Assessing Structural VARs* (Preliminary)

Lawrence J. Christiano† Martin Eichenbaum‡ Robert Vigfusson§

May 24, 2005

Abstract

We use artificial data generated from variants of a simple real business cycle model to evaluate the ability of structural VARs to estimate the dynamic response of the economy to shocks. All the variants of the model economies considered in this paper imply that VAR-based methods that use short run restrictions are remarkably accurate. We also consider the performance of standard VAR-based estimators when long-run identifying restrictions are used. The parameterization of our model that is estimated by maximum likelihood implies that these methods also work well, in terms of bias and in terms of standard estimators of the degree of sampling variation. When we consider the models in Chari, Kehoe and McGrattan (2005), we confirm their finding that estimated impulse response functions based on long-run restrictions are distorted. We diagnose the reasons for the distortions, and build on our diagnosis to develop an improved estimator of impulse response functions based on long-run restrictions. It is not clear, however, whether the problems identified by CKM are of concern in practice. The CKM models are rejected overwhelmingly by the data.

*The first two authors are grateful to the National Science Foundation for Financial Suppport.
†Northwestern University and NBER.
‡Northwestern University and NBER.
§Federal Reserve Board of Governors.

1. Introduction

We argue that structural vector autoregressions (VARs) are useful as a guide to constructing and evaluating dynamic general equilibrium models. Given a minimal set of identifying assumptions, structural VARs allow the analyst to estimate the dynamic effects of economic shocks. These estimated response functions provide a natural way to assess the empirical plausibility of a structural model.[1] To be useful, the response function estimators must have good statistical properties. To assess these properties, we ask the following three questions. First, what are the bias properties of VAR-based estimators of response functions? Second, what are the bias properties of standard estimators of the sampling uncertainty associated with response function estimators? Third, are there easy to implement variants of standard procedures which improve the bias properties of response function estimators? We address these questions using data generated from a series of dynamic general equilibrium models.

We conclude that structural VARs do indeed provide valuable information for building empirically plausible models of aggregate fluctuations. Our analysis indicates that existing critiques of structural VARs are vastly overstated. Even in the worst case scenarios stressed by authors like Chari, Kehoe and McGrattan (2005) (CKM), the variants of standard VAR-based procedures that we develop virtually eliminate the small sample bias in estimates of dynamic response functions. Leaving aside the worst case scenarios, there is little small sample bias to correct.

There are two important traditions for constructing dynamic general equilibrium models. One tradition focusses on at most a handful of key shocks, and deliberately abstracts from the smaller shocks.[2] A classic example is Kydland and Prescott (1991), who work with a model driven only by technology shocks, even though they take the position that these shocks only account for 70 percent of business cycle fluctuations. A conundrum confronted by this modeling tradition is how to empirically evaluate models which contain only a subset of the shocks, with the data that are driven by *all* the shocks.[3] Structural VARs have the potential to provide a resolution to this challenge by allowing the analyst to assess the empirical performance of a model relative to a particular set of shocks.

[1] Two important early papers in this line of research are Sims (1980) and Sims (1989). There are many others papers in this tradition including Eichenbaum and Evans (1995), Rotemberg and Woodford (1997), Gali (1999), Francis and Ramey (2001), Christiano, Eichenbaum and Evans (2005), and Del Negro, Schorfheide, Smets, and Wouters (2005).

[2] The view that aggregate dynamics are dominated by the effects of a few shocks only, appears to receive confirmation from the literature on factor models. See for example Sargent and Sims (1977), Quah and Sargent (1993), Uhlig (1992), Forni, Giannone, Lippi and Reichlin (2004) and Giannone, Reichlin, and Sala (2005).

[3] Aiyagari (1994) and Prescott (1991) draw attention to the challenge by pointing to a difficulty with the standard RBC strategy for evaluating a model. In this strategy, one compares the second moment properties of the data with the second moment properties of the model. Prescott famously asserted that if a model matches the data, then that is bad news for the model. The argument is that since good models leave some things out of the analysis, good models should *not* match the data. Of course, lots of models do not match the data. This raises the question: how can we use the data to differentiate between good models and bad models?

A second tradition to building macroeconomic models incorporates large numbers of shocks in order to provide a complete characterization of the stochastic processes generating the data.[4] This tradition avoids the Kydland and Prescott conundrum. Still, for diagnostic purposes it is useful to assess the implications of these models for particular shocks to the economy.

In practice, the literature uses two types of identifying restrictions in structural VARs. Blanchard and Quah (1989), Gali (1999) and others have exploited the implications that many models have for the long-run effects of shocks.[5] Other authors have exploited short-run restrictions.[6] There is a growing literature that questions the ability of structural VARs to uncover the dynamic response of macroeconomic variables to structural shocks. This literature focuses on identification strategies that exploit long-run restrictions. Perhaps the first critique of these strategies was provided by Sims (1972). Although this paper was written before the advent of VARs, it articulates clearly why we should be concerned about the accuracy of identification based on long-run restrictions. To implement this strategy for identifying the consequences of shocks the analyst must have a reliable estimate of sums of coefficients in distributed lag regressions. These sums are hard to reliably estimate even if the individual coefficients are reasonably precisely estimated. Faust and Leeper (1997) and Pagan and Roberston (1998) make an important related critique of identification strategies based on long-run restrictions. More recently Erceg, Guerrieri and Gust (2004) and Chari, Kehoe and McGrattan (2005) (CKM) have also examined the reliability of VAR-based inference using long-run identifying restrictions. CKM are particularly critical and argue that structural VARs are very misleading.[7]

We examine the reliability of inference using structural VARs based on long-run and short-run identifying assumptions. Throughout, we suppose that the data generating mechanism corresponds to variants of a standard Real Business Cycle (RBC) model. We focus on the question, how do hours worked respond to a technology shock?

We find that structural VAR's perform remarkably well when identification is based on short-run restrictions. This is comforting for the vast literature that has exploited short-run identification schemes to identify the dynamic effects of shocks to the economy. Of course, one can question the particular short-run identifying assumptions used in any given analysis. But our results strongly support the view that if the relevant short-run assumptions are satisfied in

[4]See, for example, Smets and Wouters (2003) and Christiano, Motto and Rostagno (2004).

[5]See, for example, Basu, Fernald, and Kimball (2004), Christiano, Eichenbaum and Vigfusson (2003, 2004), and Francis and Ramey (2001)

[6]This list is particularly long and includes at least Blanchard and Watson (1986), Bernanke and Blinder (1992), Bernanke and Mihov (1995), Christiano and Eichenbaum (1992), Christiano, Eichenbaum and Evans (2005), Cushman and Zha (1997), Hamilton (1997) and Sims and Zha (1995).

[7]See also Fernandez-Villaverdez, Rubio-Ramirez and Sargent (2005) who analyze the circumstances under which the impulse response functions of infinite order VARs resemble impulse response functions associated with an economic model. They provide model-based conditions for checking whether the mapping from VAR shocks to economic shocks is invertible.

the data generating mechanism, then standard structural VAR procedures reliably uncover and identify the dynamic effects of shocks to the economy.

Regarding identification based on long-run assumptions, we find that if technology shocks account for a substantial fraction of business cycle fluctuations in output (say, over 50 percent), then structural VARs perform well. We do find some evidence of bias when the fraction of output variance accounted for by technology shocks is very small relative to estimates in the standard RBC literature. We develop and implement an adjustment to the standard VAR estimation strategy that virtually eliminates small sample bias, even in the worst case scenario when technology shocks play only a small role in aggregate fluctuations. The standard VAR strategy for implementing long-run identifying restrictions requires factorizing an estimate of the frequency-zero spectral density matrix of the data. The standard estimator used for this purpose is the zero-frequency spectral density implicit in the estimated VAR itself. Distortions arise in the worst case scenario because the quality of this estimator deteriorates. To deal with the problem, we adjust the standard VAR estimator by working with a Newey-West non-parametric estimator of the frequency-zero spectral density. The effect of this adjustment is minor when we are not in the worst-case scenario. However, in cases when standard VAR procedures entail some bias, our adjustment substantially reduces the bias.

Our conclusions regarding the value of identified VARs differ sharply from those recently reached by CKM. There are two primary reasons for why we reach different conclusions. First, CKM do not consider the case when VARs are identified using short-run restrictions. They only consider the case of long-run restrictions. Second, CKM's examples are based on exotic and empirically uninteresting data generating processes.

The CKM examples are exotic in the sense that they are based on an RBC model which has very different properties than those stressed in literature. In sharp contrast to Kydland and Prescott (1991), CKM's 'standard business cycle' model implies that technology shocks play a very small role in cyclical output fluctuations, roughly 20 percent. The remaining 80 percent of cyclical output fluctuations arise from shocks to the representative consumer's marginal utility of leisure.

This property would fine if it was defensible on empirical grounds. However, we show that CKM's parameterization is overwhelmingly rejected by the data. Our estimated version of the model has the property that technology shocks account for roughly 70% of cyclical fluctuations in output. In this empirically relevant case, standard VAR procedures lead to reliable inference about the effects of a technology shock.

It is always possible to identify *some* data generating process for which *any* econometric estimator has poor properties. All applied econometricians understand this. The hope of practitioners is that their estimators have good properties in the empirically relevant cases. The evidence suggests that CKM' examples do not satisfy this condition. Of course, a skeptic of

RBC theory might also question the empirical relevance of our results. Fortunately the good performance of structural VARs is not limited to the case when the data generating process is given by an RBC model. Altig, Christiano, Eichenbaum and Linde (2005) examine the reliability of identified VAR's when the data generating process is given by an estimated general equilibrium embodying nominal frictions as well as real and monetary shocks. There too we find that structural VARs perform well. Taken as a whole, our results provide strong support for the view that structural VARs are a useful guide for formulating and estimating business cycle models.

