

Waggoner, Daniel F.; Zha, Tao

Working Paper

Likelihood-preserving normalization in multiple equation models

Working Paper, No. 2000-8

Provided in Cooperation with:

Federal Reserve Bank of Atlanta

Suggested Citation: Waggoner, Daniel F.; Zha, Tao (2000) : Likelihood-preserving normalization in multiple equation models, Working Paper, No. 2000-8, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:

<https://hdl.handle.net/10419/100790>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



FEDERAL
RESERVE
BANK
of ATLANTA

**Likelihood-Preserving Normalization in
Multiple Equation Models**

Daniel F. Waggoner and Tao Zha

Working Paper 2000-8
June 2000

Working Paper Series

Likelihood-Preserving Normalization in Multiple Equation Models

Daniel F. Waggoner and Tao Zha

Federal Reserve Bank of Atlanta
Working Paper 2000-8
June 2000

Abstract: Causal analysis in multiple equation models often revolves around the evaluation of the effects of an exogenous shift in a structural equation. When taking into account the uncertainty implied by the shape of the likelihood, we argue that how normalization is implemented matters for inferential conclusions around the maximum likelihood (ML) estimates of such effects. We develop a general method that eliminates the distortion of finite-sample inferences about these ML estimates after normalization. We show that our likelihood-preserving normalization always maintains coherent economic interpretations while an arbitrary implementation of normalization can lead to ill-determined inferential results.

JEL classification: C32, E52

Key words: Bayesian methods, causal analysis, supply and demand, simultaneity, likelihood shape, equilibrium effects

The authors are indebted to John Geweke and Chris Sims, whose encouragement and suggestions have led to significant improvement of this paper. The authors have also benefited from discussions with Roberto Chang, Tom Cunningham, Joel Horowitz, Clive Granger, Eric Leeper, Adrian Pagan, Ellis Tallman, Chuck Whiteman, Arnold Zellner, and seminar participants at the University of Iowa, Indiana University, UCSD, and UCLA. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Daniel Waggoner, Economist, Research Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713, 404/521-8278, daniel.f.waggoner@atl.frb.org, or Tao Zha, Senior Economist and Policy Adviser, Research Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713, 404/521-8353, tao.zha@atl.frb.org.

The full text of Federal Reserve Bank of Atlanta working papers, including revised versions, is available on the Atlanta Fed's Web site at http://www.frbatlanta.org/publica/work_papers/index.html. To receive notification about new papers, please use the publications order form, or contact the Public Affairs Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713, 404/521-8020.

Likelihood preserving normalization in multiple equation models

1. Introduction

It is widely recognized that “economic data are generated by systems of relations that are in general stochastic, dynamic, and simultaneous” (Marschak 1950). Stochastic multiple-equation models, therefore, play a central role in understanding causal relationships among economic variables (Tobin 1970). In the money market, for example, the money stock (M) and the interest rate (R) are jointly and interdependently determined in equilibrium. Such a simultaneous feedback mechanism is common across all other quantity and price variables and has been traditionally expressed as the following set of multiple equations:

$$\begin{aligned} a_{11}M + a_{21}R &= \alpha X_{MS} + \varepsilon_{MS} \quad (\text{Money Supply}) \\ a_{12}M + a_{22}R &= \beta_1 y + \beta_2 P + \beta_3 X_{MD} + \varepsilon_{MD} \quad (\text{Money Demand}) \end{aligned} \quad (1)$$

where y represents output or income, P is the price level, X_{MS} is a vector of variables entering the money supply equation, and X_{MD} is a vector of variables that enter the money demand equation. The exogenous disturbances ε_{MS} and ε_{MD} are independent random variables.

One unresolved issue in the inference of multiple equation models like (1) is normalization. It is well known that how normalization is implemented has no consequence on the inference of, say, the interest elasticity $-a_{21}/a_{11}$. But if the purpose is to obtain the statistical reliability of the estimated effects of an exogenous shift in a structural equation (such as money supply), arbitrary implementation of normalization can distort finite-sample inferences about the maximum likelihood (ML) estimates in multiple equation models. This point is related to Dent and Geweke (1980), who argue for careful normalization; it is recently iterated in Sims and Zha (1999):

[It] is widely understood that ... all coefficients seem ill-determined by normalizing the equation on a variable whose coefficient is insignificantly different from zero. ... Casual choice of normalization can lead to estimates that all responses to, say, a policy shock are “insignificant,” when a better normalization would make it clear that the responses are actually sharply determined.

In this paper we address some important practical issues related to normalization:

(I) Causal analysis in multiple equation models often involves controlled experiments by examining the effects on economic variables of an exogenous shift in, say, the supply equation while holding all other structural equations fixed. Take the example of simultaneous system (1), where the liquidity effect of an exogenous disturbance ε_{MS} on money and the interest rate has been an important issue in policy analysis.¹ Suppose the money demand is taken as given with the negative slope in the M - R plane. An expansionary monetary policy disturbance should simultaneously increase the money stock (quantity) and decrease the interest rate (price). When the uncertainty of a_{11} and a_{21} in the money supply equation is taken into account, the direction of this expansionary effect — the increase of M and the decrease of R — should not depend on particular values of these parameters. As the values of a_{11} and a_{21} are drawn from the likelihood function, however, naive implementation of normalization can lead to inferentially ambiguous conclusions (i.e., the effect of an expansionary policy shift has the opposite direction for some supply equations — decreasing M and increasing R). We explain how this anomalous result can occur and develop a general method, called likelihood preserving (LP) normalization, that resolves such an anomaly for both recursive and simultaneous-equation systems.

(II) When allowing the parameters in the other equations to vary, we show that the LP normalization minimizes the distance between the ML estimate of the effect of an exogenous shift and the normalized value of this effect sampled from the likelihood distribution. This theoretical result implies that the LP normalization eliminates the needless distortion of finite-sample inferences about the ML estimates. An accurate characterization of how sharp the ML estimates are is important in scientific reporting of empirical results. We give an applicable example to show the differences in results produced by the LP normalization and by naïve implementation of the standard rule.

In a nutshell, this paper makes two points. First, it shows that the way normalization is implemented matters not only in principle but in practice as well. We use the familiar work of Sims (1986) as an empirical example for illustration. Second, we offer a solution to the implementation of normalization for both recursive and simultaneous systems.

¹ See, for example, Poole (1970); Sargent and Wallace (1975); Leeper and Gordon (1992); and Pagan and Robertson (1995).

To elaborate these points, we organize the rest of the paper as follows. Section 2 sets out the general framework. Section 3 provides an empirical example of money demand and supply to show that the problem associated with arbitrary implementation of normalization can be serious. While other researchers adopt ad hoc ways to fix the potential problem, Section 4 presents a general method. Section 5 shows that the method works for the application considered in Section 3.

2. Likelihood and normalization

The stochastic, dynamic, and simultaneous models studied in this paper have the general form of a system of multiple equations:

$$\mathbf{y}'_t \mathbf{A} = \sum_{\ell=1}^p \mathbf{y}'_{t-\ell} \mathbf{A}_\ell + \mathbf{z}'_t \mathbf{D} + \boldsymbol{\varepsilon}'_t, \quad t = 1, \dots, T, \quad (2)$$

where \mathbf{y}_t is an $n \times 1$ vector of endogenous variables; \mathbf{z}_t is an $h \times 1$ vector of exogenous variables; \mathbf{A} or \mathbf{A}_ℓ is an $n \times n$ parameter matrix; \mathbf{D} is an $h \times n$ parameter matrix; $\boldsymbol{\varepsilon}_t$ is an $n \times 1$ vector of structural shocks; p is the lag length and T is the sample size. The structural shocks are assumed to be Gaussian with²

$$E(\boldsymbol{\varepsilon}_t | \mathbf{y}_{t-s}, s > 0) = \mathbf{0}_{n \times 1}, \quad E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t | \mathbf{y}_{t-s}, s > 0) = \mathbf{I}_{n \times n}.$$

Note that the columns of \mathbf{A} , \mathbf{A}_ℓ , and \mathbf{D} correspond to the individual equations of the model.

Let $\mathbf{F} = [\mathbf{A}'_1 | \dots | \mathbf{A}'_n | \mathbf{D}']'$. As shown in Sims and Zha (1998a), the likelihood function (or the likelihood function multiplied by the widely used reference prior of Sims and Zha) takes the form $p(\text{vec}(\mathbf{A}))p(\text{vec}(\mathbf{F}) | \mathbf{A})$ where

$$p(\text{vec}(\mathbf{A})) |\det \mathbf{A}|^T \exp\left(-\frac{T}{2} \text{vec}(\mathbf{A})' (\mathbf{I} \otimes \mathbf{S}) \text{vec}(\mathbf{A})\right); \quad (3)$$

$$p(\text{vec}(\mathbf{F}) | \mathbf{A}) = \varphi((\mathbf{I} \otimes \boldsymbol{\Xi}) \text{vec}(\mathbf{A}); \mathbf{I} \otimes \boldsymbol{\Sigma}). \quad (4)$$

² The independence of structural disturbances is a standard assumption for unambiguous causal analysis in the identified VAR literature. For detailed discussions, see Leeper, Sims, and Zha (1996).

