilde{C}_1 A' + \tilde{C}_0$$
(A. 11)

⁶⁷¹ Hence, the solution of \hat{C}_0 is given by

$$vec(\hat{C}_0) = [I - (A \otimes A)]^{-1} vec(A\tilde{C}_1 + \tilde{C}_1 A' + \tilde{C}_0)$$
 (A. 12)

where $vec(\bullet)$ is the vectorization operation turning an *m* by *n* matrix into an *mn* by 1 vector. Based on the definition of \tilde{C}_0 and \tilde{C}_1 ,

$$vec(C_0) = [I - (A \otimes A)]^{-1} vec(A\tilde{C}_1 + \tilde{C}_1 A' + \tilde{C}_0)$$
 (A. 13)

674 Therefore,

$$\hat{C}_0 = C_0.$$
 (A. 14)

Given the equivalence between C_0 and \hat{C}_0 , it is easy to see that

$$\hat{C}_1 = A\hat{C}_0 + \tilde{C}_1 = AC_0 + \tilde{C}_1 = C_1$$
 (A. 15)

676 and

$$\hat{C}_i = A^{i-1}\hat{C}_1 = A^{i-1}C_1 = C_i, \ \forall i \ge 2.$$
 (A. 16)

⁶⁷⁷ Hence, we if a representation satisfies (2)-(4), it is covariance equivalent to the structural
⁶⁷⁸ form.

form. As for the number of valid Z_j s, there are $\begin{pmatrix} 2k \\ k \end{pmatrix}$ solutions to equation (c). The form of $Z_j Z'_j$ requires it to be symmetric and positive definite; thus, the valid solution is less than $\begin{pmatrix} 2k \\ k \end{pmatrix}$. With an alternative approach, we can show there are a total of 2^k valid representations. Furthermore, we show that among all the valid covariance-equivalent representations, there is one presentation which is invertible. The detail of this alternative approach is included in appendix B. **Q.E.D**

$\mathbf{B}_{\mathrm{oss}}$ **B** The equivalence between Blaschke Matrices and the Potter Equation

Lippi and Reichlin (1994) show every noninvertible stationary VARMA(p,q) model has one invertible representation by multiplying an appropriate Blaschke matrix. A Blaschke matrices, B(z), is a special matrix satisfying the following property:

689

$$B(z)B(z^{-1})' = I.$$
 (B. 17)

As we know, every orthonormal matrix is a Blaschke matrix. In the remaining part of this section, we show how to use Blaschke matrices to get an invertible representation and how this alternative procedure is related to the proposed procedure in the main text.

693 *Lemma* Every covariance-equivalent form can be achieved by multiplying an appropriate Blaschke

695 **Proof:**

$$r_{t+1} = WI - QL^{-1}Ue_t + Ze_{t+1}$$
(B. 18)
$$= W\sum_{i=0}^{\infty} Q^i Ue_{t-i}$$

$$= WQ\bar{W}(r_t - Ze_t) + WUe_t + Ze_{t+1}$$

$$= WQ\bar{W}y_t + Ze_{t+1} + (WU - WQ\bar{W}Z)e_t.$$

⁶⁹⁶ For simplicity in notations, define $M = WQ\bar{W}$, $N_0 = Z$ and $N_1 = WU - WQ\bar{W}Z$. There-⁶⁹⁷ fore, we have the autocovariance generating function of r_t is given by:

$$G_r(z) = ([I - Mz])^{-1} (N_0 + N_1 z) (N_0 + N_1 z^{-1})' [I - M'^{-1}]^{-1}$$
(B. 19)