The remainder of this draft is organized as follows. Section 2 presents the general equilibrium model used in our examples. Section 3 discusses our results for standard VAR-based estimators of impulse response functions. Section 5 introduces and motivates our modified estimator of impulse response functions. Section 6 displays its operating characteristics. Section 7 evaluates the difference between our findings and those of CKM. Finally, section 7 contains concluding comments.

2. A Simple Real Business Cycle Model

In this section, we display the real business cycle model that serves as the data generating process in our analysis. The model has the property that the only shock that affects labor productivity in the long-run is a shock to technology. This property lies at the core of the identification strategy used by Gali (1999) and others to identify the effects of a shock to technology. We also consider a variant of the model in which we impose additional timing restrictions on agents' actions. In particular, we assume that agents choose hours worked before the technology shock is realized. This assumption allows us to identify the effects of a shock to technology using 'short-run restrictions', that is, restrictions on the variance-covariance matrix of the disturbances to a vector autoregression. We describe the conventional VAR-based strategies for estimating the dynamic impact on hours worked of a shock to technology. Finally, we discuss several parameterizations of our model that are used in the experiments we perform.

2.1. The Model

The representative agent maximizes expected utility over per capita consumption, c_t, and per capita employment, l_t

$$E_0 \sum_{t=0}^{\infty} \left(\beta\left(1+\gamma\right)\right)^t \left[\log c_t + \psi \frac{\left(1-l_t\right)^{1-\sigma}}{1-\sigma}\right],$$

subject to the budget constraint:

$$c_t + \left(1+\tau_{x,t}\right)\left[\left(1+\gamma\right)k_{t+1} - \left(1-\delta\right)k_t\right] \leq \left(1-\tau_{lt}\right)w_t l_t + r_t k_t + T_t.$$

Here, k_t denotes the per capita capital stock at the beginning of period t, w_t is the wage rate, r_t is the rental rate on capital, $\tau_{x,t}$ is an investment tax, τ_{lt} is the tax rate on labor, $\delta \in (0,1)$ is the depreciation rate on capital, γ is the net growth rate of the population, and T_t represents lump-sum taxes. Finally, $\sigma > 0$ is a curvature parameter.

The representative competitive firm's production function is:

$$y_t = k_t^{\alpha} (Z_t l_t)^{1-\alpha},$$

where Z_t is the time t state of technology and $\alpha \in (0,1)$. The stochastic processes for the shocks are:

$$\begin{aligned} \log z_t &= \mu_Z + \sigma_z \varepsilon_t^z \\ \tau_{lt+1} &= (1-\rho_l)\bar{\tau}_l + \rho_l \tau_{lt} + \sigma_l \varepsilon_{t+1}^l \\ \tau_{xt+1} &= (1-\rho_x)\bar{\tau}_x + \rho_x \tau_{xt} + \sigma_x \varepsilon_{t+1}^x, \end{aligned} \tag{2.1}$$

where $z_t = Z_t / Z_{t-1}$. In addition, ε_t^z, ε_t^d, and ε_t^x are independent random variables with mean zero and unit standard deviation. The parameters, σ_z, σ_l and σ_x are non-negative scalars. The constant, μ_Z, is the mean growth rate of technology, $\bar{\tau}_l$ is the mean labor tax rate, τ_x is the mean tax on capital. We restrict the autoregressive coefficients, ρ_l and ρ_x, to be less than unity in absolute value.

Finally, the resource constraint is:

$$c_t + (1+\gamma) k_{t+1} - (1-\delta) k_t \leq y_t.$$

We consider two versions of the model, differentiated according to timing assumption. In the *standard version,* all time t decisions are taken after the realization of the time t shocks. This is the conventional assumption in the real business cycle literature. For pedagogical purposes, we also consider a second version of the model, we call the *recursive version* of the real business cycle model. Here, the timing assumptions are as follows. First, τ_{lt} is observed, after which labor decisions are made. Next, the other shocks are realized. Then, agents make their investment and consumption decisions. Finally, labor, investment, consumption, and output occur. We first discuss the standard version of our model.

2.1.1. The Standard Version of the Model

The log-linearized policy rule for capital can be written as follows:

$$\log \hat{k}_{t+1} = \gamma_0 + \gamma_k \log \hat{k}_t + \gamma_z \log z_t + \gamma_l \tau_{lt} + \gamma_x x_t,$$

where $\hat{k}_t \equiv k_t / Z_{t-1}$. The policy rule for hours worked is:

$$\log l_t = a_0 + a_k \log \hat{k}_t + a_z \log z_t + a_l \tau_{lt} + a_x \tau_{xt}.$$

>From this expression, it is clear that all shocks have only a temporary impact on l_t and $\hat{k}_t$. Since ε_t^z is the only shock that has a permanent effect on Z_t, it follows that ε_t^z is the only shock that has a permanent impact on the level of the capital stock, k_t. Similarly, ε_t^z is the only shock that has a permanent impact on output and labor productivity, $a_t \equiv y_t/l_t$. Formally, this *exclusion restriction* is given by:

$$\lim_{j\to\infty} \left[E_t a_{t+j} - E_{t-1}a_{t+j}\right] = f\left(\varepsilon_t^z \text{ only}\right), \tag{2.2}$$

where in our linear approximation to the model solution, f is a linear function. The model also implies the *sign restriction* that f is an increasing function. In (2.2), E_t is the expectation operator, conditional on $\Omega_t = \left(\log \hat{k}_{t-s}, \log z_{t-s}, \tau_{l,t-s}, \tau_{x,t-s};\ s \geq 0\right)$. The exclusion and sign restrictions have been used by Gali (1999) and others to identify the dynamic impact on macroeconomic variables of a positive shock to technology.

In practice, researchers impose the exclusion and sign restrictions on a vector autoregression to compute ε_t^z and identify its dynamic effects on macroeconomic variables. To describe this procedure, denote the variables in the VAR by Y_t :

$$\begin{aligned} Y_{t+1} &= B(L)\, Y_t + u_{t+1},\ \ E u_t u_t' = V, \\ B(L) &\equiv B_1 + B_2 L + ... + B_p L^{p-1}, \\ Y_t &= \begin{pmatrix} \Delta \log a_t \\ \log l_t \\ x_t \end{pmatrix}, \end{aligned} \tag{2.3}$$

where x_t is an additional vector of variables that may be included in the VAR. It is assumed that the fundamental economic shocks are related to u_t in the following way:

$$u_t = C\varepsilon_t,\ E\varepsilon_t\varepsilon_t' = I,\ \ CC' = V, \tag{2.4}$$

where the first element in ε_t is ε_t^z. It is easy to verify that:

$$\lim_{j\to\infty} \tilde{E}_t[a_{t+j}] - \tilde{E}_{t-1}[a_{t+j}] = \tau\left[I - B(1)\right]^{-1} C\varepsilon_t, \tag{2.5}$$

where τ is a row vector with all zeros, except unity in the first location. Here, $B(1)$ is the sum, $B_1 + ... + B_q$. Also, $\tilde{E}_t$ is the expectation operator, conditional on $\tilde{\Omega}_t = \{Y_t, ..., Y_{t-q+1}\}$. To compute the dynamic effects of ε_t^z, we require $B_1, ..., B_q$ and C_1, the first column of C.

The symmetric matrix, V, and the B_i's can be computed by an ordinary least squares regression. However, it is well known that the requirement $CC' = V$ is not sufficient to determined a unique value of C_1. Adding the exclusion and sign restrictions does uniquely determine C_1. These restrictions are:

$$\text{exclusion restriction: } \left[I - B(1)\right]^{-1} C = \begin{bmatrix} \text{number} & \underset{1\times(N-1)}{0} \\ \text{numbers} & \text{numbers} \end{bmatrix},$$

and

$$\text{sign restriction: } (1,1) \text{ element of } [I - B(1)]^{-1} C \text{ is positive.}$$

Although there are many matrices, C, that satisfy $CC' = V$ as well as the exclusion and sign restrictions, they all have the same C_1. To see this, first let $D \equiv [I - B(1)]^{-1} C$, so that

$$DD' = [I - B(1)]^{-1} V [I - B(1)']^{-1} = S_0, \tag{2.6}$$

say. Note that S_0 (the spectral density of Y_t at frequency zero) can be computed directly from the VAR coefficients and the variance-covariance matrix of the VAR disturbances. The exclusion restrictions require that D have the following structure:

$$D = \begin{bmatrix} \underset{1\times 1}{d_{11}} & \underset{1\times(N-1)}{0} \\ \underset{(N-1)\times 1}{D_{21}} & \underset{(N-1)\times(N-1)}{D_{22}} \end{bmatrix}.$$