In (3) and (4), $vec(\cdot)$ denotes the operation of stacking the columns of a matrix into a single column vector; $\varphi(x; y)$ is the normal probability density function with mean x and covariance matrix y ; \mathbf{S}_i is an $n \times n$ symmetric, positive definite matrix; Ξ is an $(np + h) \times n$ matrix; Σ is an $(np + h) \times (np + h)$ symmetric, positive semidefinite matrix. All three matrices \mathbf{S}_i , Ξ , and Σ are functions of the data (and the prior information when an informative prior is imposed).

The existing literature often considers deterministic linear restrictions, especially exclusion restrictions, on the parameters in each structural equation. Specifically, $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ satisfies linear restrictions so that \mathbf{a}_i belongs to the set:

$$R_i = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{Q}_i \mathbf{x} = \mathbf{0}\}, \text{ for } i = 1, \dots, n, \quad (5)$$

where $\mathbf{Q}_1, \dots, \mathbf{Q}_n$ are $n \times n$ matrices and \mathbf{R}^n is Euclidean n -space. It is assumed that the restrictions are non-degenerate in the sense that there exists at least one non-singular matrix \mathbf{A} satisfying the restrictions. Waggoner and Zha (2000) show that the likelihood function has a similar form as (3) and (4) under such linear restrictions.

It is clear from (2) that the simultaneous effect of an exogenous shift s_k in the k th equation on endogenous variables \mathbf{y}_t is $s_k \mathbf{e}_k' \mathbf{A}^{-1}$ where \mathbf{e}_k is the k^{th} column of the $n \times n$ identity matrix. The dynamic effects of such a shift are the nonlinear functions of \mathbf{B}_ℓ and \mathbf{A}^{-1} where $\mathbf{B}_\ell = \mathbf{A}_\ell \mathbf{A}^{-1}$. As will be show, normalization amounts to a rule that determines the sign of all parameters in each structural equation. Clearly, a sign change in the equation switches the sign of the corresponding column in \mathbf{A} and thus of the corresponding row of \mathbf{A}^{-1} , but has no effect on the value of \mathbf{B}_ℓ . Therefore, our normalization analysis concerns \mathbf{A} only.

To focus on the subject of normalization, we only consider the model void of identification problems. That is, there is a unique ML estimate of \mathbf{A} up to sign changes in columns. Because the ML value is the same as the sign of each column changes, there are a total of 2^n ML estimates. The first step of normalization is to arbitrarily choose one of these ML estimates. Denote the normalized ML estimate by $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1 | \dots | \hat{\mathbf{a}}_n]$. To derive the statistical inferences involving of $\hat{\mathbf{A}}^{-1}$ or functions of $\hat{\mathbf{A}}^{-1}$, the second step of normalization is to determine the sign of the value of \mathbf{A} randomly drawn from the likelihood function. It is this second step that is

subject to how normalization is implemented. Before we demonstrate this point, we note that a common practice in the literature is to consider the model given by

$$\mathbf{y}'_t \Gamma = \sum_{\ell=1}^p \mathbf{y}'_{t-\ell} \Gamma_{\ell} + \mathbf{z}'_t \Delta + \eta'_t,$$

where $\eta_t \sim N(\mathbf{0}, \Lambda)$ and Λ is a diagonal matrix with positive diagonal. This model maps into the model given by (2) via $\Gamma = \mathbf{A}\Lambda^{1/2}$, $\Gamma_{\ell} = \mathbf{A}_{\ell}\Lambda^{1/2}$, $\Delta = \mathbf{D}\Lambda^{1/2}$, and $\eta'_t = \varepsilon'_t\Lambda^{1/2}$, where $\Lambda^{1/2}$ is the positive square root of Λ . The conventional choice of normalization arbitrarily divides the structural equation by one nonzero parameter (or the negative of it) and thus restricts the corresponding parameter in each column of Γ to 1 (or -1). Since $\mathbf{A} = \Gamma\Lambda^{-1/2}$ and $\Lambda^{-1/2}$ is a positive diagonal matrix, this is equivalent to restricting the corresponding parameter in each column of \mathbf{A} to be positive (or negative). Thus the conventional normalization is an arbitrary choice of sign restriction on the parameters in the structural equation (Zellner 1971, 250-252). The standard practice for the inference involving $\hat{\mathbf{A}}^{-1}$, coded in the widely used software ESTIMA (Doan 1992), chooses $\hat{\mathbf{A}}$ such that all diagonal entries are positive (the first step of normalization) and then restricts all diagonal entries to be positive for every \mathbf{A} drawn from the likelihood function (the second step of normalization). In the following section we show that this standard procedure suitably designed for recursive systems, when adopted naively, can generate inferential results that are ill determined for simultaneous systems.

3. An empirical example

We apply the standard normalization rule designed for recursive systems to Sims's (1986) second six-variable simultaneous equation system. The model has four lags and uses quarterly data with the sample period 1948:1-1979:3. The six variables are the 3-month Treasury Bill rate (R), the M1 stock (M), real GNP (y), GNP deflator (P), the unemployment rate (U), and gross domestic business investment (I). All variables are in logarithm except the interest rate and the unemployment rate, which are divided by 100. For the reader's convenience, the following table presents the exclusion restrictions on \mathbf{A} in the original paper.

Table 1. Sims's (1986) Second Identification

	<i>MS</i>	<i>MD</i>	<i>Output</i>	<i>Price</i>	<i>Unemp</i>	<i>ID</i>
M	X	X		X		
R	X	X	X	X	X	
Y		X	X	X	X	
P		X		X	X	
U					X	
I		X	X		X	X

The first column in Table 1 lists all the endogenous variables. Starting from the second column, each column represents a behavioral equation labeled on the top of the column. Specifically, “MD” stands for money demand, “MS” for money supply, “Unemp” unemployment, and “ID” investment demand. If a cell is filled with “X”, the variable labeled on the left of the row enters the equation labeled on the top of the column. Empty cells correspond to zero restrictions. For example, the second column describes the money supply equation in which the Federal Reserve responds to the interest rate (R) and money (M) but not to other variables contemporaneously. The description for other behavioral equations is in Sims’s original paper.

It can be seen from Table 1 that the first two contemporaneous structural equations are in the form of (1). The ML estimate for the interest elasticity $-a_{21}/a_{11}$ in money supply is 1.11 and for the interest elasticity $-a_{22}/a_{12}$ in money demand is -1.25 . The estimates are close to Sims’s original ones.³ As pointed out before, statistical inferences of these estimates are invariant to how normalization is implemented. But the invariance breaks down if one is interested in the simultaneous and dynamic effects of an exogenous shift in money supply. Figure 1 reports the effects of a one-standard-deviation shift in money supply for the 32-quarter time horizon under the $a_{11} > 0$ normalization. The middle solid line represents the ML estimate; the outer two bands represent .68 and .90 equal-tail probability bands, which are commonly used in the existing

³ The small discrepancies stem from the fact that the original paper uses a smooth prior that was not explicitly laid out for duplication.

literature.⁴ Two error bands, instead of one, are reported to give the reader better information about the overall shape of the likelihood.

The ML estimates (the solid lines in Figure 1) show the equilibrium effects expected by economists: a liquidity effect (the interest rate falls and the money stock increases in the initial period), an expected inflation effect (the nominal interest rate tends to rise in two years in anticipation of a rise in inflation), and an expansionary effect (output, employment, and investment all rise within two years while the price level rises steadily through the entire horizon). The constructed error bands are meant to measure how sharp these estimates are. Clearly, the .68 and .90 equal-tail probability bands cover both the positive and negative regions so that all the ML estimates are ill determined.⁵ If these error bands were an accurate measure of how sharp the ML estimates are, one would conclude that the estimates are uninformative or “insignificant.”

The error bands reported in Figure 1 are exact finite-sample intervals derived from the likelihood function. The derivation, however, depends on how normalization is implemented. Here, the $a_{11} > 0$ rule is used. While this rule is widely used for recursive systems in the empirical literature, it can result in misleading inferences for simultaneous systems (Sims and Zha 1999). We now show that it is this normalization rule that leads to an inaccurate description of the uncertainty around the ML estimates.

To understand the inferential results derived from the overall uncertainty, it is necessary to first examine the contemporaneous effect of an exogenous shift s_1 in the money supply equation while holding all the other equations fixed. Let the value s_1 to be 1, implying a one-standard-

⁴ The effects are the nonlinear functions of \mathbf{A}^{-1} and \mathbf{B}_ℓ . The probability distributions for the effects can be obtained by first simulating Monte Carlo (MC) draws of \mathbf{A} and \mathbf{A}_ℓ from (3) and (4) and then calculating \mathbf{B}_ℓ for each draw of \mathbf{A} and \mathbf{A}_ℓ . The exact finite-sample inference such as the 90% probability interval is computed through the MC integration. The simulation method is the Gibbs sampler of Waggoner and Zha (2000). Each simulation conducted in this paper generates 1.2 million draws to secure the high numerical accuracy. For an overview of the Gibbs sampling technique, see Chib and Greenberg (1995) and Geweke (1995).

