FISCAL FORESIGHT: ANALYTICAL ISSUES

ERIC M. LEEPER, TODD B. WALKER, AND SHU-CHUN SUSAN YANG

ABSTRACT. Fiscal foresight—the phenomenon that legislative and implementation lags ensure that private agents know the tax rates they face in the future—is intrinsic to the tax policy process. Although acknowledged in empirical work, theoretical analysis of its implications is scant. This paper develops an analytical framework to study the econometric implications of fiscal foresight. A simple example shows that foresight produces equilibrium time series with a non-invertible moving average component, which misaligns the agents' and the econometrician's information sets in estimated VARs. Economically meaningful shocks to taxes, therefore, cannot be easily extracted from statistical innovations. Non-invertibility arises as a natural outgrowth of the fact that optimal decisions discount future tax obligations; non-invertibility, therefore, is likely to be endemic to the study of fiscal policy.

1. INTRODUCTION

Fiscal policy presents researchers with a unique empirical challenge: how to identify and quantify the impacts of foreseen "shocks" to taxes. The challenge posed by taxes is unique because no other economic phenomenon so clearly signals to economic agents how important economic margins will change in the future. Intrinsic to the process of changing taxes are two kinds of lags: the legislative lag between when new tax law is proposed and when it is passed and the implementation lag between when the legislation is signed into law and when it actually takes place. Estimates of the total lag range from a couple of months to several years, depending on the particular legislation being considered.

Public finance economists recognize the possibility of fiscal foresight and have accumulated empirical evidence of its importance using a variety of econometric and event-study techniques.[1] Macroeconomists sometimes acknowledge the possibility of fiscal foresight in empirical work and occasionally study it in theoretical models, but the empirics are typically not grounded in theory. This paper is the first analytical study of the econometric implications of fiscal foresight. Theory suggests that fiscal foresight poses a substantial challenge to econometric analysis of fiscal policy.

Date: February 15, 2008.

[1]Evidence of foresight leading up to the Tax Reform Act of 1986 is documented in Auerbach and Slemrod (1997) and Burman, Clausing, and O'Hare (1994).

Two lines of attack on fiscal foresight appear in the empirical literature. The first estimates conventional VARs, identified in a variety of creative ways to isolate "anticipated taxes," and then examines the impacts of fiscal foresight *ex post* [Sims (1988), Blanchard and Perotti (2002), Mountford and Uhlig (2005), Yang (2007b)]. A second line rejects VAR identification schemes *ex ante*, arguing that they cannot adequately measure the impacts of foreseen changes in fiscal policy, and takes a different—narrative—approach to identification [Ramey and Shapiro (1998), Edelberg, Eichenbaum, and Fisher (1999), Ramey (2007), Romer and Romer (2007a), Mertens and Ravn (2008)].

This paper argues that fiscal foresight cannot be confronted *ex post*. Even very creative identification schemes are unlikely to correctly extract the tax news in agents' information sets from the information embedded in conventional VARs. In addition, *ex ante* approaches, while correctly skeptical of the efficacy of VAR methods, tend to achieve identification through a variety of heroic—and often implicit—identifying assumptions. Once those assumptions are laid bare, it is easy to be equally skeptical of both empirical approaches to fiscal foresight.

Fiscal foresight poses a formidable challenge because, as Yang (2005) shows, it generates an equilibrium with a non-invertible VARMA representation. Non-invertibility, in turn, implies that the fundamental shocks to tax policy cannot be recovered from current and past observable data, a central assumption of conventional econometric methods. This difficulty was pointed out in the early rational expectations econometric literature by Hansen and Sargent (1980, 1991) and recently emphasized by Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007). To the extent that fiscal foresight is an intrinsic feature of the tax policy process, conventional fiscal VARs ascribe to the econometrician an information set that is strictly smaller than the information set on which agents base their decision. A smaller information set can lead the econometrician to label as "tax shocks" objects that are linear combinations of all the exogenous disturbances at various leads and lags.

We present a simple analytical example of how fiscal foresight affects the nature of equilibrium time series and use the example to examine the nature of the problems that foresight creates for econometric analysis. As the analytical example makes clear, the non-invertibility that arises in the presence of fiscal foresight is a natural by-product of the fact that optimal decisions discount future tax obligations. Non-invertibility, therefore, is a general problem in the analysis of fiscal policy.

To be more precise, let ε_t denote the vector of exogenous disturbances that the agents observe, with $\varepsilon_{\tau,t}$ the tax element of the shock vector. At time t, the agents' information set is Ω_t, the span of $\{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$. An econometrician who estimates a conventional VAR with macro variables identifies exogenous shocks ε_t^*, with the associated tax disturbance $\varepsilon_{\tau,t}^*$. The econometrician's information set at t is Ω_t^*, the span of $\{\varepsilon_t^*, \varepsilon_{t-1}^*, \ldots\}$. In the presence of tax foresight, Ω_t^* is strictly smaller than Ω_t. Analytical examples, moreover, show that typical VAR specifications are likely to obtain measures of $\varepsilon^*_{\tau,t}$ that are linear combinations of current and past realizations of *all* the exogenous shocks in ε. Foresight about tax changes can lead to econometric estimates of tax disturbances that confound taxes with other sources of disturbance and treat as "news" old information to which rational agents have already reacted.

The purpose of this paper is to use analytical models to show explicitly how foresight makes the linear representation of the equilibrium follow a non-invertible VARMA process. We derive the mapping between the true tax news that agents observe and the "tax shocks" that the econometrician identifies. The models illustrate the kinds of mistaken inferences an econometrician is likely to draw The analysis leads to a framework to explain why existing empirical approaches cannot resolve the foresight issue when estimating tax effects, despite the fact that they yield plausible results even when foresight is ignored completely. Companion papers develop the econometric approach that the analytical work implies and then execute the empirical analysis.

2. Evidence of Fiscal Foresight

Plenty of reduced-form and anecdotal evidence suggests that private agents possess and react to expected changes in taxes. Steigerwald and Stuart (1997) develop a residual-based test statistic, which allows comparison of predictions errors for econometric models under different assumptions about the period of tax foresight. By studying investment behavior around major tax legislation that altered corporate income tax rates, investment tax credits, or the deductions for the depreciation allowances, Steigerwald and Stuart conclude that firms had 4.5-month foresight for tax changes in 1954, 1962, and 1981, and 16.5-month foresight for the tax reform of 1986.