⁶⁹⁴ *matrix on the original model*

Equation () is a VARMA(1,1) representation of the structural model, which might be invertible or non-invertible. Next, we show that there is an alternative VARMA(1,1) representation of the same model, and furthermore, this representation is invertible. To this end, we construct a square matrix A(L) of dimension *m*. This matrix depends on the matrix lag polynomial $N(L) = N_0 + N_1L$. More specifically, let $\{\lambda_i\}_{i=1}^m$ be the eigenvalues of N(L). Define a matrix $R(\lambda_i, z)$ as follows:

$$R(\lambda_{i}, z) = \begin{cases} \begin{pmatrix} I_{i-1} & 0 & 0 \\ 0 & \frac{1-\bar{\lambda_{i}z}}{1-\lambda_{i}z} & 0 \\ 0 & 0 & I_{m-i} \end{pmatrix}, & |\lambda_{i}| > 1 \\ & & I_{m}, & \text{otherwise} \end{cases}$$
(B. 20)

The matrix $R(\lambda_i, z)$ is known as a Blaschke matrix. It satisfies the property $R(\lambda_i, z)R'(\bar{\lambda}_i, z^{-1}) =$ *I.* Now, we defines another matrix K_i . This matrix is an orthonormal matrix, whose *i*th column is the normalized solution of $N(\lambda_i)x = 0$.

Firstly, we can construct another lag polynomial $N^i(L) = N_0^i + N_1^i L = (N_0 + N_1 L) K_i R(\lambda_i, L)$. By right multiplying N(L) with K_i , one can move all the entries containing the factor $1 - \lambda_i L$ on the *i*th column. By further right multiplying $R(\lambda_i, L)$, one replaces $1 - \lambda_i L$ with $\lambda_i - L$ but leave other elements untouched, in other words, "flips" a particular eigenvalue of the lag polynomial. At the same time, we even have:

$$G_{r}^{i}(z) = ([I - Mz])^{-1} (N_{0}^{i} + N_{1}^{i}z) (N_{0}^{i} + N_{1}^{i}z^{-1})' [I - M'^{-1}]^{-1}$$

$$= ([I - Mz])^{-1} (N_{0} + N_{1}z) K_{i} R_{i} (\lambda_{i}, L) R' (\bar{\lambda}_{i}, L^{-1}) K'_{i} (N_{0} + N_{1}z^{-1})' [I - M'^{-1}]^{-1}$$

$$= ([I - Mz])^{-1} (N_{0} + N_{1}z) (N_{0} + N_{1}z^{-1})' [I - M'^{-1}]^{-1}$$

$$= G_{r}(z)$$
(B. 21)

⁷¹² Therefore, we construct another VARMA(1,1) representation of the structural model:

$$r_{t+1} = Mr_t + N_0^i e_{t+1}^i + N_1^i e_t^i.$$
(B. 22)

Compared to the model in equation (B), model (B. 22) has the same variance-covariance structure and the same likelihood. Based on construction, we know that the eigenvalues of the covariance-equivalent forms are either the eigenvalues of the structural form or the reciprocal of them. Therefore, if there are eigenvalues outside the unit circle (noninvertible), there has to be a covariance-equivalent form "flipping" all the explosive eigenvalues while keeping the stable eigenvalues untouched.

- 719 Q.E.D
- 720

Lemma The method with Blaschke matrices gives the same result as the procedure based on the
 Potter equation

723

Proof: The proof applies to a general VARMA(p,q)model, $M(L)x_t = N(L)w_t$, where M(L) is stable. (*i*) Any solution implied by Blaschke matrices is a solution implied by the Potter equation. This is obvious. Based on construction, a representation generated by using Blaschke matrices have the same covariance structure as the structural form. Hence, it is satisfies conditions (2) to (4)

729

Any solution satisfying conditions (2) to (4) is a solution by using Blaschke matrices This is based on Theorem 2 in Lippi and Reichlin (1994). Assume the invertible VARMA(p,q)model is given by $M(L)x_t = N(L)u_t$. an arbitrary solution from the potter equation is given by $M(L)x_t = \tilde{N}(L)w_t$. Based on definition, $x_t = M(L)^{-1}\tilde{N}(L)w_t$ is a MA representation of the original VARMA model. Therefore, we have to have $M(L)^{-1}\tilde{N}(L) = M(L)^{-1}N(L)B(L)$, where B(L) is a Blaschke matrix. Thus, $\tilde{N}(L) = N(L)B(L)$.