⁵ In the early stage of this research, some researchers suggested that the normalization rule be modified to keeping positive the diagonal entries of \mathbf{A}^{-1} rather than \mathbf{A} . That is, if $\mathbf{A}^{-1}(k, k) < 0$, change the sign of the k^{th} column of \mathbf{A} . The statistical inferences under this rule are nearly identical to Figure 1.

deviation shift of the supply curve. In our example, we fix the parameters in all other equations at a set of likely values, in particular the interest elasticity of money demand is fixed at the ML estimate (-1.25).⁶ Figure 2 plots the probability distributions of the equilibrium effects on M and R . The distributions are distinctively bimodal. There is a substantial probability that the money stock rises and the interest rate falls, consistent with the ML estimates. There is also a nontrivial probability that the equilibrium effect is in the opposite direction: the interest rate rises while the money stock shrinks.⁷ Economic interpretations implied by Figure 2 are therefore ambiguous.

To explain how the bimodal distributions occur in this example, it is necessary to understand the unnormalized likelihood shape of a_{11} and a_{21} in the money supply equation. Figure 3 depicts the contours of such a likelihood function conditional on the other equations being fixed. The normalized ML estimates of a_{11} and a_{21} , as used in Figure 1, are in Quadrants I and II. The thick sloping line is the hyperplane on the a_{11} and a_{21} space, along which the likelihood function has zero density. The hyperplane is uniquely determined by the other columns of \mathbf{A} and we call it dividing hyperplane. Under $a_{11} > 0$, the normalized likelihood shape of a_{11} and a_{21} is bimodal (see Quadrants I and II in Figure 3). This bimodal likelihood leads to the ambiguous equilibrium outcome as shown in Figure 4. Consider Quadrant II in Figure 3 where the supply curve is upward sloping. In this case, the supply shift s_1 increases the quantity from M to M' and decreases the price from R to R' (Figure 4(a)), and the direction of the equilibrium outcome (M rises and R falls) is the same as the ML equilibrium outcome in Figure 1. Now consider Quadrant I in Figure 3 where the supply curve is downward sloping. In this situation, the same

⁶ By likely we mean that the draw is within the 0.68 probability region so that the probability of the likelihood value less than this draw's value is over 40%.

⁷ Such a bimodal phenomenon becomes even more severe when the interest elasticity of money demand is smaller. Small elasticity of money demand can be found in other empirical work (Lieberman 1979; King, Plosser, Stock, and Watson 1991). In this six-variable simultaneous system, however, the effects on M and R depend on not only the money demand equation but also all other equations. Thus, when the uncertainty of all the other columns of \mathbf{A} is taken into account (or integrated out), the effects will be in general different from those conditional on the other columns shown in Figure 2. In comparison, Figure 1 shows that the marginal impact on the initial decline of M shows little probability but there is a substantial probability of the positive response of R in the initial period.

exogenous shift s_1 has the opposite effect: the quantity falls from M to M'' and the price rises from R to R'' (Figure 4(b)). This result follows immediately from the fact that the equilibrium effect on quantity and price of this supply shift is $s_1 \mathbf{e}_1' \mathbf{A}^{-1}$ where \mathbf{e}_1 is the first column vector of the $n \times n$ identity matrix. Thus we have the bimodal results shown in Figure 2.

It is important to take into account the downward sloping supply in reporting statistical inferences of the ML estimates. Researchers, when confronted with the data, are often uncertain about the sign of the price elasticity of supplied quantity despite their a priori beliefs, especially in situations where there is a nontrivial probability that the slope of a supply curve is close to be horizontal or vertical. In our money market example, the argument for upward sloping supply is based on daily or weekly activities (Sims 1986; Bernanke and Mihov 1998). Over a longer horizon such as monthly or quarterly data, however, the Federal Reserve will set the interest rate and influence money supply in response to dynamic changes in output, employment, and the general price. The money supply behavior on the quarterly basis will not, in general, be the same as the daily or weekly behavior. Even on the weekly basis, the Federal Reserve is always concerned with liquidity problems in the banking system. In such a situation, the Federal Reserve, despite the fall in the interest rate, may continue to increase the supply of money to secure the adequate liquidity.

There are many other examples of downward sloping supply. A loan supply in markets with imperfect information is one example and a backward bending labor supply in labor markets is another (Stiglitz and Weiss 1981; Varian 1978, Chapter 6). The point here is not about what kind of a priori belief one should have regarding the slope of the supply curve; the point is how researchers should, without pretending too much a priori knowledge, report the statistical reliability of the ML estimates when the likelihood function implies a good deal of uncertainty about the parameters in the structural equation. The results presented in Figure 2 are difficult to interpret because of two distinct behaviors implied by the distributions around the two modes.

The ambiguous interpretations can be eliminated, however, if we simply change the sign of a_{11} and a_{21} in the shaded area of Quadrant I above the dividing hyperplane. Such a sign change finds the hyperplane image that lies in the shaded area of Quadrant III below the dividing hyperplane. Clearly, the normalized likelihood below the hyperplane has the unique mode and the shift in the supply curve in Figure 4(a) moves in the same direction as Figure 4(b). Consequently, the equilibrium effects of the exogenous shift s_1 , as shown in Figure 5, always increase the quantity M and decrease the price R despite the slope of the supply curve. In contrast to Figure 4, the results in Figure 5 have a clear interpretation about the effects of s_1 and are consistent with the direction of the ML estimates as shown in Figure 1.

Having seen Figures 2 and 5, one would not adopt naive implementation of the standard normalization rule if this rule generates visible bimodal distributions. In some previous work of simultaneous modeling, researchers instead search for a sophisticated ad hoc rule to change the sign of a_{11} and a_{21} whenever the negative effect on M or the positive effect on R occurs (Sims and Zha 1998b; Christiano, Eichenbaum, and Evans 1999). Operationally, this rule is equivalent to changing the sign of the parameters in the shaded area of Quadrant I and finding the hyperplane image of these parameters in the shaded area of Quadrant III. Finding such a rule is not always easy in high dimensions implied by multiple equation models. In particular, when the parameters in the other columns of \mathbf{A} vary, the equilibrium effects on M and R may not be in the same direction of the ML estimates. Take the example in which the equilibrium response of R is negative while the response of M is negligible but nonetheless negative.⁸ In this case, the sign of the M response is inconsequential. But it is computationally very inefficient to manually keep track of such instances as all columns of \mathbf{A} are moving and then decide when to change the sign of the structural equation.

⁸ The reverse case is of small or little liquidity effect, as found in Leeper (1995), Cushman and Zha (1997), and Zha (1999).

In the following section we provide a general solution to normalization, including the standard rule for recursive systems as a special case. We show the existence of the zero-density dividing hyperplane in the parameter space of each column of \mathbf{A} while holding the other columns fixed. We develop a method guaranteeing that the normalized ML estimates and the normalized parameters drawn from the likelihood always lie on the same side of the hyperplane. The method preserves the shape of the normalized likelihood by eliminating the bimodal phenomenon, and maintains the direction of the equilibrium effects of an exogenous shift in the structural equation. As the other columns of \mathbf{A} move, such a likelihood-preserving procedure adjusts the zero-density dividing hyperplane accordingly.⁹ For the inference of the impact effect of an exogenous shift, furthermore, the method minimizes the distance between the normalized ML estimate and the normalized value drawn from the likelihood function.

4. The theory of likelihood preserving normalization

A hyperplane in \mathbf{R}^m is simply a linear subspace of dimension $m - 1$. Thus a hyperplane in \mathbf{R}^2 is a line and a hyperplane in \mathbf{R}^3 is a plane. The following proposition describes the zeros of the likelihood function in terms of hyperplanes.

Proposition 1. Given \mathbf{a}_i for $i \neq k$, the set of all $\mathbf{a}_k \in R_k$ such that likelihood function defined by (3) is zero at $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_k | \dots | \mathbf{a}_n]$ is either a hyperplane in R_k or all of R_k .

Proof. Let Γ_k be the set of all $\mathbf{a}_k \in R_k$ for which the value of the likelihood function at \mathbf{A} is zero¹⁰. Note that the likelihood function is zero if and only if \mathbf{A} is singular. If Γ_k is equal to R_k , then we are done. If not, then there exists $\mathbf{a}_k \in R_k$ such that the matrix \mathbf{A} is non-singular. This

⁹ Take the case of Figure 3. The line $a_{21} = 0$ is a good approximation to the zero-density dividing hyperplane. But when the other columns of \mathbf{A} vary, this line can become a poor approximation because it fails to adjust accordingly. In Waggoner and Zha (1997), they show the similarly ill-determined inferences when the $a_{21} > 0$ rule is used.