House and Shapiro (2004) use a neoclassical growth model with fiscal foresight to calibrate the macroeconomic effects of Bush's tax cuts in 2001 and 2003. They argue that the slow recovery from the 2001 recession, especially in the job market, can be attributed to the laws' phasing-in provisions, which create incentives for workers and firms to delay production.

The Tax Reform Act of 1986 offered public finance economists a rich laboratory in which to test hypotheses that economic agents may respond to news about tax changes before the changes are enacted. Reactions to tax foresight are most evident when policy changes are phased in, incorporate sunset provisions, or include a delay between enactment and effective dates. The Tax Reform Act of 1986, which was enacted in October 1986, repealed the capital gains exclusion from ordinary income, raising the maximum capital gains tax rate from an effective rate of 20 percent to 28 percent beginning in 1987. During the lag between enactment and effective dates, capital gains realizations jumped from \$172 billion in 1985

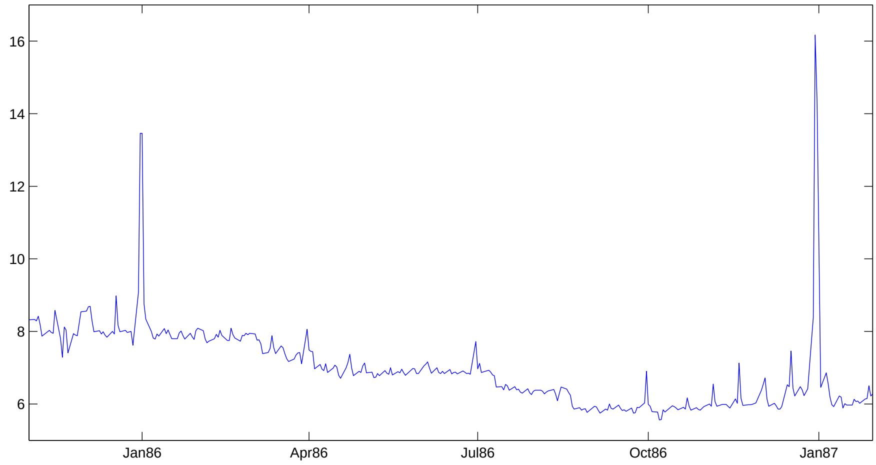

FIGURE 1. Daily Federal Funds Rate

to \$328 billion in 1986 in anticipation of the tax hike ((Congressional Budget Office, 2002, Table 1, p. 3)).

Auerbach and Slemrod (1997) thoroughly review the economic effects of the 1986 reform. They also point out the jump in long-term capital gains realizations that predated the effective date of the reform [see also Burman, Clausing, and O'Hare (1994)]. Auerbach and Slemrod argue that the 1986 Act "established once and for all that the timing of capital gains realizations behavior can be enormously sensitive to anticipated changes in the rate of taxation" [p. 605]. The act also reduced corporate tax rates from a maximum of 46 percent in 1986 to a maximum of 34 percent over 1987 and 1988. Scholes, Willson, and Wolfson (1992) find that the rate reduction induced firms to act in anticipation to defer revenue recognition or accelerate expense recognition.

Figure 1 documents two instances in which anticipations of the tax code changes associated with the Tax Reform Act of 1986 spilled over to affect monetary policy behavior. At the end of both 1985 and 1986, the federal funds rate spiked as traders shifted their portfolios to get ahead of expected restrictive tax legislation. On December 31, 1985 the funds rate rose almost 4.5 percentage points when, after the House of Representatives passed the legislation on December 17, traders placed higher probability that tax reform would take effect during 1986. As it happened, the bill was not enacted until October 22, 1986, to be effective on January 1, 1987. This is why the funds rate doubled on December 30, 1986, and remained high on the last day of 1986.[2] Evidently, the trading desk at the New York Fed underestimated the extent to which money markets would react to foresight about tax policy changes.

[2]Behavior in the reserves market is detailed in Federal Reserve Bank of New York (1986, 1987).

Evidence of the impacts of fiscal foresight extends well beyond a single piece of legislation. Yang (2007b) compares responses to a typical tax innovation from VAR systems with and without interest rates and prices. She finds that adding the three-month Treasury bill rate and the GDP price deflator (or commodity prices) to a VAR substantially reduces the responses of labor, investment, and output to a tax innovation. Since financial markets tend to be relatively sensitive to news in the economy, the result is consistent with the interpretation that information contained in financial variables reflects agents' knowledge of future tax policy changes. Sims (1988) and Garcia-Mila (1989) find similar results for government spending multipliers.

A well-documented example of behavior due to foresight underlies the controversy related to estimating the taxable income elasticity with respect to the tax hike for the high-income group that was embedded in the Omnibus Budget Reconciliation Act of 1993. Felstein and Feenberg (1996) examine the changes in income and tax data from 1992 to 1993 and find a large taxable income elasticity for the high-income group. Their finding implies that an increase in the marginal tax rate has a strong negative effect on the tax base. Goolsbee (2000), on the other hand, finds this elasticity drops to 0.4 if the data are extended to 1995. As the share of time-shiftable compensation (such as one-time bonuses, exercising stock options, and so forth) rose dramatically with income from 1991 to 1992, the contrasting estimates of a high short-run and low long-run elasticity arise because high-income taxpayers in 1992 acted upon their foresight about anticipated higher income tax rates.

Finally, anticipated shifts in government spending may also trigger changes in economic behavior before the spending shift is realized. Recent efforts by Ramey (2007) to reconcile the discrepancy in the effects of government spending using the narrative approach based on war dummies and VAR approaches lead her to conclude that the VAR-based innovations are indeed anticipated.[3] Since the war dates are identified based on the timing when the media reported future defense buildups, the news reports predate the actual rises in defense spending. Ramey finds that the war dates Granger-cause the VAR-based innovation for defense spending or aggregate government spending, but not vice versa. This supports the existence of foresight about government spending policy.

3. An Analytical Example

In this section, we specify a simple economic environment so that the econometric issues associated with tax foresight can be exposited analytically.

[3] See Ramey and Shapiro (1998) and Edelberg, Eichenbaum, and Fisher (1999) for empirical analysis of the dynamic impacts of war dummies.