Then,

$$DD' = \begin{bmatrix} d_{11}^2 & d_{11}D_{21}' \\ D_{21}d_{11} & D_{21}D_{21}' + D_{22}D_{22}' \end{bmatrix} = \begin{bmatrix} S_0^{11} & S_0^{21\prime} \\ S_0^{21} & S_0^{22} \end{bmatrix},$$

say. The sign restriction is:

$$d_{11} > 0. \tag{2.7}$$

Then, the first column of D is uniquely determined by:

$$d_{11} = \sqrt{S_0^{11}},\ D_{21} = S_0^{21}/d_{11}$$

Finally, the first column of C is determined from:

$$C_1 = [I - B(1)]\, D_1. \tag{2.8}$$

2.1.2. The Recursive Version of the Model

In the recursive version of the model, the policy rule for labor involves $\log z_{t-1}$ and x_{t-1} because they help forecast $\log z_t$ and x_t :

$$\log l_t = a_0 + a_k \log \hat{k}_t + \tilde{a}_l \tau_{lt} + \tilde{a}_z' \log z_{t-1} + \tilde{a}_x' x_{t-1}.$$

Because labor is a state variable at the time the investment decision is made, the policy rule for $\hat{k}_{t+1}$ takes the following form:

$$\begin{aligned} \log \hat{k}_{t+1} &= \gamma_0 + \gamma_k \log \hat{k}_t + \tilde{\gamma}_z \log z_t + \tilde{\gamma}_l \tau_{lt} + \tilde{\gamma}_x x_t \\ &\quad + \tilde{\gamma}_z' \log z_{t-1} + \tilde{\gamma}_x' x_{t-1}. \end{aligned}$$

It is easy to verify that these policy rules satisfy the exclusion restriction, (2.2), and the sign restriction on ε_t^z. So, the long-run identification strategy outlined above can be rationalized in this model. An alternative procedure for identifying ε_t^z that does not rely on estimating long-run responses to shocks can also be rationalized. We refer to this as the 'short-run' strategy, because it involves recovering ε_t^z using just the realized one-step-ahead forecast errors in labor productivity and hours, as well as the second moment properties of those forecast errors. According to the model, the error in forecasting a_t given Ω_{t-1}, denoted by $u_{\Omega,t}^a$, is a linear combination of ε_t^z and ε_t^l. The error in forecasting $\log l_t$ given Ω_{t-1}, $u_{\Omega,t}^l$, is proportional to ε_t^l. Specifically,

$$u_{\Omega,t}^a = \alpha_1 \varepsilon_t^z + \alpha_2 \varepsilon_t^l, \; u_{\Omega,t}^l = \gamma \varepsilon_t^l,$$

where $\alpha_1 > 0$, α_2 and γ are functions of the model parameters. It follows that $\alpha_1 \varepsilon_t^z$ is the error from regressing $u_{\Omega,t}^a$ on $u_{\Omega,t}^l$:

$$u_{\Omega,t}^a = \beta u_{\Omega,t}^l + \alpha_1 \varepsilon_t^z, \; \beta = \frac{cov(u_{\Omega,t}^a, u_{\Omega,t}^l)}{V\left(u_{\Omega,t}^l\right)},$$

where $cov(x, y)$ denotes the covariance between the random variables, x and y, and $V(x)$ denotes the variance of x. Recall, we normalize the standard deviation of ε_t^z to be unity. Consequently, the value of α_1 can be recovered as the positive square root of the variance of the forecast error in this regression:

$$\alpha_1 = \sqrt{V\left(u_{\Omega,t}^a\right) - \beta^2 V\left(u_{\Omega,t}^l\right)}.$$

In practice, we implement the previous procedure using the one-step-ahead forecast errors generated from a VAR. It is convenient to work with a version of (2.3) in which the variables in Y_t are ordered as follows:

$$Y_t = \begin{pmatrix} \log l_t \\ \Delta \log a_t \\ x_t \end{pmatrix},$$

where x_t is an additional vector of variables that may be included in the VAR. In addition, we write the vector of VAR one-step-ahead forecast errors, u_t as:

$$u_t = \begin{pmatrix} u_t^l \\ u_t^a \\ u_t^x \end{pmatrix}.$$

We identify the technology shock with the second element in e_t in (2.4). To compute the dynamic response of the variables in Y_t to the technology shock, we require $B_1, ..., B_q$ in (2.3) and the second column of the matrix, C, in (2.4). We obtain the elements of the second column of C in two steps. First, we identify the technology shock using:

$$\varepsilon_t^z = \frac{1}{\hat{\alpha}_1}\left(u_t^a - \hat{\beta} u_t^l\right),$$

where

$$\hat{\beta} = \frac{cov(u_t^a, u_t^l)}{V\left(u_t^l\right)}, \; \hat{\alpha}_1 = \sqrt{V\left(u_t^a\right) - \hat{\beta}^2 V\left(u_t^l\right)},$$

where the indicated variances and covariances are obtained from V in (2.3). Second, to obtain C_2, the second column of C, we regress u_t on ε_t^z :

$$C_2 = \begin{pmatrix} \frac{cov(u^l, e_2)}{var(e_2)} \\ \frac{cov(u^a, e_2)}{var(e_2)} \\ \frac{cov(u^x, e_2)}{var(e_2)} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{\alpha}_1 \\ \frac{1}{\hat{\alpha}_1}\left(cov(u_t^x, u_t^a) - \hat{\beta} cov\left(u_t^x, u_t^l\right)\right) \end{pmatrix}.$$

This procedure for computing C_2 can be implemented by computing $CC' = V$, where C is the lower triangular Choleski decomposition of V, and taking the second column of that matrix. This is a convenient strategy because the Choleski decomposition can be computed using widely-available software.

2.2. Parameterizing the Model

We consider different versions of the RBC model that are distinguished by the nature of the exogenous shocks. For comparability we assume, as in CKM, that:

$$\begin{aligned} \beta &= 0.98^{1/4},\ \theta = 0.33,\ \delta = 1 - (1 - .06)^{1/4},\ \psi = 2.5,\ \gamma = 1.01^{1/4} - 1 \\ \bar{\tau}_x &= 0.3,\ \bar{\tau}_l = 0.242,\ \mu_z = 1.016^{1/4} - 1,\ \sigma = 1. \end{aligned} \tag{2.9}$$

We consider various parameterizations for the shocks. These parameterizations were chosen to illustrate the key factors determining the reliability of inference based on short-run and long-run identification restrictions.. It is convenient to report, for each parameterization, the variance of HP-filtered output due to technology shocks.

KP Specification

In the *Benchmark KP Model*, the technology shock process is the same as the one estimated by Prescott (1986):[8]

$$\log z_t = \mu_Z + 0.011738 \times \varepsilon_t^z.$$

Erceg, Guerrieri and Gust (2005) update Prescott's analysis and estimate σ_z to be 0.0148. To be conservative, we use Prescott's estimate because it attributes relatively less importance to technology shocks in aggregate fluctuations. Although he concentrates on technology shocks in his analysis, Prescott (1986) argues that other shocks also affect aggregate fluctuations. To maintain comparability with CKM, we specify τ_{lt} to be the other shock in this specification.

[8] Prescott (1986) estimates that the standard deviation of the innovation to technology growth is 0.763 percent. However, he adopts a different normalization than we do, placing technology in front of the production function, rather than next to hours worked, as we do. Our standard deviation is 0.01174=0.00763/(1-.35).

For the benchmark KP specification, we estimate a law of motion for $\tau_{l,t}$ as follows. Combining the household and firm first order conditions for labor, and rearranging, we obtain:

$$\tau_{l,t} = 1 - \frac{c_t}{y_t}\frac{l_t}{1-l_t}\frac{\psi}{1-\theta}.$$

Given our parameter values, we compute a time series for $\tau_{l,t}$, and estimate the following first order autoregressive representation:[9]

$$\tau_{l,t} = (1-0.993)\times 0.242 + 0.993\times\tau_{l,t-1} + 0.0066\times\varepsilon_t^l.$$

Figure 1 depicts the time series on $\tau_{l,t}$, $l_t/(1-l_t)$ and c_t/y_t.[10]

Let p denote the percent variance in HP-filtered, log output due to technology shocks in a model. Table 1 reports that $p = 71$ in the Kydland-Prescott specification.[11] This value of p is consistent with a key claim advanced by Kydland and Prescott, namely that technology shocks account for roughly 70% of the cyclical volatility of output. The finding that p is 71% is the reason we refer to this version of our model as the Kydland-Prescott specification. For reference, Table 1 reports other standard business cycle statistics for the KP specification.

To assess robustness, we also estimated the parameters, ρ_l, σ_l, and σ_z, by maximizing the Gaussian likelihood function of the vector, $X_t = (\Delta\log y_t, \log l_t)'$, subject to the parameter values in (2.9). We do so using the standard, Kalman filter strategy discussed in Hamilton (1994, section 13.4).[12] Our results are given by:

$$\begin{aligned}\log z_t &= \mu_Z + 0.00953\times\varepsilon_t^z\\ \tau_{l,t} &= (1-0.986)\times\bar{\tau}_l + 0.986\times\tau_{l,t-1} + 0.0056\times\varepsilon_t^l\end{aligned}$$

For this specification, we find $p = 67$. We infer that our estimate of the benchmark KP specification is robust.

We also consider a *Three Shock KP Model*, obtained by adding the investment tax shock, τ_{xt}. We estimated this version of the model by maximizing the Gaussian likelihood function of

[9] Consumption, c_t, is the sum of nondurables, services and government consumption. We measure the ratio, c_t/y_t, as the ratio of dollar quantities. Total hours worked, l_t, is nonfarm business hours worked divided by a measure of the population, aged 16 and older. The data cover the period 1959QIII to 2001QIV. For the purpose of these calculations, we scaled per capita hours worked so that the sample mean coincides with steady state hours worked in the model. We obtained our estimates of ρ_l and σ_l by regressing $\tau_{l,t} - \bar{\tau}_l$ on $\tau_{l,t-1} - \bar{\tau}_l$. In this way, we imposed that the constant term in the regression is consistent with the value of $\bar{\tau}_l$ in CKM. When we estimate the regression without imposing this restriction, we obtain essentially the same results.