⁷³⁷ Q.E.D

738 References

- Alessi, B., M. Barigozzi and M. Capasso (2011), "Nonfundamentalness in Structural Eco-
- ⁷⁴¹ nomic Models: A Review," International Statistical Review, forthcoming
- Beaudry, P., and F. Portier (2006), "Stock Prices, News and Economic Fluctuation," Amer *ican Economic Review* 96, pp. 1293-1307.
- ⁷⁴⁴ Bernanke, B., Boivin, J. and Eliasz, P. (2005), "Measuring the Effects of Monetary Policy: A
- ⁷⁴⁵ Factor-Augmented (FAVAR) Approach," *Quarterly Journal of Economics* 120, pp. 387-422.
- Blanchard, O. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Sup ply Disturbances," *American Economic Review* 79, 655-680.
- ⁷⁴⁸ Blanchard, O. and R. Perroti (2002), "An Empirical Characterization of the Dynamic Ef-
- ⁷⁴⁹ fects of Changes in Government Spending and Revenue on Output," *Quarterly Journal of*
- 750 *Economics* 117, pp. 1329-1368.
- ⁷⁵¹ Blanchard, O., G. Lorenzoni and J. L'Huillier, "News, Noise and Fluctuations: An Empir ⁷⁵² ical Exploration," manuscript, MIT.
- 753 Brockwell, P. and R. Davis, *Time Series: Theory and Method*, Springer.
- ⁷⁵⁴ Cardosa-Mendonca, M., L. Medrano and A. Sachsida (2008), "Evaluating Fiscal Policy
 ⁷⁵⁵ Effects in Brazil Using Agnostic Restrictions," Institute of Applied Economic Research,
 ⁷⁵⁶ working paper.
- ⁷⁵⁷ Chari, V.V., P. Kehoe and E. McGrattan (2008), "Are Structural VARs with Long-Run Re ⁷⁵⁸ strictions Useful in Developing Business Cycle Theory?" *Journal of Monetary Economics* 50,
 ⁷⁵⁹ pp. 1337-1352.
- ⁷⁶⁰ Christiano, L., R. Motto and M. Rostagno (2009), "Financial Factors in Economic Fluctua ⁷⁶¹ tions," manuscript, European Central Bank.
- ⁷⁶² Clarida, R., J. Gali and M. Gertler (2000), "Monetary Policy Rules and Macroeconomic
 ⁷⁶³ Stability: Evidence and Some Theory," *Quarterly Journal of Economics* 115, 147-180.
- ⁷⁶⁴ Dufour, J.M. and D. Pelletier (2008), "Practical Methods for Modelling Weak VARMA Pro-
- ⁷⁶⁵ cesses: Identification, Estimation and Specification with a Macroeconomic Application,"
- ⁷⁶⁶ working paper, McGill University.

- Faust, J. (1998), "The Robustness of Identified VAR Conclusions about Money," Carnegie-767 Rochester Conference Series on Public Policy 49, 207-244. 768
- Fernandez-Villaverde, J., J. Rubio-Ramirez, T. Sargent and M. Watson, "ABCs (and Ds) of 769
- Understanding VARs," American Economic Review 97, 1021-1026 770
- Fisher, J. (2006), "The Dynamic Effects of Neutral and Investment-Specific Technology 771 Shocks," Journal of Political Economy 114, pp. 413-451. 772
- Forni, M., Giannone, D., Lippi, M., and Reichlin, L. (2009), "Opening the black box: struc-773 tural factor models with large cross-sections," Econometric Theory 25, pp. 1319–1347. 774
- Fry, R. and A. Pagan (2010), "Sign Restrictions in Structural Vector Autoregressions: A 775 Critical Review," Australian National University, manuscript. 776
- Giannone, D. and L. Reichlin (2006), "Does Information Help Recovering Structural Shocks 777
- from Past Observations?" Journal of the European Economic Association 4, pp. 455-465. 778
- Goloslov, M., and R. Lucas (2007), "Menu Cost and Phillips Curve," Journal of Political 779 Economy 117, pp. 171-199. 780
- Hannan, E. J., And J. Rissanen (1982), "Recursive Estimation of Mixed Autoregressive-781 Moving- Average Order," Biometrika 69, pp. 81-94. 782
- Hamilton, J. (1994), Time Series Analysis, Princeton. 783