3.1. **Model.** Consider a standard growth model with log preferences, inelastic labor supply, and complete depreciation of capital. A proportional tax is levied against income or, equivalently, output. The equilibrium conditions are well known and given by

$$\frac{1}{C_t} = \alpha\beta E_t(1-\tau_{t+1})\frac{1}{C_{t+1}}\frac{Y_{t+1}}{K_t} \tag{1}$$

$$C_t + K_t = Y_t = A_t K_{t-1}^{\alpha} \tag{2}$$

where C_t, K_t, and Y_t denote time-t consumption, capital, and output, respectively, and $\{A_t\}$ is an exogenous technology shock that obeys $A_t = e^{\varepsilon_{A,t}}$.[4] The government sets the tax rate according to a time-invariant rule and then adjusts lump-sum transfers to satisfy the constraint, $T_t = \tau_t Y_t$. Note that government spending is identically zero.

We imagine that news about tax rates arrives q periods before the news directly affects rates. For simplicity, we take tax rates to be exogenous and determined by the rule $\tau_t = \bar{\tau}e^{\varepsilon_{\tau,t-q}}$, or in log-linearized form

$$\hat{\tau}_t = \varepsilon_{\tau,t-q} \tag{3}$$

where $\hat{\tau}$ denotes percentage deviations of the tax rate from its steady state value. We assume the representative agent's information set at t consists of variables dated t and earlier, including the exogenous shocks, $\{\varepsilon_{A,t}, \varepsilon_{\tau,t}\}$. Given the tax rule in (3), this implies that the agent at t has (in this simple example, perfect) knowledge of $\{\hat{\tau}_{t+q}, \hat{\tau}_{t+q-1}, \ldots\}$.

After log linearizing equations (1)–(2), the equilibrium is characterized by a second-order difference equation in capital

$$\alpha\beta(1-\tau)E_t k_{t+1} - (1+\alpha^2\beta(1-\tau))k_t + \alpha k_{t-1} = \left\{1-\alpha\beta(1-\tau)\left(\frac{\tau}{1-\tau}\right)\right\}E_t\hat{\tau}_{t+1} - \varepsilon_{A,t}. \tag{4}$$

The solution to (4) satisfies the well-known saddlepath property and is given by

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\sum_{i=0}^{\infty}\theta^i E_t\varepsilon_{\tau,t-q+1+i}. \tag{5}$$

where $\theta = \alpha\beta(1-\tau)$. If there is no foresight, $q = 0$, we get the usual result that *i.i.d.* shocks to tax rates have no effect on capital accumulation. When there is some degree of tax foresight ($q > 0$), we obtain the unusual implication that even serially uncorrelated tax shocks affect capital accumulation (higher expected taxes lowers current capital).

It is useful to display the equilibrium for various degrees of fiscal foresight.

$q = 0$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t}. \tag{6}$$

[4]Appendix A generalizes this setup to satisfy any serial correlation property for the exogenous processes and demonstrates that the issues associated with foresight are generic.

$q = 1$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - [1 - \alpha\beta(1-\tau)]\left(\frac{\tau}{1-\tau}\right)\varepsilon_{\tau,t} \tag{7}$$

$q = 2$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - [1 - \alpha\beta(1-\tau)]\left(\frac{\tau}{1-\tau}\right)\left\{\varepsilon_{\tau,t-1} + \alpha\beta(1-\tau)\varepsilon_{\tau,t}\right\}. \tag{8}$$

$q = 3$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - [1 - \alpha\beta(1-\tau)]\left(\frac{\tau}{1-\tau}\right)\left\{\varepsilon_{\tau,t-2} + \alpha\beta(1-\tau)\varepsilon_{\tau,t-1} + [\alpha\beta(1-\tau)]^2\varepsilon_{\tau,t}\right\}. \tag{9}$$

In this model with inelastic labor supply, *i.i.d.* tax shocks should not affect capital accumulation. As expression (7) makes clear, when there is one period of foresight, tax shocks appear to affect capital contemporaneously. Invertibility problems appear to arise only when the foresight horizon is greater than one period. This result is not general and stems from the fixed labor. When labor choice is elastic, even one period of foresight can cause invertibilitiy problems.

An interesting implication of (8) and (9) is that more *recent* news is discounted (by $\alpha\beta(1-\tau)$) relative to more distant news. Why? Because $\varepsilon_{\tau,t-1}$ affects $\hat{\tau}_t$, while $\varepsilon_{\tau,t}$ affects $\hat{\tau}_{t+1}$, so the news that affects tax rates farther into the distance receives the heaviest discount.

3.2. **Econometrics of Foresight.** We now pose the following question that underlies the premise of this paper: What erroneous conclusions might an econometrician draw if they ignore foresight and proceed with the usual vector autoregression (VAR) analysis?

The main concern here is that with fiscal foresight the private agent's information set may be larger than the econometrician's information set. The Wold representation theorem tells us that the econometrician's information set from estimating a VAR will be the sigma algebra generated by current and past *observable* variables. In our simple setup, agents with fiscal foresight will observe tax shocks before the taxes are realized, suggesting the information set of the private agents will not coincide with information generated by observable variables. More specifically, suppose agents have 2-quarter foresight and the econometrician estimates a VAR employing current and past capital and technology, $\{k_{t-j}, a_{t-j}\}_{j=0}^{\infty}$,

$$\begin{bmatrix} a_t \\ k_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{-\kappa(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L} \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t} \\ \varepsilon_{A,t} \end{bmatrix}$$
$$\mathbf{y}_t = \mathcal{M}(L)\boldsymbol{\epsilon}_t. \tag{10}$$

where $\kappa = (1-\theta)(\tau/(1-\tau))$, $\theta = \alpha\beta(1-\tau)$ and L is a lag operator (i.e., $L^s x_t = x_{t-s}$). If this system is invertible in nonnegative powers of L, then $\boldsymbol{\epsilon}_t$ may be obtained as a square-summable linear combination of current and past $\mathbf{y}_t$. This suggests that knowledge of past and present $\mathbf{y}$ is equivalent to knowledge of past and present $\boldsymbol{\epsilon}$.[5] A necessary and sufficient condition for (10) to be a Wold (invertible) representation is that the determinant of $\mathcal{M}(z)$ be analytic with no zeros inside the open unit disk. By direct calculation, $\det \mathcal{M}(z) = \frac{\kappa}{1-\alpha z}(z+\theta)$ and hence this representation has a zero inside the unit circle (at $z = -\theta$). Thus fiscal foresight implies the information set generated by $\{\mathbf{y}_t, \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, ...\}$ is smaller than the information set generated by $\{\boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_{t-1}, \boldsymbol{\epsilon}_{t-2}, ...\}$.