[10] We found some evidence of serial correlation in the fitted disturbances, ε_t^l. We decided to stay with the AR(1) representation, in order to maintain comparability with CKM.

[11] The 71 percent figure is a population value, computed using the spectral integration approach described in Christiano (2002).

[12] We removed the sample mean from X_t prior to estimation and set the measurement error in the observer equation to zero.

the vector, $X_t = (\Delta \log y_t, \log l_t, \Delta \log i_t)'$, subject to the parameter values in (2.9). The results are:

$$\begin{aligned}
\log z_t &= \mu_Z + 0.00968 \times \varepsilon_t^z \\
\tau_{l,t} &= (1-\rho_l) \times \bar{\tau}_l + \rho_l \times \tau_{l,t-1} + 0.00631 \times \varepsilon_t^l, \ \ \rho_l = 0.9994 \\
\tau_{x,t} &= (1-\rho_x) \times \bar{\tau}_x + \rho_x \times \tau_{x,t-1} + 0.00963 \times \varepsilon_t^x, \ \ \rho_x = 0.9923
\end{aligned}$$

For this specification, $p = 59$. Note that the estimated values of ρ_x and ρ_l are close to unity. This is consistent with the experience of other authors who, when they estimate general equilibrium models, find that shocks exhibit high serial correlation.[13]

CKM Benchmark Specification

The *Benchmark CKM Model* has two shocks, z_t and τ_{lt}, which have the following time series representations:

$$\begin{aligned}
\log z_t &= \mu_Z + \log z_t = \mu_Z + 0.00568 \times \varepsilon_t^z \\
\tau_{lt} &= (1 - 0.940)\, \bar{\tau}_l + 0.940 \times \tau_{l,t-1} + 0.0080 \times \varepsilon_t^l.
\end{aligned}$$

In sharp contrast to the Kydland and Prescott specification, the benchmark CKM model implies that p is only 20 (see Table 1). Other business cycle implications of the CKM benchmark specification are reported in Table 1. Notice that productivity exhibits a counterfactualy, near-perfect negative correlation between productivity and hours worked. This negative correlation reflects that non-technology shocks are the primary driver of business cycle fluctuations in the CKM benchmark specification.

We also consider a *Three Shock CKM Model*, obtained by adding the investment shocks, $\tau_{x,t}$. We assume that the law of motion for $\tau_{x,t}$ is the one used by CKM:

$$\tau_{xt} = (1 - 0.98)\, \bar{\tau}_x + 0.98 \times \tau_{x,t-1} + 0.00847 \times \varepsilon_t^x. \tag{2.10}$$

The law of motion of z_t and τ_{lt} in this model is the same as it is in the CKM benchmark model.

Other Specifications

For diagnostic purposes we consider a series of perturbations of the KP and CKM models where we vary the values of σ and σ_l. Varying these parameters allows us to change the value of p, the fraction of business cycle variance in output due to technology shocks. When we perturb σ, we always adjust ψ so that steady state labor is what it is in the associated benchmark model.

[13] See, for example, Christiano (1988), Christiano, Motto and Rostagno (2004), Smets and Wouters (2003).

3. Results Based on Conventional Estimation Strategies

In this section we analyze the properties of conventional VAR-based strategies for identifying the effects of a technology shock. Our basic strategy is to simulate artificial time series using variants of the economic model discussed above as the data generating process. By construction we know the actual response of hours worked to a technology shock. We then consider what an econometrician using VARs would find, on average over repeated small samples.

In this section we focus on two key questions. First, is there substantial bias associated with the estimated dynamic response functions of hours to a technology shock? Second, is there substantial bias in a standard estimator of sampling uncertainty? The first subsection presents our answers to these questions when we use the recursive version of the model. The second subsection presents our results when we use the long-run properties of the standard version of the model to identify technology shocks.

3.1. Recursive Identification

Analysis of the KP Specification

We begin by discussing the results we obtained using variants of the KP specification as the data generating mechanism. Throughout we proceed as follows. Using the economic model as the data generating mechanism,we simulate 1000 data sets, each of length 180 observations. The shocks ε_t^z, ε_t^l and possibly ε_t^x are drawn from *i.i.d.* standard normal distributions.

On each data set we estimate a four lag VAR. In data generated from the benchmark KP and CKM models, the variables in the VAR are $\Delta \log a_t$ and $\log l_t$. In data generated from the three-shock KP and CKM models, the variables in the VAR are $\Delta \log a_t$, $\log l_t$ and $\Delta \log i_t$, where

$$i_t = (1+\gamma)\, k_{t+1} - (1-\delta)\, k_t.$$

Given the estimated VAR, we calculate the dynamic response of hours to a technology shock based on the short-run identifying restriction and method discussed in section 2.1.2 above. The solid lines in Figure 2 are the average dynamic response function obtained over the 1000 synthetic data sets in the different specifications. The starred lines are the true dynamic response function of hours worked implied by the economic model that is being used as the data generating process. The grey areas in the figure are measures of the sampling uncertainty associated with the estimated dynamic response functions. We obtain these measures by first calculating the standard deviation of the points in the estimated impulse response functions across the 1000 synthetic data sets. The grey areas correspond to a two standard deviation band about the relevant solid black line. The dashed lines corresponds to the top 2.5% and bottom 2.5% of the estimated coefficients in the dynamic response functions across the 1000 synthetic data sets. To the extent that the dashed lines coincide with boundaries of the grey area, there is support for the notion that the coefficients of estimated impulse response functions are normally distributed.

An important question is whether an econometrician would correctly estimate the true uncertainty associated with the estimated dynamic response functions. To address this question we proceed as follows. For each synthetic data set and corresponding estimated impulse response function, we calculated the bootstrap standard deviation of each point in the impulse response function. Specifically, for a given synthetic data set, we estimate a VAR and use it as the data generating process to construct 200 synthetic data sets, each of length 180 observations, by randomly drawing from the fitted VAR disturbances. For each of these 200 synthetic data sets, we estimate a new VAR and impulse response function. We then calculate the standard deviation of the coefficients in the impulse response functions across the 200 data sets. Finally, we take the average of these standard deviation across the 1000 synthetic data sets that were generated using the economic model as the data generating process. The lines with 0's in Figure 2 correspond to a two standard deviation band about the solid black line and are a measure of the average standard deviations that a econometrician would construct.

The top left graph in Figure 2 exhibits the properties of the VAR estimator of the response of hours to a technology shock when the data are generated by the benchmark KP specification. The 2,1 graph in Figure 2 corresponds to the case when the data generating mechanism is the KP specification with $\sigma = 0$. This case is of interest, because utility is roughly linear in leisure, corresponding to Hansen (1985)'s indivisible labor model. The 3,1 graph in Figure 2 shows what happens when σ is increased above its benchmark specification, to $\sigma = 6$, in which case the Frisch labor supply elasticity is 0.63. This case is of interest, because this is the same Frisch elasticity used in the model studied by Erceg, Guerrieri and Gust (1004). The 4,1 graph in Figure 2 shows what happens in the three shock KP model. In each case, the impact effect on hours worked and associated sampling variance is also reported, for convenience, in Table 2.

The first column of Figure 2 exhibits two striking features. First, regardless of which variant of the KP specification we work with, there is no evidence whatsoever of bias in the estimated impulse response functions. In all cases, the solid lines virtually coincide with the starred lines. Second, Figure 2 indicates that an econometrician would not be misled in inference using standard procedures for constructing confidence intervals. This conclusion reflects the fact that the average value of the econometrician's confidence interval (the line with the 0's) coincides closely to the actual range of variation in the impulse response function (the grey area). Finally, it is also interesting to note that the estimated coefficients of the impulse response functions appear normally distributed: in all cases the boundaries of the grey area coincide closely with the dashed lines.

Analysis of the CKM Specification

The right hand column of Figure 2 reports our results when the data generating mechanism is given by variants of the CKM specification. The top right hand graph in Figure 2 corresponds to the CKM specification. The 2,2 and 2,3 graphs in Figure 2 correspond versions of the benchmark CKM model with $\sigma = 0$ and $\sigma = 6$, respectively. Finally, the 4,1 graph in Figure 2 corresponds to the three variable CKM model.

Notice that the second column of Figure 2 contains the same striking features as the first column. First, there is no evidence whatsoever of bias in the estimated impulse response functions. Second, the average value of the econometrician's confidence interval coincides closely to the actual range of variation in the impulse response function (the grey area).

In sum, our analysis of the recursive identification scheme reveals that structural VAR's perform remarkably well. This is extremely comforting for the vast literature that has exploited recursive identification schemes to identify the dynamic effects of shocks to the economy. Of course, one can criticize the particular short-run identifying assumptions used in any given analysis. But our results strongly support the view that if the relevant recursive assumptions are satisfied in the data generating mechanism, standard structural VAR procedures will reliably uncover and identify the dynamic effects of shocks to the economy.

Finally, note that we did *not* include capital as a variable in the VAR. Despite this omission, the structural VAR procedure performs very well. This demonstrates that, claims in CKM to the contrary, omitting the economically relevant state variable capital does not in and of itself pose a problem for inference using structural VAR's.