- Hansen, L., W. Roberds and T. Sargent (1991), "Time Series Implications of Present Value 784 Budget Balance," in Rational Expectations Econometrics (L. Hansen and T. Sargent, eds), 785 Westview, pp. 121-161. 786
- Hansen, L. and T. Sargent (1991), "Two Difficulties in Interpreting Vector Autoregres-787 sions," in Rational Expectations Econometrics (L. Hansen and T. Sargent, eds), Westview, 788 pp. 77-120. 789
- Kascha, C. (2007), "A Comparison of Estimation Methods for VARMA Models," manuscript, 790 European University Institute.
- Leeper, E., T. Walker and S. Yang (2009), "Fiscal Foresight and Information Flows," Indi-792 ana University, manuscript 793

- Lippi, M., and L. Reichlin (1994), "VAR Analysis, Nonfundamental Representations, Blaschke
 Matrices," *Journal of Econometrics* 63, pp. 307-325.
- Lippi, M., and L. Reichlin (2003), "The Dynamic Effects of Aggregate Demand and Supply
 Disturbances: Comment," *American Economic Review* 83, pp. 644-652.
- ⁷⁹⁸ Lutkepohl, H., (2010), *New Introduction to Multiple Time Series Analysis*, Springer.
- ⁷⁹⁹ Mertens, K. and M. Ravn (2010), "Measuring the Impact of Fiscal Policy in the Face of ⁸⁰⁰ Anticipation: A Structural VAR Approach," *Economic Journal*, 120, pp. 393-413.
- Mountford, A. and H. Uhlig (2009), "What Are the Effects of Fiscal Policy Shocks?" *Journal of Econometrics* 24, pp. 960-992.
- ⁸⁰³ Owyang, M. (2002), "Modeling Volcker as a non-absorbing state: agnostic identification of
- a Markov-switching VAR," Working Papers 2002-018, Federal Reserve Bank of St. Louis.
- Potter, J. (1966), "Matrix Quadratic Solutions," SIAM Journal on Applied Mathematics, 14,
 pp. 496-501.
- Romer, D. and C. Romer (2010), "The Macroeconomic Effects of Tax Changes: Estimates
 Based on a New Measure of Fiscal Shocks," *American Economic Review* 100, pp. 763-801.
- ⁸⁰⁹ Romer, D. and C. Romer (2004), "A New Measure of Monetary Shocks: Derivation and ⁸¹⁰ Implications," *American Economic Review* 94, pp. 1055-1084.
- ⁸¹¹ Sargent, T. (1987), *Macroeconomic Theory*, Academic Press.
- Scholl, A. and H. Uhlig (2008), "New evidence on the puzzles: Results from agnostic
 identification on monetary policy and exchange rates," *Journal of International Economics*76, pp. 1-13.
- Sims, C. (2002), "Solving Linear Rational Expectation Models," *Computational Economics*,
 20, pp. 1-20.
- Sims, E. (2009), "Non-invertibilities and Structural VARs," manuscript, University of Notre
 Dame.
- Stock, J. and M. Watson (2010), "Dynamic Factor Models," in Oxford Handbook of Eco-
- nomic Forecasting (M. Clements and D. Henry, eds), forthcoming.

- ⁸²¹ Uhlig, H. (2005). "What are the effects of monetary policy on output? Results from an ⁸²² agnostic identification procedure," *Journal of Monetary Economics*, 52, pp. 381-419.
- ⁸²³ Whiteman, C. (1983), *Linear Rational Expectations Models: A User's Guide*, University of ⁸²⁴ Minnesota Press.