It is useful to think about this invertibility condition for each shock process individually. By observing the process for technology, the econometrician will be able to recover current and past technology shocks $\{\varepsilon_{A,t}, \varepsilon_{A,t-1}, ...\}$. However, conditioning on current and past capital will not reveal current and past tax shocks.[6] The reason is that while the technology process is a Wold (invertible) representation, the process for capital is not. The solution for capital with 2-quarter foresight is given by

$$(1-\alpha L)k_t = -\kappa(L+\theta)\varepsilon_{\tau,t}$$

and note that invertibility of the stochastic process requires $|\theta| > 1$, so that

$$\left[\frac{1-\alpha L}{1+\theta^{-1}L}\right] k_t = -\kappa\theta\varepsilon_{\tau,t}$$

is a convergent sequence. But recall $\theta = \alpha\beta(1-\tau) < 1$ and thus the process is not invertible. Therefore, capital spans a *smaller* space than that generated by the fiscal shock. Foresight concerning fiscal policy introduces a non-invertible moving average representation into the equilibrium. This non-invertibility implies the econometrician's information set will be smaller than the agent's.

In order to determine the econometrician's information set, we need to find the corresponding Wold representation for capital. This is achieved by flipping the root of the moving average representation from inside the unit circle to outside the unit circle via the Blaschke factor, $[(L+\theta)/(1+\theta L)]$. The Wold representation for system (10) is given by

$$\underbrace{\begin{bmatrix} a_t \\ k_t \end{bmatrix}}_{\mathbf{y}_t} = \underbrace{\begin{bmatrix} 0 & 1 \\ \frac{-\kappa(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L} \end{bmatrix} \begin{bmatrix} \frac{1+\theta L}{L+\theta} & 0 \\ 0 & 1 \end{bmatrix}}_{\mathcal{M}^*(L)} \underbrace{\begin{bmatrix} \frac{L+\theta}{1+\theta L} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t} \\ \varepsilon_{A,t} \end{bmatrix}}_{\boldsymbol{\epsilon}_t^*}. \quad (11)$$

[5]More specifically, the Hilbert space generated by $\{\mathbf{y}_{t-j}\}_{j=0}^{\infty}$ is equivalent (in mean-square norm) to the Hilbert space generated by $\{\boldsymbol{\epsilon}_{t-j}\}_{j=0}^{\infty}$.

[6]In fact, one can show that no endogenous variable dated t and earlier nor any combination of variables will be able to recover current and past tax shocks.

and hence the Wold representation for capital is given by

$$\begin{aligned}(1-\alpha L)k_t &= \underbrace{-\kappa(L+\theta)\left[\frac{1+\theta L}{L+\theta}\right]}_{=\quad -\kappa(1+\theta L)}\underbrace{\left[\frac{L+\theta}{1+\theta L}\right]\varepsilon_{\tau,t}}_{\varepsilon^*_{\tau,t}} \\ &= -[1-\alpha\beta(1-\tau)]\left(\frac{\tau}{1-\tau}\right)\left\{\alpha\beta(1-\tau)\varepsilon^*_{\tau,t-1}+\varepsilon^*_{\tau,t}\right\} \end{aligned} \tag{12}$$

Several points are worth noting here. First, by observing current and past capital, the econometrician recovers $\varepsilon^*_{\tau,t}$, the statistical shocks of the VAR, which represent information that is "old news" to the agents of the economy. To see this more clearly note that

$$\begin{aligned}\varepsilon^*_{\tau,t} = \left[\frac{L+\theta}{1+\theta L}\right]\varepsilon_{\tau,t} &= (L+\theta)\sum_{j=0}^{\infty}-\theta^j\varepsilon_{\tau,t-j} \\ &= \theta\varepsilon_{\tau,t}+(1-\theta^2)\varepsilon_{\tau,t-1}-\theta(1-\theta^2)\varepsilon_{\tau,t-2}+\theta^2(1-\theta^2)\varepsilon_{\tau,t-3}+\cdots\end{aligned} \tag{13}$$

Therefore what the econometrician believes to be the tax innovation at time t, is in fact a discounted sum of tax shocks observed by the agents at date t and earlier. Thus the impulse response functions based upon $\varepsilon^*_{\tau,t}$ will not represent how *agents* respond to innovations in tax policy and will lead to incorrect inference. Second, comparing (12) with (8) one sees that the equilibrium based upon the econometrician's information set (i.e., ignoring foresight) discounts innovations incorrectly, $\alpha\beta(1-\tau)\varepsilon^*_{\tau,t-1}+\varepsilon^*_{\tau,t}$. This discounting rule implies that yesterday's innovation affects tomorrow's capital choice and date t innovations only have contemporaneous effects. However, as discussed above, agents with foresight discount innovations according to $\varepsilon_{\tau,t-1}+\alpha\beta(1-\tau)\varepsilon_{\tau,t}$ because yesterday's innovation affects contemporaneous capital accumulation and date t innovations are informative for tomorrow's capital tax rate. These first two points can be made clearer by examining the impulse response functions (IRF) based upon the agent's and econometrician's information sets given by Figure 2 ($\alpha = 0.36, \beta = 0.99, \tau = 0.25$).

With foresight, agents know exactly when the innovation in fiscal policy is going to translate into deviations in the tax rule. Hence the sharp decline in capital one quarter following the shock and output at two quarters (recall that the tax is levied on output which depends upon last quarter's capital stock). Conversely, the econometrician who ignores foresight yields a very different response to the tax shock. The discrepancy can be quite dramatic at short horizons; note that the difference in output between the two IRFs 1 quarter following the tax shock is more than a factor of 4.