3.2. Long-run Identification

Analysis of the KP Specification

We begin by discussing results associated with variants of the KP specification. As above we use versions of the KP specification as the data generating mechanism to simulate 1000 data sets, each of length 180 observations. The shocks ε_t^z, ε_t^l and possibly ε_t^x are drawn from *i.i.d.* standard normal distributions. On each data set we estimate a four lag VAR. Two or three variables are included in the VAR depending on the specification being analyzed. Given the estimated VAR, we calculate the dynamic response of hours to a technology shock based on the long-run identifying restriction and method discussed in section 2.1.1 above. The different lines, as well as the grey areas in the Figure 3 are the analogs of the corresponding objects in Figure 2.

The top left graph in Figure 3 exhibits the properties of the VAR estimator of the response of hours to a technology shock, when the data are generated by the KP specification. Notice that there is virtually no bias in the estimate of the response of hours worked to a technology shock. While there is considerable sampling uncertainty in the estimator, the econometrician would not be substantially misled with respect to inference. This is because, while there is some tendency to understate sampling uncertainty, the average value of the econometrician's confidence interval (the line with the 0's) coincides reasonably closely to the actual range of variation in the impulse response function (the grey area).

Consider next the 2,1 graph in Figure 3, which corresponds to the version of the benchmark KP model with $\sigma = 0$. Note that there is a slight increase in bias, although the bias is very small compared with sampling uncertainty. To understand the reason for the appearance of some (small) bias in this case, it is interesting to note that $p = 60$, which is somewhat smaller than the corresponding value of 71 in the benchmark KP specification (recall, p is the percent of the business cycle variance in output due to technology shocks). Reducing σ increases the response of hours worked to both technology and labor tax shocks. However, the impact on the response of hours worked to a labor tax shock is greater than the impact on the response to a technology shock.

The 3,1 graph in Figure 3 shows what happens when σ is increased above its benchmark specification, to $\sigma = 6$.In this case the bias in the VAR-based estimator of the impulse response function almost disappears, and the sampling uncertainty shrinks drastically. To understand the reason for this, simply apply the discussion underlying the 2,1 graph in reverse. In the parameterization underlying the 3,1 graph, $p = 92$, so that technology shocks account for the vast majority of cyclical fluctuations in output.

The 4,1 graph in Figure 3 shows what happens in the three shock KP model. In this case there is an increase in bias relative to the benchmark case, although the bias is still small relative to sampling uncertainty. Consistent with the $\sigma = 0$ case, the rise in the bias is associated with a fall in p in the three shock case to $p = 59$.

Analysis of the CKM Specification

Consider the left column in Figure 4. The top left graph reports results for the benchmark specification. Consistent with results reported in CKM, there is substantial bias in the estimated dynamic response function. In the model, the contemporaneous response of hours to a one-standard-deviation technology shock is 0.14 percent, while the mean estimated response is 0.65 percent. This large bias is due to the fact that, in CKM's parameterization of the model, technology shocks play a very small role in output fluctuations, with $p = 20$ (see Table 1). The second and third rows of Figure 4 report results with $\sigma_l = 0.0080/2$ and $\sigma_l = 0.0080/3$, respectively. In these two cases, the percent business cycle variance in output, p, is $p = 50$ and $p = 70$, respectively. Note how the accuracy of the impulse response functions improves as p increases.

Figure 5 presents the analog results to Figure 2 for the CKM model. The first row of Figure 5 reproduces the first row of Figure 4, for convenience. The second row corresponds to indivisible labor case, $\sigma = 0$. The third row corresponds to the low Frisch elasticity case, $\sigma = 6$. The fourth row corresponds to the case with three shocks. Notice that there is a considerable increase in the bias in the indivisible labor and three shock cases. In these two cases, $p = 11$ and $p = 18$, respectively. So, large bias is associated with low values of p. The bias is reduced when $\sigma = 6$, in which case $p = 60$.

In summary, our results for the CKM model show that there exists parameterizations for which large biases emerge when the conventional long-run estimation strategy is used. Below, we argue that these parameterizations are empirically uninteresting. In the next section we develop a variant of the long-run estimation strategy that virtually eliminates the bias, even in these empirically uninteresting cases.

4. An Improved Long-Run Estimator

In the previous section we have shown that with the short run identification strategy, the conventional estimator of impulse response functions is remarkably accurate. In contrast, we found that for some parameterizations of the data generating process the conventional estimator based on long-run identification restrictions leads to substantial bias. In this section we explore why recursive identication leads to such accurate results, whereas long-run identification may lead to biases. We build on this analysis to develop an improved long-run estimation strategy.

We begin by considering a simple analytic expression due to Sims (1972), which approximates what an econometrician who fits a misspecified VAR will find. The expression is an approximation, because it assumes a large sample of data.[14] Let $\hat{B}_1, ..., \hat{B}_q$ and $\hat{V}$ denote the parameters of the $q-th$ order VAR fit by the econometrician to the data. Then,

$$\hat{V} = V + \min_{\hat{B}_1,...,\hat{B}_q} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[B\left(e^{-i\omega}\right) - \hat{B}\left(e^{-i\omega}\right) \right] S_Y(\omega) \left[B\left(e^{i\omega}\right) - \hat{B}\left(e^{i\omega}\right) \right]' d\omega, \tag{4.1}$$

where $B\left(e^{-i\omega}\right)$ is $B(L)$ with L replaced by $e^{-i\omega}$.[15] Here, B and V are the parameters of the actual VAR representation of the data, and $S_Y(\omega)$ is the associated spectral density, at frequency ω. [16] According to (4.1), estimation of a VAR approximately involves choosing VAR lag matrices

[14] For additional discussion of the Sims formula, see Sargent (1979). One interpretation of (4.1) is that it provides the probability limit of our estimators: what they would converge to as the sample size increases to infinity. We do not adopt this interpretation, because in practice an econometrician uses a consistent lag length selection method. As a result, the probability limit of our estimators corresponds to the true impulse response functions in all cases considered in this paper. We verified this by solving (4.1) with $q = 300$.

[15] The minimization is actually over the trace of the indicated integral.

[16] The derivation of this formula is straightforward. Suppose that the true VAR representation of the covariance stationary process, Y_t, is:

$$Y_t = B(L)Y_{t-1} + u_t,$$

where $B(L)$ is a possibly infinite-ordered matrix polynomial in non-negative powers of L and $Eu_tu_t' = V$. Suppose the econometrician contemplates a particular parameterization of $B(L)$, $\hat{B}(L)$. Let the fitted disturbances to minimize a quadratic form in the difference between the estimated and true lag matrices. The quadratic form assigns greatest weight to the frequencies where the spectral density is the greatest. If the econometrician's VAR is correctly specified, then $\hat{B} = B$ and $\hat{V} = V$ and the estimator is consistent. If there is specification error, then $\hat{B} \neq B$ and $V > \hat{V}$.[17] In our context, there is specification error because the true VAR implied by our models has $q = \infty$, but the econometrician uses a finite value of q. Specifically, we work with $q = 4$.

We have found that $\hat{V}$ is an accurate estimate of V, even for low values of q. For example, in the case of the benchmark CKM model, the value of $\hat{V}$ that solves (4.1) is the same, to three significant digits, for $q = 4$, 8, 16, and greater.[18] This result is perhaps not surprising in view of (4.1), which shows that a principle objective of least squares is to get $\hat{V}$ as close as possible to V. We may also infer from the fact that impulse response functions based on recursive identification are so accurate, that the levels of $\hat{B}_1, ..., \hat{B}_4$ are reasonably well estimated.

It may seem that if $\hat{V}$ and the levels of $\hat{B}_1, ..., \hat{B}_4$ are estimated reasonably well, then estimation based on long-run restrictions should have worked well too. But, such an impression would be incorrect. Estimation based on long run restrictions requires, in addition to $\hat{V}$ and the levels of $\hat{B}_1, ..., \hat{B}_4$, an accurate estimate of the sum of the VAR coefficients, $B(1)$. As emphasized by Sims (1972), it is possible for the levels of regression coefficients to be estimated reasonably accurately, and yet for the estimate of the sum to be way off.[19] Expression (4.1) indicates that $\hat{B}(1)$ will be a good approximation for $B(1)$ only if $S_Y(\omega)$ happens to be relatively large in a neighborhood of $\omega = 0$. This is simply not something one can count on.[20]

The previous reasoning suggests that estimation based on long-run restrictions may be improved if the zero-frequency spectral density in (2.6) is replaced by an estimator that is specifically designed for the task. With this in mind, we replace S_0 with a standard Newey-West

associated with this parameterization be denoted $\hat{u}_t$. Simple substitution implies:

$$\hat{u}_t = \left[B(L) - \hat{B}(L)\right] Y_{t-1} + u_t.$$

The two random variables on the right of the equality are orthogonal, so that the variance of $\hat{u}_t$ is just the variance of the sum of the two:

$$var(\hat{u}_t) = var\left(\left[B(L) - \hat{B}(L)\right] Y_{t-1}\right) + V.$$

Expression (4.1) in the text follows immediately.

[17]By $V > \hat{V}$, we mean that $V - \hat{V}$ is a positive definite matrix.

[18]This explains why lag length selection methods, such as the Akaike criterion, almost never suggest values of q greater than 4 in artificial data sets of length 180, regardless of which of our data generating methods we used. These lag length selection methods focus on $\hat{V}$.