¹⁰ Though not explicit in our notation, Γ_k depends on \mathbf{a}_i for $i \neq k$.

implies that $\text{span}\{\mathbf{a}_i | i \neq k\}$ is of dimension $n-1$ and that Γ_k is equal to $R_k \cap \text{span}\{\mathbf{a}_i | i \neq k\}$, which is a hyperplane in R_k . QED

Because the restrictions are non-degenerate, the set of all $\prod_{i \neq k} \mathbf{a}_i \in \prod_{i \neq k} R_i$ such that Γ_k is all of R_k is of measure zero in $\prod_{i \neq k} R_i$. Thus we need only consider the case that Γ_k is a hyperplane in R_k . The hyperplane Γ_k divides R_k into two regions and the shape of the likelihood is identical and unimodal on each of these regions¹¹. A normalization will be likelihood preserving if normalized values of \mathbf{a}_k either lie on or on the same side of Γ_k .

Definition 1. A normalization rule is likelihood preserving (LP) if every normalized value of $\mathbf{a}_k \in R_k$ either lies on or on the same side of Γ_k .

The following definition makes precise the notion of two points being on the same side of a hyperplane and Proposition 2 and its corollary give easy conditions for testing this condition.

Definition 2. Two points are on the same side of a hyperplane if the line that connects them does not intersect the hyperplane.

Proposition 2. For $1 \leq i \leq n$, let $\mathbf{a}_i \in R_i$, with $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ non-singular. If $\hat{\mathbf{a}}_k = \sum_{i=1}^n \alpha_i \mathbf{a}_i$, then the vectors \mathbf{a}_k and $\hat{\mathbf{a}}_k$ lie on the same side of Γ_k if and only if α_k is positive.

Proof. Proceed by contradiction. The line between \mathbf{a}_k and $\hat{\mathbf{a}}_k$ is

$$t\hat{\mathbf{a}}_k + (1-t)\mathbf{a}_k = (1-t(1-\alpha_k))\mathbf{a}_k + \sum_{i \neq k} t\alpha_i \mathbf{a}_i,$$

for $0 \leq t \leq 1$. Since Γ_k is contained in $\text{span}\{\mathbf{a}_i | i \neq k\}$, the line between \mathbf{a}_k and $\hat{\mathbf{a}}_k$ will intercept Γ_k if and only if there exists t such that $1-t(1-\alpha_k) = 0$ or $t = 1/(1-\alpha_k)$. But $0 \leq 1/(1-\alpha_k) \leq 1$ if and only if $\alpha_k \leq 0$. QED.

¹¹ This follows easily from Theorem 2 in Waggoner and Zha [2000].

Proposition 2 completes the establishment of our LP normalization rule, which produces the zero-density dividing hyperplane Γ_k for the k th column of \mathbf{A} and selects the region that always contains both the ML estimate $\hat{\mathbf{a}}_k$ and the normalized value \mathbf{a}_k . The following corollary and algorithm provide an efficient method for implementing the LP normalization.

Corollary 1. The vectors \mathbf{a}_k and $\hat{\mathbf{a}}_k$ lie on the same side of the hyperplane Γ_k if and only if $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_k > 0$.

Proof. Note that $\hat{\mathbf{a}}_k = \sum_{i=1}^n \alpha_i \mathbf{a}_i$ if and only if $\mathbf{A}^{-1} \hat{\mathbf{a}}_k = [\alpha_1, \dots, \alpha_n]'$. The corollary follows directly from Proposition 2. QED.

Let R be the set of all $n \times n$ nonsingular matrices $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_n]$ such that $\mathbf{a}_i \in R_i$.

Algorithm 1. For $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_n] \in R$ and $1 \leq k \leq n$,

- (a) keep \mathbf{a}_k if $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_k > 0$ and replace \mathbf{a}_k with $-\mathbf{a}_k$ if $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_k < 0$;
- (b) if $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_k = 0$, successively compute $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_i$ for $i = 1, \dots, k-1, k+1, \dots, n$;
- (c) stop at the first i such that $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_i \neq 0$ and replace \mathbf{a}_k with $-\mathbf{a}_k$ if $\mathbf{e}'_k \mathbf{A}^{-1} \hat{\mathbf{a}}_i < 0$.

In Algorithm 1, steps (b) and (c) are needed merely for the mathematical completion so that the sign of \mathbf{a}_k is always uniquely determined even if $\hat{\mathbf{a}}_k$ happens to lie in the separating hyperplane Γ_k . In practice, however, this situation will not occur because the set of all $\mathbf{A} \in R$ such that $\hat{\mathbf{a}}_k \in \Gamma_k$ has measure zero. Hence, step (a) of Algorithm 1 is all we need to implement the LP normalization.

In the special case where \mathbf{A} is restricted to be triangular, step (a) of Algorithm 1 implies that the sign of \mathbf{a}_k is so chosen as to make

$$\frac{\hat{\mathbf{A}}(k, k)}{\mathbf{A}(k, k)} > 0.$$

This result accords with the standard practice for recursive systems: the ML estimate $\hat{\mathbf{A}}(k, k)$ is first normalized to be positive and the sign of the k th equation is then chosen so that the diagonal element after normalization is always positive.

The LP normalization, by definition, ensures zero density on the boundary of the normalized likelihood function. We now show that such normalization also leads to unambiguous causal analysis. Let $s_k > 0$ be the amount of an exogenous shift in the k th structural equation. The following proposition states that despite the parameter uncertainty in the equation, the equilibrium effects of s_k on each variable of \mathbf{y} under the LP normalization are always in the same direction.

Definition 3. Two vectors $\mathbf{u} = [u_1, \dots, u_n]'$ and $\mathbf{v} = [v_1, \dots, v_n]'$ point in the same direction if $u_i v_i \geq 0$, for $i = 1, \dots, n$.

Proposition 3. If

$$\mathbf{A}_{(j)} = [\mathbf{a}_1 | \dots | \mathbf{a}_{k-1} | \mathbf{a}_k^{(j)} | \mathbf{a}_{k+1} | \dots | \mathbf{a}_n], \quad j = 1, 2,$$

are LP normalized elements of R , then the effects under these structural parameters, given by $s_k \mathbf{e}'_k \mathbf{A}_{(j)}^{-1}$, are in the same direction.

Proof. Since $\mathbf{A}_{(j)}$ invertible, there exists a non-zero vector $\rho_k \in R_k$ such that ρ_k is perpendicular to $\text{span}\{\mathbf{a}_i | i \neq k\}$ and $\rho'_k \hat{\mathbf{a}}_k > 0$. By the corollary of Proposition 2, we have

$$0 < \mathbf{e}'_k \mathbf{A}_{(j)}^{-1} \hat{\mathbf{a}}_k \equiv \frac{\rho'_k \hat{\mathbf{a}}_k}{\rho'_k \mathbf{a}_k^{(j)}}, \quad \text{for } j = 1, 2.$$

This result implies that $\rho'_k \mathbf{a}_k^{(j)} > 0$ for $j = 1, 2$. Because

$$s_k \mathbf{e}'_k \mathbf{A}_{(j)}^{-1} = \frac{\rho'_k}{\rho'_k \mathbf{a}_k^{(j)}}, \quad \text{for } j = 1, 2,$$

and ρ_k is independent of j , the proof follows from Definition 3. QED.

The above analysis concerns the statistical inference of one structural equation given all other equations. To gauge the overall statistical reliability of the ML-estimated effects on economic variables to exogenous shifts, we now take into account the parameter uncertainty of all structural equations. As discussed before, the impact effect on the vector of variables \mathbf{y} of an exogenous shift in the k th equation is proportional to the k th row of \mathbf{A}^{-1} . The following proposition provides a powerful result about the relationship between the normalized \mathbf{A}^{-1} and the ML estimate $\hat{\mathbf{A}}^{-1}$.

Proposition 4. For $\mathbf{A} = [\mathbf{a}_1 | \dots | \mathbf{a}_n] \in R$, \mathbf{a}_k and $\hat{\mathbf{a}}_k$ lie on the same side of Γ_k ($k = 1, \dots, n$) if and only if

$$\|\mathbf{A}^{-1} - \hat{\mathbf{A}}^{-1}\|_{\hat{\Omega}^{-1}} \geq \|g(\mathbf{A})^{-1} - \hat{\mathbf{A}}^{-1}\|_{\hat{\Omega}^{-1}}$$

where $g(\mathbf{A})$ is the normalized value of \mathbf{A} and $\|\cdot\|_{\hat{\Omega}^{-1}}$ is defined as

$$\|[\mathbf{x}_1 | \dots | \mathbf{x}_n]\|_{\hat{\Omega}^{-1}} = \sum_{i=1}^n \mathbf{x}_i' \hat{\Omega}^{-1} \mathbf{x}_i, \quad (6)$$

in which $\hat{\Omega} = (\hat{\mathbf{A}}\hat{\mathbf{A}}')^{-1}$.