Thus far, we have assumed that the econometrician is able to correctly identify the technology shock by conditioning on the technology process, a_t. However, it is more natural for the econometrician to condition on observable taxes (or some function of taxes such as

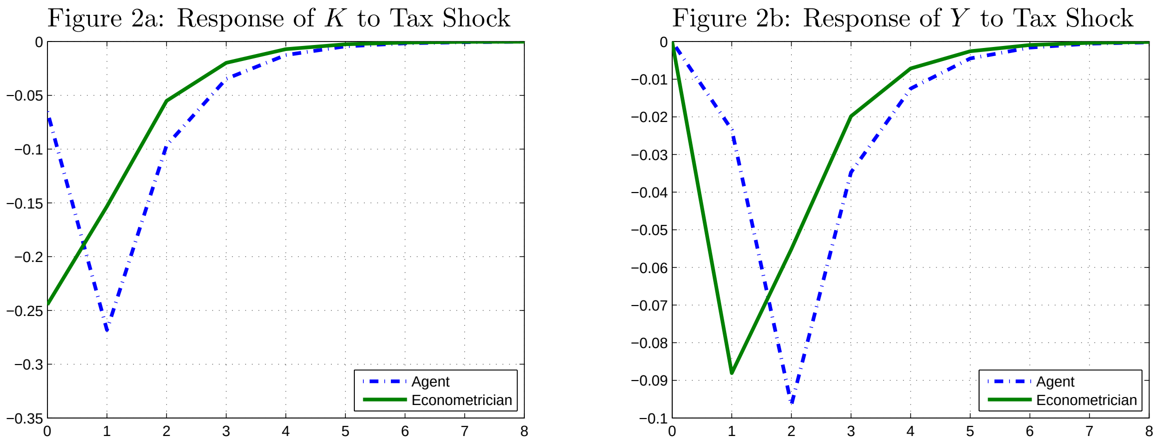

FIGURE 2. Impulse Response Functions Conditioning on Capital and Technology

revenues) as opposed to the technology process. Under this scenario, the econometric problems associated with fiscal foresight become even more pernicious. Consider the case with 2-quarter foresight and suppose the econometrician estimates the following VAR

$$\begin{aligned}\begin{bmatrix}\tau_t\\ k_t\end{bmatrix} &= \begin{bmatrix} L^2 & 0\\ \frac{-\kappa(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L}\end{bmatrix}\begin{bmatrix}\varepsilon_{\tau,t}\\ \varepsilon_{A,t}\end{bmatrix}\\ \mathbf{x}_t &= \mathcal{H}(L)\boldsymbol{\epsilon}_t.\end{aligned} \tag{14}$$

As with the previous case, foresight implies a zero inside the unit circle (now at $z = 0$), and hence the information set generated by $\{\mathbf{x}_t, \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ...\}$ is smaller than the information set generated by $\{\boldsymbol{\epsilon}_t, \boldsymbol{\epsilon}_{t-1}, \boldsymbol{\epsilon}_{t-2}, ...\}$. Following Hansen and Sargent (1991), the corresponding Wold representation for (14) is derived by flipping the zeroes outside the unit circle as in the previous example,

$$\begin{bmatrix}\tau_t\\ k_t\end{bmatrix} = \underbrace{\begin{bmatrix} L^2 & 0\\ \frac{-k(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L}\end{bmatrix} W\mathcal{B}(L)\tilde{W}\mathcal{B}(L)}_{\mathcal{H}^*(L)}\underbrace{\mathcal{B}(L^{-1})\tilde{W}'\mathcal{B}(L^{-1})W'\begin{bmatrix}\varepsilon_{\tau,t}\\ \varepsilon_{A,t}\end{bmatrix}}_{\boldsymbol{\epsilon}_t^*}$$
$$\mathbf{x}_t = \mathcal{H}^*(L)\,\boldsymbol{\epsilon}_t^* \tag{15}$$

where

$$W = \begin{bmatrix}\frac{1}{\sqrt{1+(\theta\kappa)^2}} & \frac{-\kappa\theta}{\sqrt{1+(\theta\kappa)^2}}\\ \frac{\kappa\theta}{\sqrt{1+(\theta\kappa)^2}} & \frac{1}{\sqrt{1+(\theta\kappa)^2}}\end{bmatrix}, \tilde{W} = \begin{bmatrix}\frac{1+\kappa^2\theta^2}{\mathcal{S}} & \frac{-\kappa}{\mathcal{S}}\\ \frac{\kappa}{\mathcal{S}} & \frac{1+\kappa^2\theta^2}{\mathcal{S}}\end{bmatrix}, \mathcal{B}(L) = \begin{bmatrix}L^{-1} & 0\\ 0 & 1\end{bmatrix}$$

and $\mathcal{S} = (1+2\kappa^2\theta^2+\kappa^4\theta^4+\kappa^2)^{1/2}$. The difference here is that the orthonormal W matrix must be employed to ensure the representation remains causal (i.e., preserving the assumption that the econometrician does not observe future values of the variables).

The statistical shocks recovered by the VAR analysis are now linear combinations of *both* technology and tax shocks. From (15), the statistical shocks can be written as

$$\varepsilon^*_{\tau,t} = c_1\varepsilon_{\tau,t-2} + c_2\varepsilon_{A,t-2}, \quad \varepsilon^*_{A,t} = c_3\varepsilon_{\tau,t} + c_4\varepsilon_{A,t}$$

where the c's are complicated functions of α, β and τ. The econometrician will not only misinterpret the response to a tax shock but also the response to a technology shock. It is important to emphasize that even though these statistical shocks are combinations of both economic shocks, the orthogonal nature of the Blaschke and W matrices ($WW' = \tilde{W}\tilde{W}' = I$ and $\mathcal{B}(L)\mathcal{B}(L^{-1}) = I$) ensures that the statistical shocks remain uncorrelated. To see this more clearly, assume that the covariance matrix for the economic shocks is the identity matrix and note that the covariance generating function for $\boldsymbol{\epsilon}^*_t$, given by

$$\mathcal{B}(z^{-1})\tilde{W}'\mathcal{B}(z^{-1})W'W\mathcal{B}(z)\tilde{W}\mathcal{B}(z),$$

is also equal to the identity matrix.[7] Hence, fiscal foresight is a "structural" problem and not a "reduced-form" problem. As we will see in the next section, even identification strategies that acknowledge the existence of fiscal foresight may lead to incorrect inference. Handling foresight econometrically can be very challenging for empirical researchers because the correlation structure of the econometrician's representation (15) is identical to that of the agents' (14) (i.e., the covariance generating functions are identical), but the conditional projections (or expectations) implied by the two representations are not identical. The presence of fiscal foresight suggests that interpreting IRFs for *any* shock sequence could be problematic.