[19]See Sims (1972, p. 169) for a simple illustration of this point.

[20]Consistent with these observations, we found some modest improvement in estimators when we applied a band pass filter to remove the very highest frequencies of the data, prior to analysis.

estimator:

$$S_0 = \sum_{k=-(T-1)}^{T-1} g(k)\hat{C}(k), \; g(k) = \begin{cases} 1 - \frac{|k|}{r} & |k| \le r \\ 0 & |k| > r \end{cases}, \tag{4.2}$$

and (after removing the sample mean from Y_t)

$$\hat{C}(k) = \frac{1}{T} \sum_{t=k+1}^{T} Y_t Y_{t-k}'.$$

We use essentially all possible covariances in the data by choosing a large value of r, $r = 150$.[21]

In the next section, we show that our modified procedure for implementing long run restrictions does a good job at correcting distortions that occur in the estimation of impulse response functions based on long-run restrictions. In some respects, our modified estimator is equivalent to running a VAR with longer lags. The crucial difference, however, is that our method does not require choosing a VAR lag length. This is an important advantage, because standard lag length selection procedures are notoriously unreliable.

There are various conjectures in the literature, concerning the impact of persistence in non-technology shocks on long-run identification of technology shocks. For example, there is an intuition that the more persistence there is in a non-technology shock, the greater will be the distortions in long-run identification. The idea is that if a non-technology shock is highly persistent, then it will have a long-lived effect on labor productivity, and this will be confounded with the effects of technology shocks. The formula in (4.1) shows that there is another consideration that works in the other direction. Greater persistence in a non-technology shock shifts power in the spectrum towards the lower frequencies. But, the more weight there is in the low frequency component of the data, *i.e.*, the larger is $S_Y(\omega)$ for ω close to zero, the more likely it is that $\hat{B}(1)$ is a good approximation to $B(1)$. Other things the same, this implies that long-run identification is improved. At the same time, a process for which $S_Y(0)$ is large may also have the property that $\hat{V}$ poorly approximates V. As a result, the net effect of persistence in non-technology shocks is in fact ambiguous. To see this, consider Figure 6. That graphs the ratio of the econometrician's estimator of the contemporaneous impact on hours worked of a technology shock, to its true value, against ρ_l, for $\rho_l \in [-0.5, 0.9999]$. We use the benchmark CKM model as the data generating mechanism. Also, for each value of ρ_l, σ_l is adjusted so that the unconditional variance of output is held fixed at the value implicit in the benchmark CKM parameterization. The dot-dashed line in the figure corresponds to the solution of (4.1), with $q = 4$, using the standard VAR-based estimator.[22] The star in the figure indicates the value of ρ_l

[21] Setting the bandwidth, r, equal to sample size does not rovide a consistent estimator of the spectral density at frequency zero. We assume that as sample size is increased beyond $T = 180$, the bandwidth is increased sufficiently slowly that consistency obtains.

[22] Since (4.1) is a quadratic function, we solved the optimization problem by solving the linear first order

in the benchmark CKM model. Note how in the neighborhood of this value of ρ_l, the distortion falls sharply as ρ_l increases. Indeed, for $\rho_l = 0.9999$, there is essentially no distortion. For values of ρ_l in the region, $(-0.5, 0.5)$, distortion increases with increases in ρ_l.

The results in Figure 6 also allow us to assess the accuracy of the approximation in (4.1). The solid line shows the small sample ($T = 180$) mean of the standard estimator corresponding to each value of ρ_l. For each ρ_l this mean was computed as the mean across 1,000 Monte Carlo replications. That the curve is not completely smooth reflects the presence of a small amount of Monte Carlo sampling error in the calculations. Although the solid line is uniformly below the dot-dashed line, note that the basic patterns in the two curves is the same. We conclude that (4.1) is a reliable guide to the operating performance of the standard VAR-based estimator of impulse response functions.

The results in Figure 6 also allow us to assess the value of our proposed modification to the standard estimator. The distortions in the standard estimator are quite large, particularly for small degrees of persistence in the non-technology shock. When the standard estimator works well, i.e., for large values of ρ_l, then the modified and standard estimators produce similar results. However, when the standard estimator works poorly, as for ρ_l near 0.5, then our modified estimator shrinks bias by more than a factor of two. In the following section we discuss the performance of our modified estimator in the whole set of data generating mechanisms analyzed in section 3.

5. Results For The Improved Long-Run Estimator

We now evaluate the performance of our modified estimator, using the examples in section 3. Consider Figure 3, which presents results when the data generating mechanism corresponds to versions of the KP model. In that case, the standard estimator (see the left column) has relatively little bias, and our modified estimator also has little bias (right column). We noted before that the econometrician's estimator of standard errors understates somewhat the degree of sampling uncertainty. Interestingly, the modified estimator represents an improvement in this dimension. Note how the lines with circles roughly coincide with the boundary of the grey area. It is also interesting to note that the degree of sampling uncertainty with the modified estimator is not greater than the sampling uncertainty of the standard estimator. In fact, in some cases there is a slight reduction in sampling uncertainty. Now consider Figure 4. These are variants of the CKM model, in which evidence of bias appears in the standard estimator. Note in the right column how bias is greatly reduced with the modified estimator. Recall that Figure 5

conditions. These are the Yule-Walker equations, which rely on population second moments of the data (solvingt these equations is the strategy for computing population projections suggested by Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005)). We obtained the population second moments by complex integration of the reduced form of the model used to generate the data, as suggested in Christiano (2002).

documents what happens when σ_l is reduced. When it is reduced to the point that technology shocks account for 70 percent of fluctuations, then the standard estimator performs quite well. Our modified estimator also does well in this case. The results in the top two graphs on the left exhibit examples where the standard estimator results in distortions. These distortions are virtually eliminated with our modified estimator.

In sum, when the standard estimator works well, then the modified estimator also works well. When there was evidence of distortion in the standard estimator, then the modified estimator virtually eliminated the distortion.

6. Relation to CKM

In the preceding sections, we showed that conventional VAR-based methods using long run restrictions work well in data generated by the KP model. However, substantial biases emerge when we used the CKM model. We showed that the key reason the CKM model has such different implications for VARs is that it attributes only a small fraction of output fluctuations to technology shocks. This feature of the CKM model reflects that when they estimate it, CKM impose a highly unusual assumption. We show that when this assumption is dropped, the likelihood function jumps by orders of magnitude and the resulting estimated model is similar to the benchmark KP model in that it attributes a substantial fraction of business fluctuations to technology shocks. We conclude that the CKM model is empirically uninteresting. The implications of that model serve as a potentially important warning to users of VARs that they should consider using the modified VAR approach that we propose. However, the CKM results do not constitute a basis for their conclusion that VARs are useless in practice.

At the heart of the CKM estimation strategy is the remarkable assumption that technology growth can be measured with a high degree of accuracy by the growth rate of government purchases. Specifically, CKM assume that technology growth equals government consumption plus a measurement error which has a very small, exogenously fixed variance. The consequences of this modeling assumption are not surprising. It is well known that government purchases are at best weakly correlated with output. Since CKM's estimation criterion also includes output growth, the model must therefore rely on other shocks to account for output fluctuations.

CKM adopt a Kalman filter framework for estimating their model. Let

$$Y_t = \left(\Delta \log a_t, \log l_t, \Delta \log i_t, \Delta \log g_t\right),$$

where g_t denotes government spending. CKM suppose:

$$Y_t = X_t + u_t,\ Eu_t u_t' = R, \tag{6.1}$$

where R is diagonal, u_t is a 4×1 vector of iid measurement error and X_t is a 4×1 vector containing the model's implications for the variables in Y_t. The model being estimated only has two shocks (τ_{lt} and z_t). The parameters estimated are the ones listed in Table 3. In CKM, government spending is modeled as:

$$g_t = g_{z,t} \times Z_t,$$

where $g_{z,t}$ is in principal an exogenous stochastic process. However, when estimating the parameters of the technology and preferences processes, τ_{lt} and z_t, the variance of the government spending shock is set to zero, so that g_{zt} is a constant. As a result, CKM assume

$$\Delta \log g_t = \log z_t.$$

During estimation, the elements on the diagonal of R are fixed at 0.01^2. The optimized value of the likelihood function, as well at the parameter estimates appear in the first column of Table 3. The second column shows what happens when we allow the diagonal elements of R to be free. Twice the difference of the optimized likelihoods is 427. Under the null hypothesis that CKM's setting for R is correct, this is the realization of a chi-square distribution with 4 degrees of freedom. Few hypotheses have been rejected as overwhelmingly as this one! Interestingly, the model with the elements of R freely estimated implies that the fraction of output variance due to technology is 61, three times greater than CKM's estimate. The third column in Table 3 shows what happens when we simply drop government spending from the analysis, but keep the measurement error variance fixed at its value in the first column. Note that the resulting estimated model is like our KP model in that it implies that 70 percent of the cyclical variation in output is due to technology shocks. Ultimately, CKM also estimate the parameters for two other stochastic processes, τ_{xt} and g_{zt}. However, in this estimation, they always fix the parameters of the stochastic processes underlying τ_{lt} and z_t to the estimates reported in the first column of Table 2.[23]

7. Concluding Remarks

In this paper we studied the performance of structural VARs for uncovering the response of hours worked to a technology shock. For pedagogical reasons, we only considered very simple data generating processes, based on variants of a prototype RBC model. We find that with short-run restrictions, structural VARs perform remarkably well. With long-run restrictions we find that structural VARs work well when technology shocks play an important role in business cycle fluctuations, as in the tradition of the RBC literature. This is a property of an RBC model that we fit by maximum likelihood to the data. We confirmed the results in CKM, who display models which imply that VARs do less well. These models have this property because they imply that technology shocks play only a minor role in business cycle fluctuations. We developed a

[23] To ensure comparability of our results with CKM, the calculations underlying Table 2 use their computer code and data, available on Ellen McGrattan's web page.

modification to the usual VAR methods which works well in artificial data generated by the CKM models (the modification also works well in the cases where standard methods already work well). However, it is not clear the CKM examples should be taken seriously as evidence that VARs do not work in practice. We showed that the CKM model is rejected overwhelmingly by the data.