Proof. Let $g(\mathbf{A}) = [\xi_1 \mathbf{a}_1 | \dots | \xi_n \mathbf{a}_n]$ for $\xi_i = \pm 1$. Since \mathbf{A} is non-singular, the set $\xi_1 \mathbf{a}_1, \dots, \xi_n \mathbf{a}_n$ forms a basis for \mathbf{R}^n . Thus, there exists unique scalars $\alpha_1, \dots, \alpha_n$ such that

$$\hat{\mathbf{a}}_k = \sum_{i=1}^n \alpha_i \xi_i \mathbf{a}_i.$$

By Proposition 2, \mathbf{a}_k and $\hat{\mathbf{a}}_k$ lie on the same side of Γ_k if and only if $\alpha_k > 0$; otherwise $\hat{\mathbf{a}}_k \in \Gamma_k$, which is a trivial case. The proof follows once we show that $\alpha_k > 0$ if and only if

$$\|\mathbf{A}^{-1} - \hat{\mathbf{A}}^{-1}\|_{\hat{\Omega}^{-1}} \leq \|g(\mathbf{A})^{-1} - \hat{\mathbf{A}}^{-1}\|_{\hat{\Omega}^{-1}}.$$

The distance between $g(\mathbf{A})^{-1}$ and $\hat{\mathbf{A}}^{-1}$ is

$$\begin{aligned}
\left\| g(\mathbf{A})^{-1} - \hat{\mathbf{A}}^{-1} \right\|_{\hat{\Omega}^{-1}} &= \sum_{i=1}^n \left(\xi_i \mathbf{e}_i' \mathbf{A}^{-1} - \mathbf{e}_i' \hat{\mathbf{A}}^{-1} \right) \left(\hat{\mathbf{A}} \hat{\mathbf{A}}' \right) \left(\xi_i \mathbf{A}^{-1'} \mathbf{e}_i - \hat{\mathbf{A}}^{-1'} \mathbf{e}_i \right) \\
&= \sum_{i=1}^n \mathbf{e}_i' \left(\mathbf{A}^{-1} \hat{\mathbf{A}} \right) \left(\mathbf{A}^{-1} \hat{\mathbf{A}} \right)' \mathbf{e}_i - 2 \xi_i \mathbf{e}_i' \mathbf{A}^{-1} \hat{\mathbf{a}}_i + \mathbf{e}_i' \mathbf{e}_i.
\end{aligned}$$

The necessary and sufficient condition for minimizing this distance is that $\xi_k \mathbf{e}_k' \mathbf{A}^{-1} \hat{\mathbf{a}}_k \geq 0$. Thus,

$$0 < \xi_k \mathbf{e}_k' \mathbf{A}^{-1} \hat{\mathbf{a}}_k = \xi_k \mathbf{e}_k' \left(\sum_{i=1}^n \alpha_i \xi_i \mathbf{e}_i \right) = \xi_k^2 \alpha_k = \alpha_k, \quad (7)$$

which completes the proof. QED

In Proposition 4, the distance is weighted by the inverse of $\hat{\Omega}$ which is the ML estimate of the variance-covariance matrix of the residuals for \mathbf{y} .¹² This weight ensures that the minimized distance is invariant to variance changes introduced by not only the scale of the variables but any non-singular, linear transformation of the variables.

Proposition 5. The effects of the LP normalization are invariant to the transformation $\mathbf{z}_t = \mathbf{P}' \mathbf{y}_t$ where \mathbf{P} is an $n \times n$ non-singular matrix.

Proof. Substituting $\mathbf{z}_t = \mathbf{P}' \mathbf{y}_t$ into (2), the contemporaneous parameter matrix becomes

$\mathbf{H} = \mathbf{P}^{-1} \mathbf{A}$. Because the transformation is linear, the ML estimate of \mathbf{H} is $\hat{\mathbf{H}} = \mathbf{P}^{-1} \hat{\mathbf{A}}$. Let $\hat{\Sigma} = (\hat{\mathbf{H}} \hat{\mathbf{H}}')^{-1}$. We thus have

$$\begin{aligned}
\left\| \mathbf{H}^{-1} - \hat{\mathbf{H}}^{-1} \right\|_{\hat{\Sigma}^{-1}} &= \sum_{i=1}^n \left(\mathbf{e}_i' \mathbf{H}^{-1} - \mathbf{e}_i' \hat{\mathbf{H}}^{-1} \right) \left(\hat{\mathbf{H}} \hat{\mathbf{H}}' \right) \left(\mathbf{H}^{-1'} \mathbf{e}_i - \hat{\mathbf{H}}^{-1'} \mathbf{e}_i \right) \\
&= \sum_{i=1}^n \left(\mathbf{e}_i' \mathbf{A}^{-1} - \mathbf{e}_i' \hat{\mathbf{A}}^{-1} \right) \mathbf{P} \left(\mathbf{P}^{-1} \hat{\mathbf{A}} \hat{\mathbf{A}}' (\mathbf{P}')^{-1} \right) \mathbf{P}' \left(\mathbf{A}^{-1'} \mathbf{e}_i - \hat{\mathbf{A}}^{-1'} \mathbf{e}_i \right) \\
&= \sum_{i=1}^n \left(\mathbf{e}_i' \mathbf{A}^{-1} - \mathbf{e}_i' \hat{\mathbf{A}}^{-1} \right) \left(\hat{\mathbf{A}} \hat{\mathbf{A}}' \right) \left(\mathbf{A}^{-1'} \mathbf{e}_i - \hat{\mathbf{A}}^{-1'} \mathbf{e}_i \right) \\
&= \left\| \mathbf{A}^{-1} - \hat{\mathbf{A}}^{-1} \right\|_{\hat{\Omega}^{-1}}.
\end{aligned}$$

¹² This weighted measurement resembles that used in the generalized least squares method. A recent paper of Cogley and Nason (1994) suggests a similar measurement.

The last equality proves the proposition. QED

5. The empirical example revisited

We revisit the empirical example discussed in Section 3. When the $a_{11} > 0$ rule is used to implement normalization, the bimodal distributions of the equilibrium effects on the variables (Figure 2) are the primary source of the ill-determined probability bands displayed in Figure 1. The normalized value of the effect randomly drawn from the likelihood function is arbitrarily far away from the normalized ML estimate. Under the LP normalization, the bimodal conditional distributions are eliminated and consequently the normalized value of the effect is closest to the normalized ML estimate among all possible values produced by sign changes, even when the parameters in all equations are allowed to vary (Proposition 4).

Figure 6 displays the impact and dynamic effects of a one-standard-deviation shift in money supply under the LP normalization implemented by Algorithm 1. The .68 and .90 equal-tail probability bands take account of the parameter uncertainty in all the structural equations. In sharp contrast to Figure 1, the error bands around the ML estimates in Figure 6 provide an accurate measurement about how informative these estimates are. The positive response of M is persistent for the entire time horizon; The interest rate falls initially as a result of the liquidity effect, and rises in two years in anticipation of a rise in inflation; Output, employment, and investment all rise within two years, but the response of the price level has a lag of four years as gauged by both error bands. The error bands indicate the asymmetric, long-tail distributions of the effects on many variables in various time periods.

The .90 bands in Figure 6 indicate some small probability of the negative impact on M and the positive impact on R in the initial period. This phenomenon occurs partly because of the positive interest elasticity of money demand. There are situations in which a demand curve can be upward sloping, such as the demand for used-cars (Wilson 1979). In such a case, the probability bands give us an accurate measure of the effects due to upward sloping demand curves. In our example, it is difficult to justify an upward sloping demand for money despite the great uncertainty about the elasticity. The probability bands, therefore, give us a useful piece of information about how badly the model is specified in this dimension. The results shown in Figure 6 suggest that the probability of such a “misspecification” is quite small.

6. Conclusion

Despite an essential role in understanding the causal relationships among economic variables, multiple equation modeling is not yet widely practiced because of many difficulties associated with simultaneity. One difficult issue is normalization. The standard normalization rule designed to work for recursive systems may lead to misleading inferential results when applied to simultaneous systems. Previous empirical work has adopted ad hoc rules to avoid distorting the shape of the normalized likelihood. But ad hoc rules are limited in scope and may be difficult to find in highly simultaneous systems.

In this paper we have developed a general method, the LP normalization, that always preserves the shape of the likelihood after normalization for both recursive and simultaneous systems. The LP normalization is easy and inexpensive to implement, especially in large systems of multiple equations. It always maintains an unambiguous economic interpretation of inferential results about the impact effects of an exogenous shift in the structural equation. Because the LP-normalized values of these effects derived from the likelihood function are closest to their ML estimates among all other normalized values, the resulting statistical inference gives an accurate picture of how informative the estimates are. We hope that the general method developed here helps advance the application of multiple equation modeling in economic analysis.