3.3. **Practical Implications.** The nature of the econometric problem is such that important aspects of empirical research regarding fiscal policy can be misinformed. One such example is the short-run forecasts of macroeconomic aggregates in response to legislative changes. This is an important empirical issue if, as is now going on, the tax change is designed to operate at business cycle frequencies. Econometricians who do not adequately handle foresight may get the forecast of aggregates wrong, especially in the short run. Figure 2 demonstrates that most of the discrepancy between the IRFs of the agent and econometrician dissipates only after 6 quarters. This is due to the moving-average component of the equilibrium and the degree of foresight. As the degree of foresight increases, this discrepancy will become more prominent at longer horizons and the econometrician's short-run forecast will deviate further from the truth. Hence, it is essential that foresight be confronted when examining the response of macroeconomic aggregates to fiscal stimulus.

[7]If the covariance matrix for the economic shocks is not equal to the identity matrix, this result continues to hold but the factorization is slightly more complicated.

Short-run dynamics are also crucial for forecasting revenue consequences of proposed legislation. The PAYGO system was reestablished as a budget rule on January 4th 2007 by the 110th Congress. It requires that new spending legislation be "budget neutral" and places great importance on forecasting tax revenues over short horizons. Moreover, several states are required to balance budgets every fiscal year and rely heavily on accurate forecasts of revenues.

4. Background

Vector autoregressive (VAR) models have become the most important tool in applied macroeconomics since the seminal work by Sims (1980). The prevalent use of the VAR approach to identify monetary policy shocks has recently spurred interest to estimate the tax policy effects using the same framework. Unlike monetary policy making, tax policy changes in the U.S. often go through a lengthy legislative delay. Lags between policy announcements and enactment or implementation mean that private agents know a lot about their future tax liabilities over the next few quarters. This foresight suggests that the information structure underlying data is at variance with the one imposed by a VAR or a single-equation regression model, which assumes isolated tax shocks are unanticipated by private agents.

4.1. **Recognizing the Foresight Problem.** The foresight issue in estimating tax policy effects provides one example where current and past variables cannot recover meaningful structural shocks.[8] Sims (1988) first points out that when government spending decisions are known several quarters before they affect fiscal variables, the coefficient matrix of the linear model, which postulates that current observables are determined by current and past innovations, does not have a one-sided inverse. As a result, policy innovations cannot be constructed from current and past observables, as specified by a VAR or an autoregression model. Leeper (1989) formally demonstrates one possible empirical consequence if fiscal foresight is overlooked. Under active fiscal policy and passive monetary policy (in the sense of Leeper (1991)), when an econometrician identifies fiscal policy shocks through government spending and taxes, the Granger-causal ordering between deficits and money creation can be reserved; one can mistakenly attribute fiscal effects to monetary effects.

Hansen and Sargent (1991) emphasize accurate inference about the responses to economic shocks requires the equivalence of the private agents' information set to an econometrician's, which is characterized by the current and past variables in a VAR system. When estimating tax policy effects, private agents' information set contains not only current and past observables but also information about future tax policy changes because of foresight induced by the legislative process. Using simulated data from a dynamic stochastic general equilibrium

[8]See Lippi and Reichlin (1994) and Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) for other examples of macroeconomic models, which produce non-invertible VARMA structures.

model with tax policy foresight, Yang (2005) shows that structural VAR fails to recover the true tax effects even if correct policy rules are imposed to identify tax policy innovations.

4.2. **Empirical Work Estimating Fiscal Effects.** Despite the theoretical recognition of the difficulty that foresight creates for estimating fiscal policy effects, empirical work tends to deemphasize the issue. Some studies ignore the problem altogether (Fatas and Mihov (2001b) and Canzoneri, Cumby, and Diba (2001)). Some acknowledge that ignoring foresight can produce biased estimates but still pursue the estimation without considering it. And others recognize the issue but propose an unsatisfactory method to deal with foresight.

Perotti (2004) hypothesizes that if private agents make decisions based on policy announcements unobserved to an econometrician, then the estimated the VAR innovations should be correlated with those announcements, which are reflected in the OECD forecasts of government spending and output growth. His regression results find little evidence that VAR innovations are predictable by the forecast of these variables, with the exception of the residuals in government spending for the UK and in the tax revenues less transfers for the U.S. As the VAR innovations he uses are reduced-form residuals, which is a combination of all the shocks in the economy, it is not surprising that these mongrels do not appear to be correlated with the OECD forecasts. On the other hand, even if he has identified fiscal shocks, one still needs to confront with the foresight issue with the U.S. tax effects since a significant relationship is found between forecast and tax innovations.

Blanchard and Perotti (2002) (BP, hereafter) propose to include future fiscal variables in the output equations, and then use the recovered tax or spending policy shocks as instruments for these variables. There are two reasons why this approach might incorrectly handle foresight. First, BP implicitly assume that the statistical shocks recovered by the econometrician's VAR are the same shocks observed by the agents. They then argue that forwarding these statistical shocks ahead q periods–where q is the degree of foresight–serves to reconcile the discrepancy between the econometrician's and agents' information sets. However as we showed in Section 3, even the most simplistic economic environments generate statistical shocks that are convolutions of the true economic shocks. Forwarding (13) ahead q periods will not reconcile the information set of the agents with that of the econometrician. The BP method will continue to discount the shocks incorrectly. Given this, it is not surprising that BP find anticipated tax changes have no significant effect on output and the majority of their results are estimated under the assumption of unanticipated tax policy changes. Second, by their own logic, if agents have x quarters of fiscal foresight, an econometrician has to assume that fiscal authorities cannot respond to economic conditions for $x + 1$ quarters. This assumption is especially problematic given that the tax code adjusts automatically to the level of output, reflecting some degree of progressivity of tax rates via automatic stablizers. Moreover, Yang (2007a) finds that the average time elapsed between a proposal announcement by a President and enactment for major postwar U.S. income tax legislation is seven months.