References

[1] Aiyagari, S. Rao, 1994, 'On the Contribution of Technology Shocks to Business Cycles,' Federal Reserve Bank of Minneapolis Quarterly Review, vol. 18, no. 1, Winter, pp. 22-34.

[2] Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2004. 'Are Technology Improvements Contractionary?' National Bureau of Economic Research Working Paper 10592.

[3] Bernanke, Ben S. and Alan S. Blinder (1992), 'The Federal Funds Rate and the Channels of Monetary Transmission', *American Economic Review,* Vol. 82, No. 4, pages 901 - 921.

[4] Bernanke, Ben S. and Ilian Mihov (1995), 'Measuring Monetary Policy', NBER Working Paper No. 5145.

[5] Blanchard, Olivier Jean, and Danny Quah, 1989, 'The Dynamic Effects of Aggregate Demand and Supply Disturbances,' *American Economic Review*, v. 79, issue 4, pp. 655-673.

[6] Blanchard, Olivier Jean, and Mark Waston, 1986, 'Are Business Cycles All Alike?'. In Robert Gordon (ed), Continuity and Change in the American Business Cycle. The University of Chicago Press, pp.123 - 156.

[7] Chari, V.V., Patrick Kehoe and Ellen McGrattan, 2005, 'A Critique of Structural VARs Using Real Business Cycle Theory,' Federal Reserve Bank of Minneapolis Working Paper 631, February.

[8] Christiano, Lawrence J., 2002, 'Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients,' *Computational Economics*, October, Volume 20, Issue 1-2.

[9] Christiano, Lawrence J. and Martin Eichenbaum (1992), 'Identification and the Liquidity Effect of a Monetary Policy Shock', in *Political Economy, Growth and Business Cycles,* edited by Alex Cukierman, Zvi Hercowitz, and Leonardo Leiderman, Cambridge and London: MIT Press, pages 335 - 370.

[10] Christiano, Lawrence J., Martin Eichenbaum and Charles Evans, 2005, 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,' *Journal of Political Economy.*

[11] Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson, 2003, "What Happens After a Technology Shock?", National Bureau of Economic Research working paper w9819.

[12] Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson, 2004, 'The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology,' *Journal of the European Economic Association.*

[13] Christiano, Lawrence J., Roberto Motto and Massimo Rostagno, 2004, 'The Great Depression and the Friedman-Schwartz Hypothesis,' *Journal of Money, Credit and Banking.*

[14] Cushman, David O. and Tao Zha, 'Identifying Monetary Policy in a Small Open Economy under Flexible Exchange Rates', *Journal of Monetary Economics*, August 1997, v. 39, iss. 3, pp. 433-48.

[15] Del Negro, Marco, Frank Schorfheide, Frank Smets, and Raf Wouters, 2005, 'On the Fit and Forecasting Performance of New Keynesian Models,' manuscript, University of Pennsylvania.

[16] Eichenbaum, Martin and Charles Evans, 'Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates,' *Quarterly Journal of Economics*, 110(4), November 1995, pages 975-1009.

[17] Erceg, Christopher J., Luca Guerrieri, and Christopher Gust, 2004, 'Can Long-Run Restrictions Identify Technology Shocks?', International Finance Discussion Papers Number 792, Federal Reserve Boad of Governors.

[18] Faust, Jon and Eric Leeper, 1997, 'When Do Long-Run Identifying Restrictions Give Reliable Results?', *Journal of Business and Economic Statistics*, July, vol. 15, no. 3.

[19] Fernandez-Villaverde, Jesus, Rubio-Ramirez, Juan F. and Thomas J. Sargent, 2005, A,B,C's (And D's) For Understanding VARS, manuscript, NYU.

[20] Forni, Mario, Giannone, Domenico, Lippi, Marco and Lucrezia Reichlin, (2004), 'Opening the Black Box: Structural Factor Models versus Structural VARs', manuscript, Université Libre de Bruxelles.

[21] Francis, Neville, and Valerie A. Ramey, 2001, 'Is the Technology-Driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited,' manuscript, UCSD.

[22] Giannone, Domenico, Lucrezia Reichlin, and Luca Sala, 2005, The Kydland and Prescott Legacy, European Central Bank.

[23] Hansen, Gary, 1985, Indivisible Labor and the Business Cycle,' *Journal of Monetary Economics*, November, 16(3): 309-27

[24] Gali, Jordi, 1999, 'Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?' *American Economic Review*, 89(1), 249-271.

[25] Hamilton, James D. (1997), 'Measuring the Liquidity Effect', *American Economic Review,* Vol. 87, No. 1 pages 80 - 97.

[26] Kydland, Finn.E. and Edward Prescott, 1991, 'Hours and Employment Variation in Business Cycle Theory', *Economic Theory,* , pp 63 - 82.

[27] Pagan, Adrian R. and John Robertson (1998) "Structural Models of the Liquidity Effect", *Review of Economics and Statistics*, 80, 202-217.

[28] Prescott, Edward, 1991, 'Real Business Cycle Theory: What Have We Learned?', Revista de Analisis Economico, 6 (November), pp. 3-19.

[29] Quah, Danny and Thomas Sargent, 1993, 'A Dynamic Index Model for Large Cross Sections,' in Stock, James H., Watson, Mark W., eds. *Business cycles, indicators, and forecasting*, NBER Studies in Business Cycles, vol. 28. Chicago and London: University of Chicago Press, 285-306.

[30] Rotemberg, Julio J. and Michael Woodford, 1997, 'An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,' National Bureau of Economic Research *Macroeconomics Annual*.

[31] Sargent, Thomas, 1979, Macroeconomics.

[32] Sargent, Thomas J. and Christopher Sims, 1977, Business Cycle Modleing Without Much A Priori Economic Theory', in *New Methods in Business Cycle Research*, Federal Reserve Bank of Minneapolis.

[33] Sims, Christopher, 1989, 'Models and Their Uses,' *American Journal of Agricultural Economics* 71, May, pp. 489-494.

[34] Sims, Christopher, 1972, 'The Role of Approximate Prior Restrictions in Distributed Lag Estimation,' *Journal of the American Statistical Association*, vol. 67, no. 337, March, pp. 169-175.

[35] Sims, Christopher, 1980, 'Macroeconomics and Reality', *Econometrica*, January, pp. 1 - 48.

[36] Sims, Christopher A. and Tao Zha (1995), 'Does Monetary Policy Generate Recessions?', Manuscript. Yale University.

[37] Smets, Frank, and Raf Wouters, 2003, 'An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area,' *Journal of European Economic Association* 1(5), pp. 1123-1175.

[38] Uhlig, Harald, 2002, 'What Moves Real GNP?', manuscript, Humboldt University, Berlin.

Table 1: Selected Business Cycle Statistics					
	US data	KP Specification		CKM	
		Two-Shock	Three-Shock	Two-Shock	Three-Shock
σ_y	1.6	1.5	1.4	1.4	1.5
$\frac{\sigma_c}{\sigma_y}$	0.53 (0.051)	0.58	0.66	0.37	.47
$\frac{\sigma_i}{\sigma_y}$	3.66 (0.15)	2.69	3.89	3.80	4.63
$\frac{\sigma_l}{\sigma_y}$	1.08 (0.056)	0.84	0.98	1.34	1.36
$\mathrm{corr}(h, \frac{y}{h})$	-0.39 (0.10)	-0.10	-0.30	-0.74	-0.77
% variance due to technology		71	59	20	18

Note: σ_x - standard deviation of x; $\mathrm{corr}(x, y)$ correlation between x and y; Here, x and y have been logged first, and then HP-filtered; Standard Errors in Parentheses: Computed by GMM, Appropriate Zero-Frequency Spectral Density Estimated by Newey-West Proceedure Using Three Lag Autocorrelations and Bartlett Window

Table 2: Percent Contemporaneous Impact on Hours of One Standard Deviation Shock to Technology

	Contribution of	Impact of Tech Shock On Hours Worked		
	Tech Shocks to	True Value	Standard VAR	Newey-West
Model Specification	Business Cycle	(Plim)	Mean (Std Dev)	Mean (Std Dev)
Kydland-Prescott Parameterization				
Benchmark KP	71	0.29	0.34 (0.43)	0.13 (0.31)
$\sigma = 0$ ('indivisible labor')	60	0.43	0.55 (0.56)	0.22 (0.45)
$\sigma = 6$ (Frisch elasticity=0.63)	92	0.11	0.09 (0.19)	0.03 (0.13)
Three Shocks	59	0.24	0.33 (0.44)	0.14 (0.35)
Chari-Kehoe-McGrattan Paramerization				
CKM Benchmark	20	0.14	0.65 (0.39)	0.21 (0.43)
$\sigma = 0$ ('indivisible labor')	11	0.21	1.28 (0.51)	0.38 (0.59)
$\sigma = 6$ (Frisch elasticity=0.63)	60	0.05	0.13 (0.17)	0.03 (0.17)
$\sigma^l/2$	50	0.14	0.26 (0.22)	0.12 (0.23)
$\sigma^l/3$	70	0.14	0.18 (0.15)	0.09 (0.15)
Three Shocks	18	0.14	0.56 (0.48)	0.27 (0.48)

Note: Plim is the probability limit of the standard VAR-based estimator, when the econometrician uses a consistent lag-length estimator. We verified that such a Plim is correct performing the relevant population projections with a 300-lag VAR. For other details, see the text.