References

- Barro, Robert J., 1977. "Unanticipated Money Growth and Unemployment in the United States," *American Economic Review* 67(2): 101-15.
- Bernanke, Ben S. and Ilian Mihov, 1998. "Measuring Monetary Policy," *Quarterly Journal of Economics* (August), 869-902.
- Chib, Siddhartha and Edward Greenberg, 1995. "Understanding the Metropolis-Hastings Algorithm," *The American Statistician* 49 (4), 327-335.
- Christiano, Lawrence J., and Martin Eichenbaum, and Charles Evans, 1999. "Monetary Policy Shocks: What Have We Learned and To What End?" in John Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics*. Amsterdam, New York: Elsevier Science.
- Cogley and Nason, 1994. "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics* 9, S37-S70.
- Cushman, David O. and Tao Zha, 1997. "Identifying Monetary Policy in a Small Open Economy Under Flexible Exchange Rates," *Journal of Monetary Economics* 39, 433-448.
- Dent, Warren and John Geweke, 1980. "On Specification in Simultaneous Equation Models," in Jan Kmenta and James B. Ramsey (eds.), *Evaluation of Econometric Models*. Boston: Academic Press, Inc.
- Doan, Thomas A., 1992. *RATS User's Manual Version 4*, Estima.
- Geweke, John, 1995. "Monte Carlo Simulation and Numerical Integration," in H. Amman, D. Kendrick and J. Rust (eds.), *Handbook of Computational Economics*, Amsterdam: North-Holland.
- King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson, 1991. "Stochastic Trends and Economic Fluctuations," *American Economic Review* 81 (4), 819-840.
- Leeper, Eric M. and David B. Gordon, 1992. "In Search of the Liquidity Effect," *Journal of Monetary Economics* 29, 341-69.
- Leeper, Eric M., 1995. "Reducing Our Ignorance about Monetary Policy Effects," Federal Reserve Bank of Atlanta *Economic Review* 80 (4), 1-38.
- Leeper, Eric M., Christopher A. Sims, and Tao Zha, 1996. "What Does Monetary Policy Do?" *Brookings Papers On Economic Activity* 2, 1-63.
- Lieberman, Charles, "Structural and Technological Change in Money Demand," *American Economic Review* 69(2), 324-29.

- Marschak, Jacob, 1950. "Statistical Inference in Economics," *Studies in Econometric Method* (ed., William C. Hood and Tjalling C. Koopmans), New York: Wiley, 1-26.
- Pagan, Adrian and John Robertson, 1995. "Resolving the Liquidity Effect," *Federal Reserve Bank of St. Louis Review* 77 (3), 33-54.
- Poole, William, 1970. "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model," *Quarterly Journal of Economics* 84 (May), 197-216.
- Sargent, Thomas, J. and Neil Wallace, 1975. "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Monetary Supply Rule," *Journal of Political Economy* 83 (April), 241-54.
- Sims, Christopher A., 1986. "Are Forecasting Models Usable for Policy Analysis," *Quarterly Review of the Federal Reserve Bank of Minneapolis*, Winter.
- Sims, Christopher A. and Tao Zha, 1998a. "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review* 39 (4), 949-968.
- _____, 1998b. "Does Monetary Policy Generate Recession?" Federal Reserve Bank of Atlanta *Working Paper* 98-12 (July).
- _____, 1999. "Error Bands for Impulse Responses," *Econometrica* 67(5), 1113-1155.
- Stiglitz, Joseph E. and Andrew Weiss, 1981. "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71(3), 393-410.
- Tobin, James, 1970. "Money and Income: Post Hoc Ergo Propter Hoc?" *Quarterly Journal of Economics* 84(2): 301-17.
- Varian, Hal R., 1978. *Microeconomic Analysis*. 2nd ed. Norton & Company, New York.
- Waggoner, Daniel F. and Tao Zha, "Normalization, Probability Distribution, and Impulse Responses," Federal Reserve Bank of Atlanta *Working Paper* 97-11.
- _____, 2000. "A Gibbs Simulator for Restricted VAR Models," Federal Reserve Bank of Atlanta *Working Paper* 2000-3 (March).
- Wilson, Charles A., "Equilibrium and Adverse Selection," *American Economic Review* 69(2), 313-17.
- Zellner, Arnold, 1971. *An Introduction to Bayesian Inference in Econometrics*. John Wiley & Sons, New York.
- Zha, Tao, 1999. "Block Recursion and Structural Vector Autoregressions," *Journal of Econometrics* 90, 291-316.

Figure 1 Equilibrium effects of an exogenous shift in money supply under $a_{11} > 0$
(Basis points)

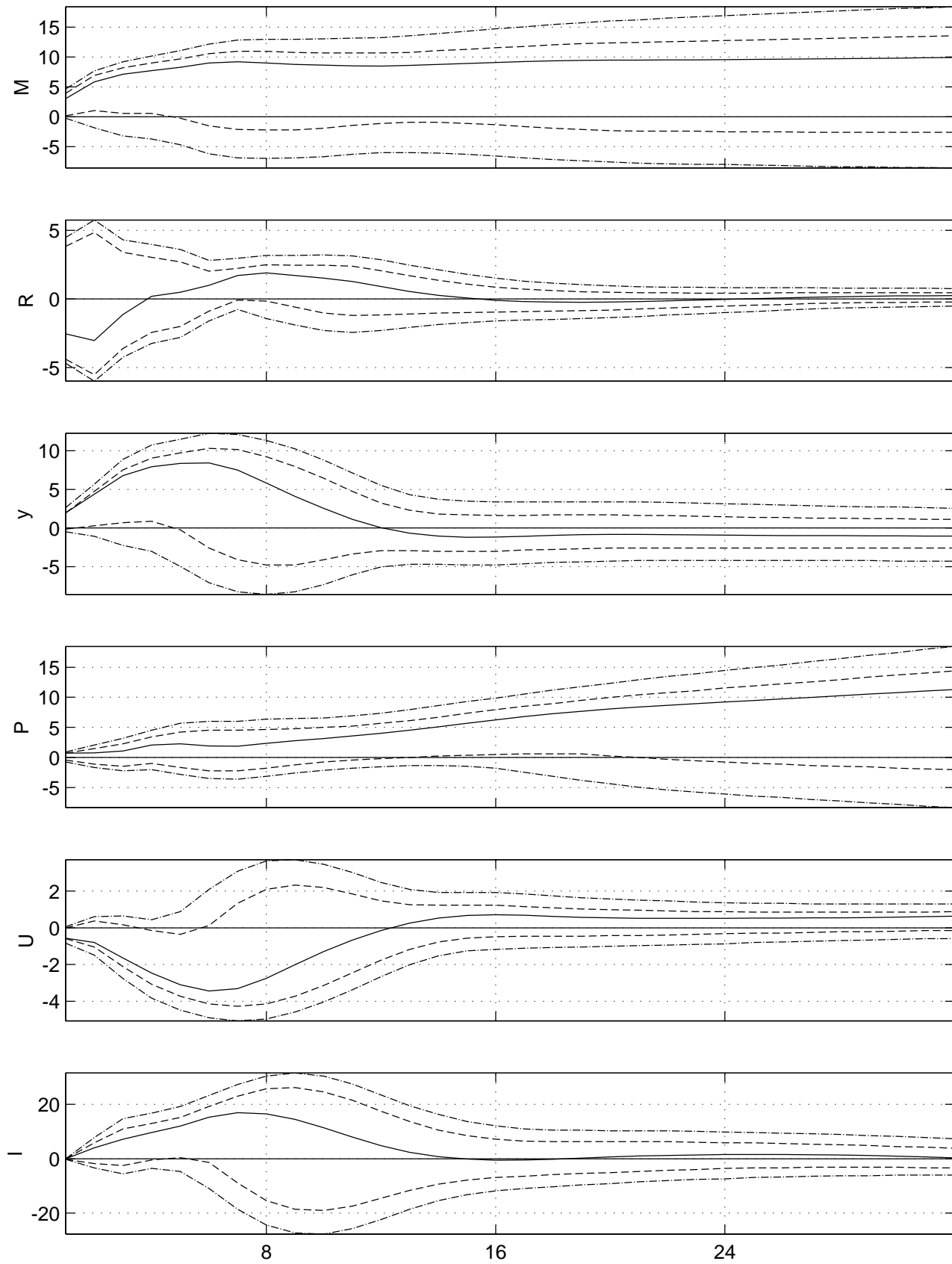


Figure 2 Equilibrium effects of an expansionary supply shift under $a_{ll} > 0$
(Basis points)