If we use this as a measure of foresight length, we have to impose a probably unacceptable identification assumption that fiscal authorities cannot respond to economic conditions for three quarters to one year.

Another proposed solution by Mountford and Uhlig (2005) is to combine the sign restriction method and zero restrictions for identification in a VAR. They argue that anticipated fiscal policy changes can be identified by imposing zero restrictions on the responses on a fiscal variable for a few quarters, which reflects that the isolated policy shock is news about a change in future policy. In the case of anticipated taxes, their identification scheme, which prevents tax revenues from moving for four quarters, delivers an eccentric result that output (a fraction of the tax base) falls while tax revenues have no changes. The only possibility for this to occur is that tax rates rise automatically by an amount to completely offset the effect of a decreased tax base on revenues. Since the automatic stabilizers built in most countries' tax code would suggest that tax rates should fall when output falls, *a priori*, it is difficult to be persuaded that the proposed methodology resolves the foresight issue.

Given the concern that a VAR model may only recover anticipated changes in policy, the literature has also developed an alternative approach to estimate fiscal effects. Ramey and Shapiro (1998) identify the three arguably exogenous war or defense build-up events based on timing when the discussion about these events became heated in *news media*.[9] Edelberg, Eichenbaum, and Fisher (1999) incorporate these war dummies in a VAR framework. While government spending can increase for a reason unrelated to domestic economic conditions such as wars, it is difficult to apply the same logic for tax policy changes.

Romer and Romer (2007a,b) construct a thorough historical account of the U.S. tax legislation from 1945 to 2006, and classify tax policy into "exogenous" and "endogenous" changes based on the stated dominant motivations by fiscal authorities. The exogenous tax changes are those driven by a desire to raise the long-run economic growth rate and to reduce deficits. The most important tax cuts to stimulate long-run growth identified are: the 1964 Kennedy-Johnson tax cut, and 1981 Reagan tax cut, and the 2001, 2003 Bush tax cut, which followed the 1960-61 recession, the 1980 and 1981-82 recession, and the 2001 recession. While the reason of stimulating long-term economic growth can be emphasized for these tax cuts, it was often the case that they were inspired by a slow recovery from these recessions.[10] As shown by their results (figure 4 and panel b of figure 6), the unusual high output multiplier of 3 with respect to an "exogenous" tax cut is completely dominated by those tax cuts under a long-run growth motivation. Their paper also concerns about the foresight issue in the same econometric framework. Since their narrative approach probably cannot identify true exogenous tax changes, the results under anticipated future tax changes are equally questionable.

[9]Ramey (2007) recently adds the fourth event, 9/11.

[10]See Yang (2007a) for the motivation of the 1964 and 2003 tax cuts.

5. Conclusion

We have explicitly shown how fiscal foresight introduces econometric difficulties and makes interpreting VAR analysis difficult. The crux of the issue is that foresight introduces non-invertible moving average representations into the linear equilibrium process. We have shown how these noninvertible moving average representations imply that the mapping between the true tax news that agents observe and the "tax shocks" that the econometrician identifies can be vastly different. The models illustrate the kinds of mistaken inferences an econometrician is likely to draw We then demonstrated why the existing empirical approaches cannot resolve the foresight issue when estimating tax effects, despite the fact that they yield plausible results even when foresight is ignored completely. Future research will focus on developing the econometric techniques necessary to handle the issues implied by the analytical framework established here.

Appendix A. Generalization of the Illustrative Model

This section demonstrates that the issues associated with fiscal foresight extend to more generalized setups. The generalization shows that fiscal foresight will lead to a non-invertible moving average equilibrium regardless of the serial correlation properties of the exogenous driving force. To that end, suppose the tax rate and technology follow

$$\hat{\tau}_t = C(L)L^q \varepsilon_{\tau,t}, \quad a_t = A(L)\varepsilon_{A,t}$$

where $C(L)$ and $A(L)$ are polynomials in the lag operator L and q is the degree of foresight. The only restriction placed on $C(L)$ and $A(L)$ is that the corresponding coefficients are square summable, which allows for any serial correlation pattern. Agents will guess that the law of motion for capital is given by a square summable linear combination of technology and tax shocks, $k_t = D(L)\varepsilon_{A,t} + F(L)\varepsilon_{\tau,t}$. The resulting difference equation in capital can be written as

$$\begin{aligned}\theta\{L^{-1}[D(L) - D_0]\varepsilon_{A,t} + L^{-1}[F(L) - F_0]\varepsilon_{\tau,t}\} - (1+\alpha\theta)(D(L)\varepsilon_{A,t} + F(L)\varepsilon_{\tau,t}) \\ +\alpha L(D(L)\varepsilon_{A,t} + F(L)\varepsilon_{\tau,t}) = A(L)\varepsilon_{A,t} + \left\{1 - \theta\left(\frac{\tau}{1-\tau}\right)\right\}E_{t+1}\hat{\tau}_{t+1}\end{aligned} \tag{16}$$

where the Weiner-Kolmogorov formula is used to take expectations (i.e., $E_t x_{t+1} = L^{-1}[D(L) - D_0]\varepsilon_{x,t})$, and $\theta = \alpha\beta(1-\tau)$. Given that the shocks are uncorrelated, the difference equation may be solved in each shock individually. Note that $D(L)\varepsilon_{A,t}$ will not change with the degree of foresight, and is given by

$$D(L)\varepsilon_{A,t} = \left[\frac{[LA(L) - \theta A(\theta)]}{(1-\alpha L)(L-\theta)}\right]\varepsilon_{A,t}. \tag{17}$$

The solution for $F(L)\varepsilon_{\tau,t}$ for various degrees of foresight is given by
$q = 0$:

$$F(L)\varepsilon_{\tau,t} = -\left[\frac{(1-\theta)(\tau/(1-\tau))[C(L) - C(\theta)]}{(1-\alpha L)(L-\theta)}\right]\varepsilon_{\tau,t}$$

$q = 1$

$$F(L)\varepsilon_{\tau,t} = -\left[\frac{(1-\theta)(\tau/(1-\tau))[LC(L) - \theta C(\theta)]}{(1-\alpha L)(L-\theta)}\right]\varepsilon_{\tau,t}$$

$q = 2$

$$F(L)\varepsilon_{\tau,t} = -\left[\frac{(1-\theta)(\tau/(1-\tau))[L^2C(L) - \theta^2 C(\theta)]}{(1-\alpha L)(L-\theta)}\right]\varepsilon_{\tau,t}$$

Therefore the optimal amount of capital given q degrees of foresight is given by

$$(1-\alpha L)(L-\theta)k_t = [LA(L) - \theta A(\theta)]\varepsilon_{A,t} - (1-\theta)(\tau/(1-\tau))[L^q C(L) - \theta^q C(\theta)]\varepsilon_{\tau,t} \quad (18)$$

It is clear from these equations how foresight impinges upon the optimal amount of capital regardless of the exogenous driving force. Moreover, it is clear from (18) that for $q \geq 2$ the equilibrium will contain moving average representations regardless of the exogenous driving force.