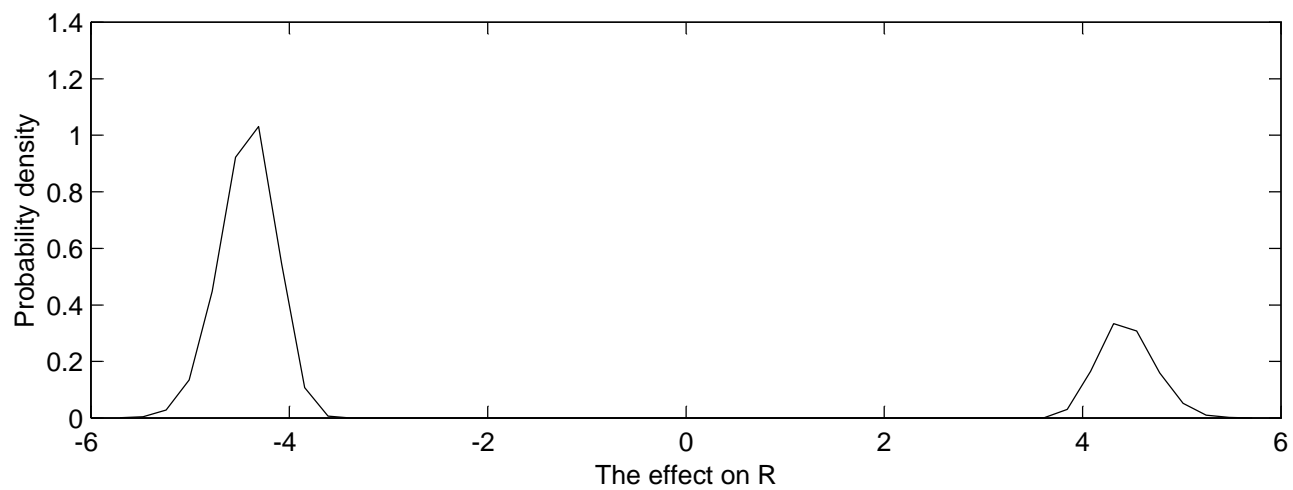
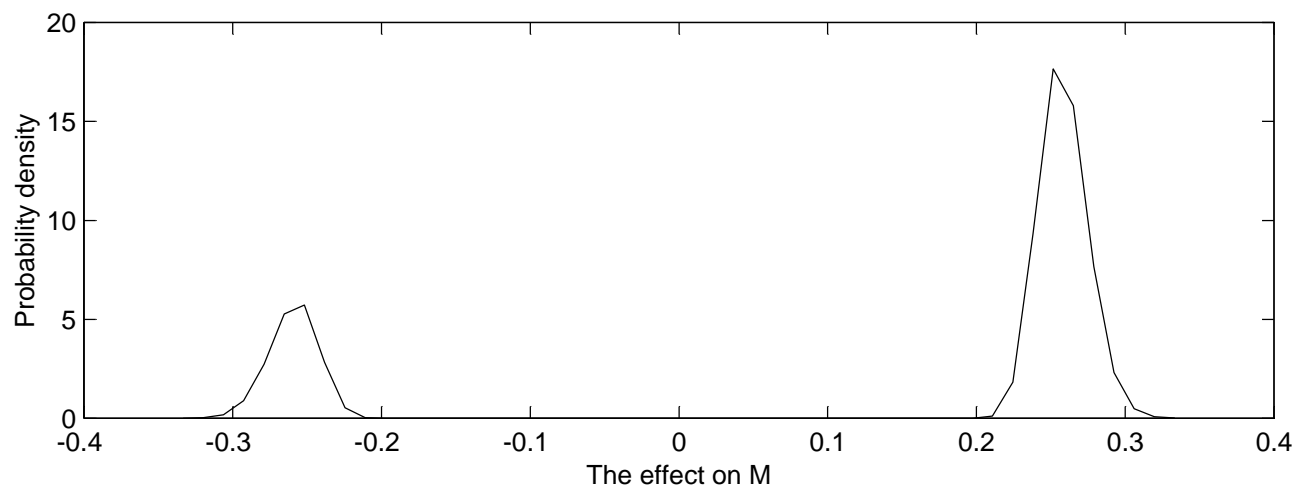