References

Auerbach, A. J., and J. Slemrod (1997): "The Economic Effects of the Tax Reform Act of 1986," *Journal of Economic Literature*, 35(2), 589–632.

Blanchard, O. J., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117(4), 1329–1368.

Burman, L. E., K. A. Clausing, and J. F. O'Hare (1994): "Tax Reform and Realizations of Capital Gains in 1986," *National Tax Journal*, 47(1), 1–18.

Canzoneri, M. B., R. E. Cumby, and B. T. Diba (2001): "Is the Price Level Determined by the Needs of Fiscal Solvency?," *American Economic Review*, 91(5), 1221–1238.

Congressional Budget Office (2002): "Capital Gains Taxes and Federal Revenues," *Revenue and Tax Policy Brief.*

Edelberg, W., M. Eichenbaum, and J. D. M. Fisher (1999): "Understanding the Effects of a Shock to Government Purchases," *Review of Economic Dynamics*, 2(1), 166–206.

Fatas, A., and I. Mihov (2001a): "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," Mimeo, INSEAD.

——— (2001b): "Politica Fiscal y Ciclos Economicos: Una Investigacion Empirica. (Fiscal Policy and the Business Cycle: An Empirical Analysis.)," *Moneda y Credito*, (212), 167–205.

Federal Reserve Bank of New York (1986): "Monetary Policy and Open Market Operations in 1985," *Federal Reserve Bank of New York Quarterly Review*, 11(1), 34–53.

——— (1987): "Monetary Policy and Open Market Operations in 1986," *Federal Reserve Bank of New York Quarterly Review*, 12(1), 35–56.

Felstein, M., and D. R. Feenberg (1996): "The Effect of Increased Tax Rates on Taxable Income and Economic Efficiency: A Preliminary Analysis of the 1993 Tax Rate Increase," in *Tax Policy and the Economy*, ed. by J. Poterba, vol. 10, pp. 89–117. MIT Press, Cambridge.

Fernandez-Villaverde, J., J. F. Rubio-Ramirez, T. J. Sargent, and M. Watson (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97(3), 1021–1026.

Garcia-Mila, T. (1989): "Some Empirical Evidence on Government Purchase Multipliers," *Economics Letters*, 31, 375–380.

Goolsbee, A. (2000): "Its Not About the Money: Why Natural Experiments Dont Work on the Rich.," in *Does Atlas Shrug: The Economic Consequences of Taxing the Rich*, ed. by J. Slemord, pp. 141–158. Harvard University Press, Cambridge, MA.

Hansen, L. P., and T. J. Sargent (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2, 7–46.

——— (1991): "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 77–119. Westview Press, Boulder, CO.

House, C. L., and M. D. Shapiro (2004): "Phased-in Tax Cuts and Economic Activity," *NBER Working Paper*, No. 10415.

Leeper, E. M. (1989): "Policy Rules, Information, and Fiscal Effects in a 'Ricardian' Model," Federal Reserve Board, International Finance Discussion Paper No. 360.

——— (1991): "Equilibria Under 'Active' and 'Passive' Monetary and Fiscal Policies," *Journal of Monetary Economics*, 27(1), 129–147.

Lippi, M., and L. Reichlin (1994): "VAR Analysis, Nonfundamental Representations, Blaschke Matrices," *Journal of Econometrics*, 63(1), 307–325.

Mertens, K., and M. O. Ravn (2008): "The Aggregate Effects of Anticipated and Unanticipated U.S. Tax Policy Shocks: Theory and Empirical Evidence," Manuscript, European University Institute, ECO 2008/05.

Mountford, A., and H. Uhlig (2005): "What Are the Effects of Fiscal Policy Shocks?," SFB 649 Discussion Paper 2005-039, Humbolt University.

Perotti, R. (2004): "Estimating the Effects of Fiscal Policy in OECD Countries," IGIER Working Paper No. 276.

Ramey, V. A. (2007): "Identifying Government Spending Shocks: It's All in the Timing," Manuscript, University of California, San Diego.

Ramey, V. A., and M. D. Shapiro (1998): "Costly Capital Reallocation and the Effects of Government Spending," *Carneige-Rochester Conference Series on Public Policy*, 48, 145–194.

Romer, C. D., and D. H. Romer (2007a): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," Manuscript, University of California, Berkeley.

——— (2007b): "A Narrative Analysis of Postwar Tax Changes," Manuscript, University of California, Berkeley.

Scholes, M. S., G. P. Willson, and M. A. Wolfson (1992): "Firms' Responses to Anticipated Reducations in Tax Rates: The Tax Reform Act of 1986," *Journal of Accounting Research*, 30(Supplement), 161–185.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48(1), 1–48.

——— (1988): "Identifying Policy Effects," in *Empirical Macroeconomics for Interdependent Economics*, ed. by R. C. Bryant, D. W. Henderson, G. Holtham, P. Hooper, and S. A. Symansky, pp. 308–321. The Brookings Institution, Washington, DC.

Steigerwald, D. G., and C. Stuart (1997): "Econometric Estimation of Foresight: Tax Policy and Investment in the United States," *Review of Economics and Statistics*, 79(February), 32–40.

Yang, S.-C. S. (2005): "Quantifying Tax Effects Under Policy Foresight," *Journal of Monetary Economics*, 52(8), 1557–1568.

——— (2007a): "A Chronology of Postwar Federal Income Tax Policy," Center for Applied Economics and Policy Research Working Paper No. 2007-021.

——— (2007b): "Tentative Evidence of Tax Foresight," *Economics Letters*, 96(1), 30–37.