Figure 3 Likelihood contours in the simultaneous case

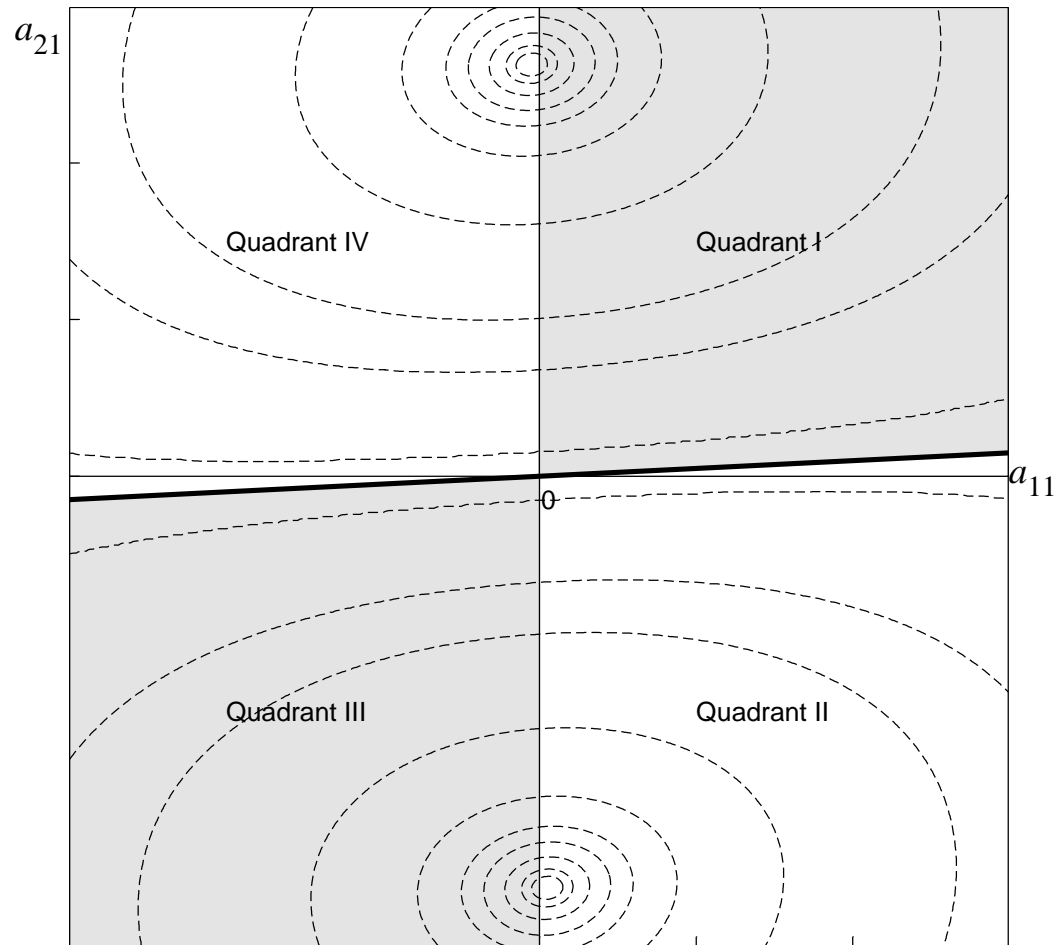


Figure 4 Equilibrium effects under $a_{II} > 0$

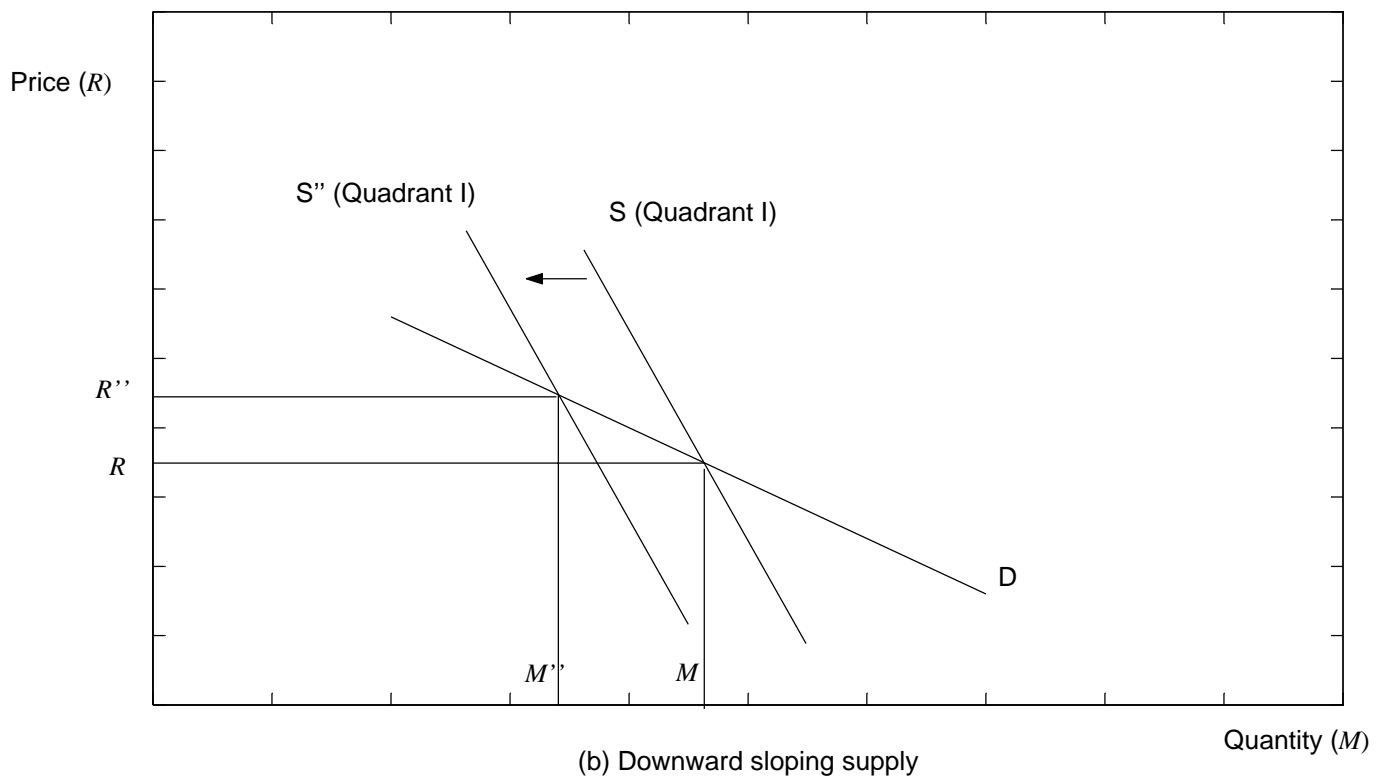
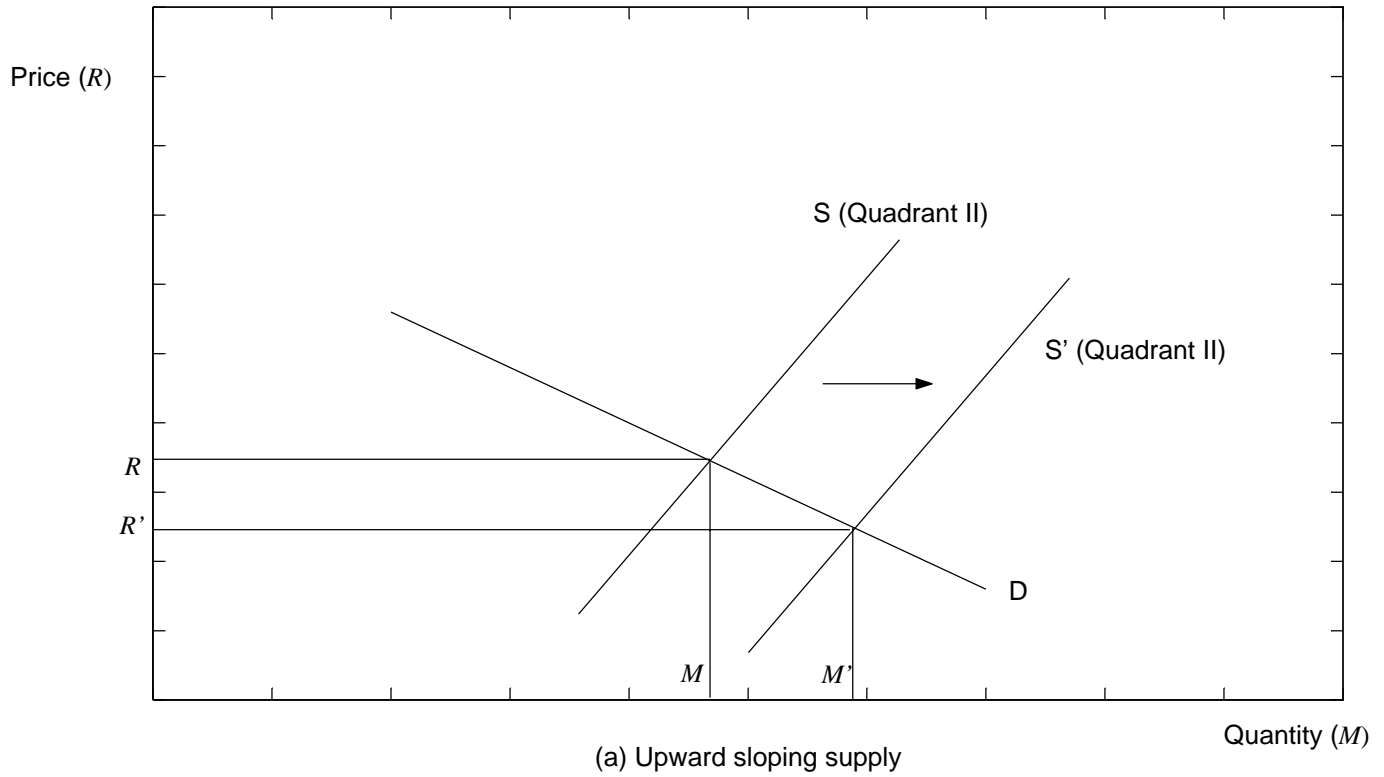


Figure 5 Equilibrium effects of an expansionary supply shift under new normalization
(Basis points)

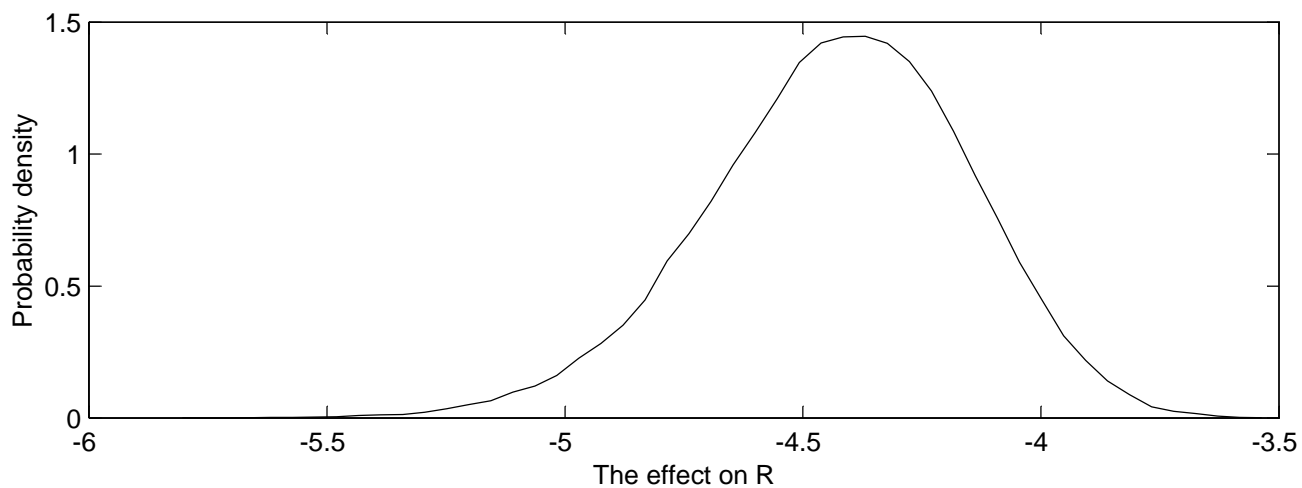
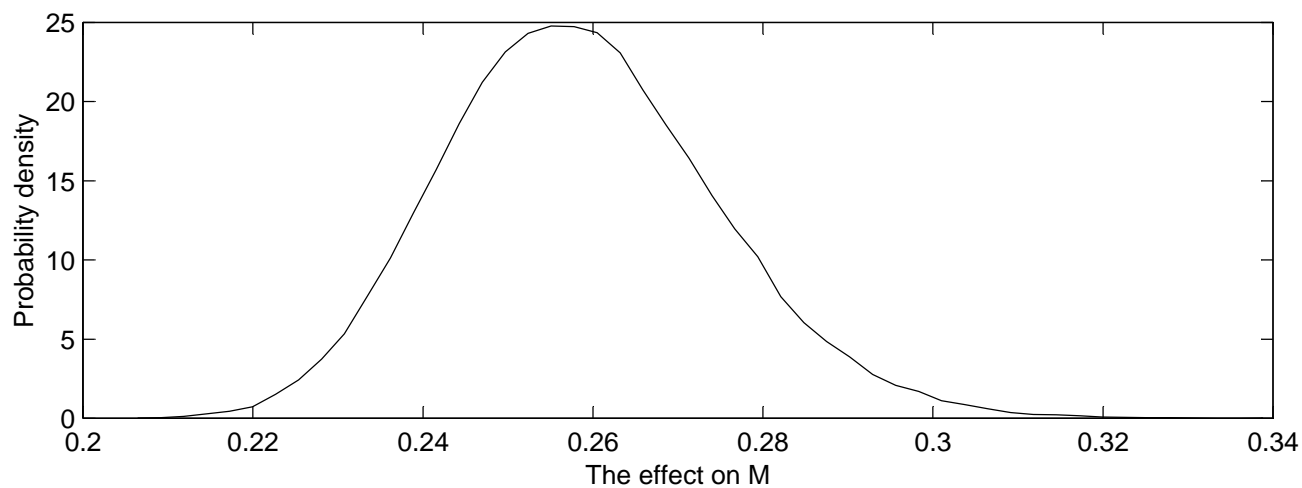


Figure 6 Equilibrium effects of an exogenous shift in money supply under LP
(Basis points)