Table 3: Estimation Results for CKM Model				
	CKM Benchmark	Free Measurement Error Variance	Fixed Measurement Error, No Gov't Spending	Free Measurement Error, No Gov't Spending
$\mathcal{L}$	-2590.3	-2803.8	-2034.1	-2188.4
$\bar{\tau}_l$	0.2415	0.2536	0.2511	0.2550
ρ_l	0.9403	0.9818	0.9745	0.9865
σ_l	0.0080	0.0057	0.0062	0.0058
μ_z	0.0032	0.0042	0.0042	0.0042
σ_z	0.0057	0.0093	0.0122	0.0097
$v_{\Delta y}$	0.01	0.00	0.01	0.0
$v_{\log l}$	0.01	0.0164	0.01	0.0169
$v_{\Delta i}$	0.01	0.0037	0.01	0.0032
$v_{\Delta g}$	0.01	0.0188	NA	NA
p	20	61	70	64

Notes: p - fraction of business cycle variance in output due to technology shocks; v_x measurement error in x
L- maximized value of log likelihood function.

Figure 1 - The Labor Tax Wedge and Its Components

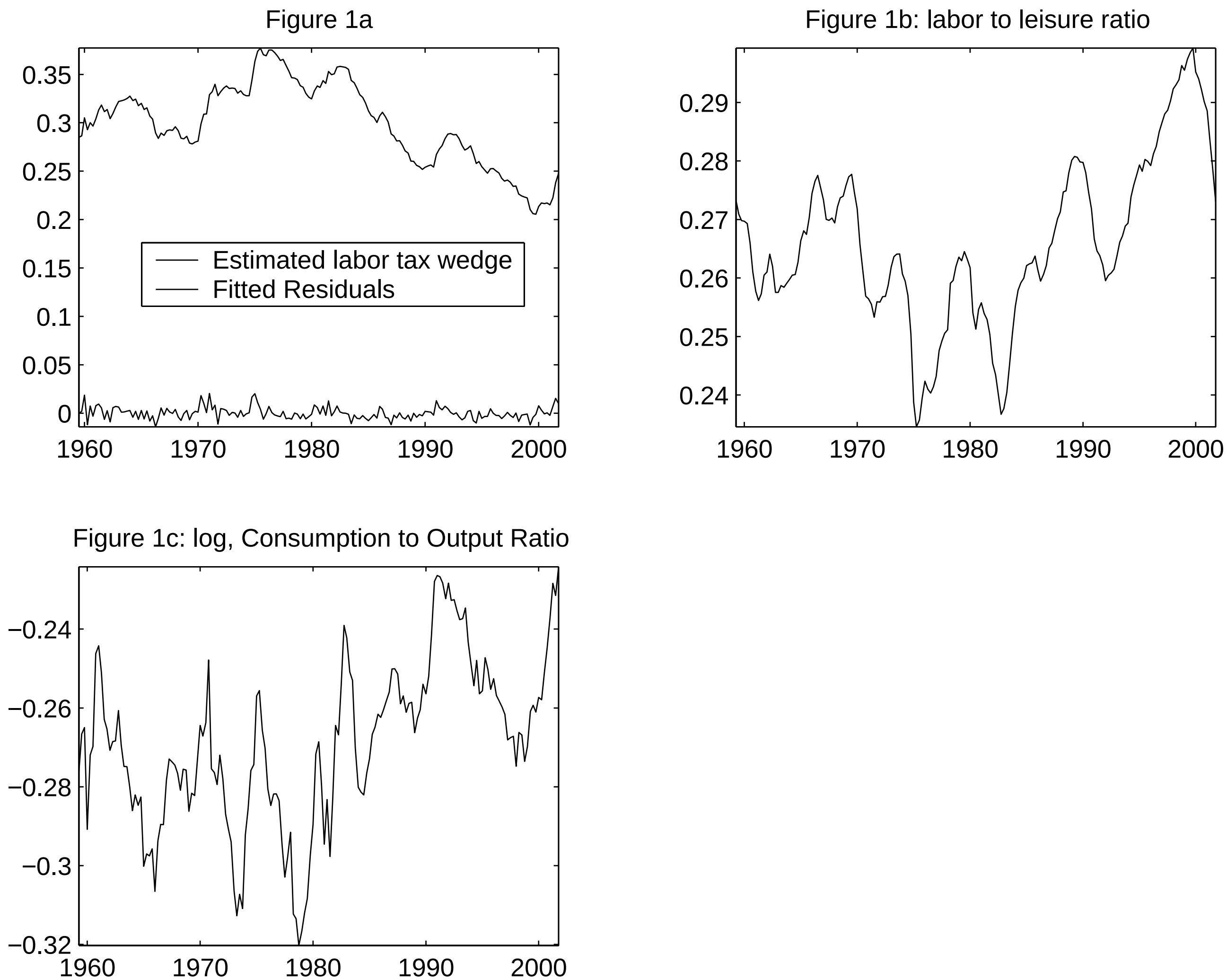
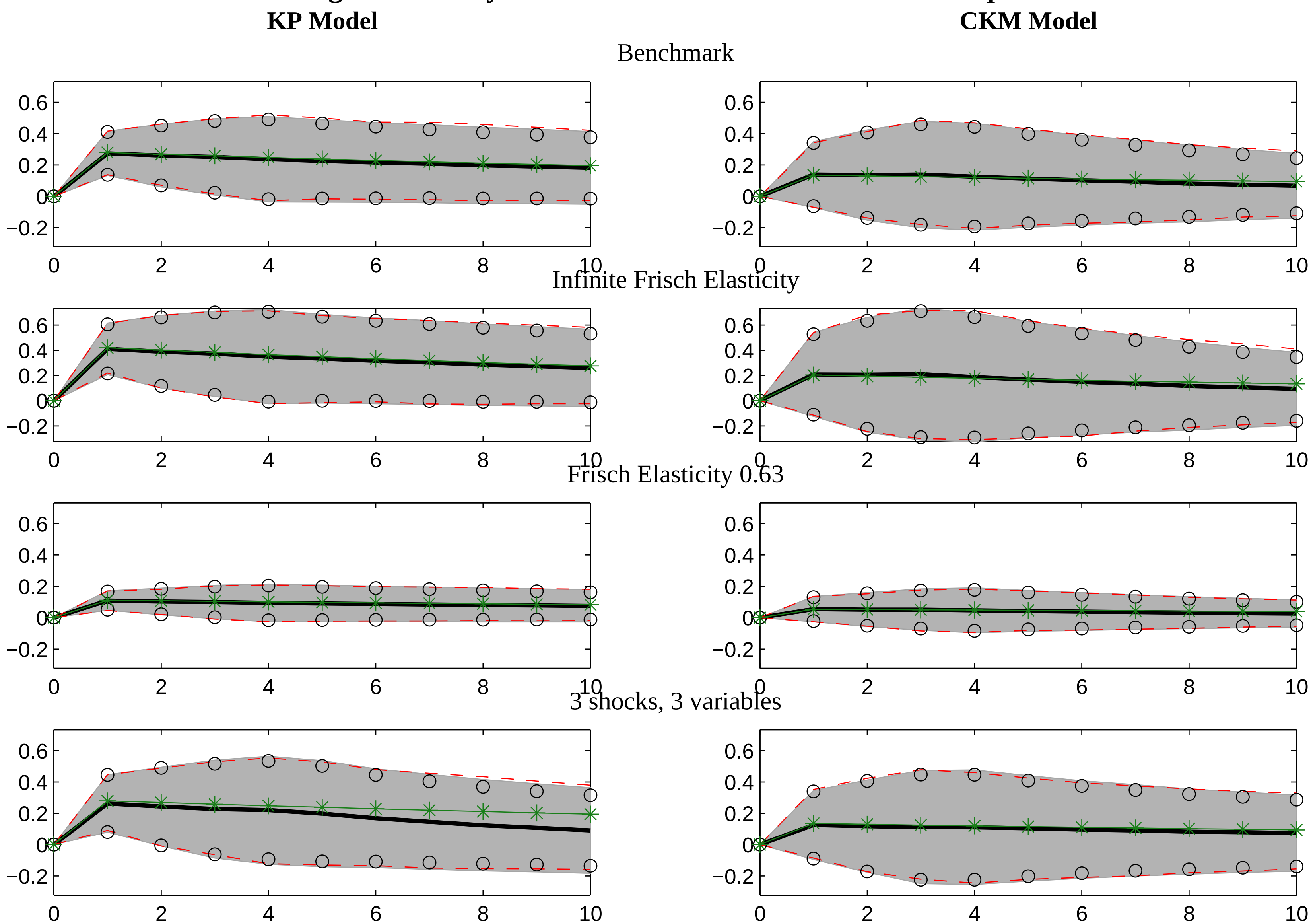


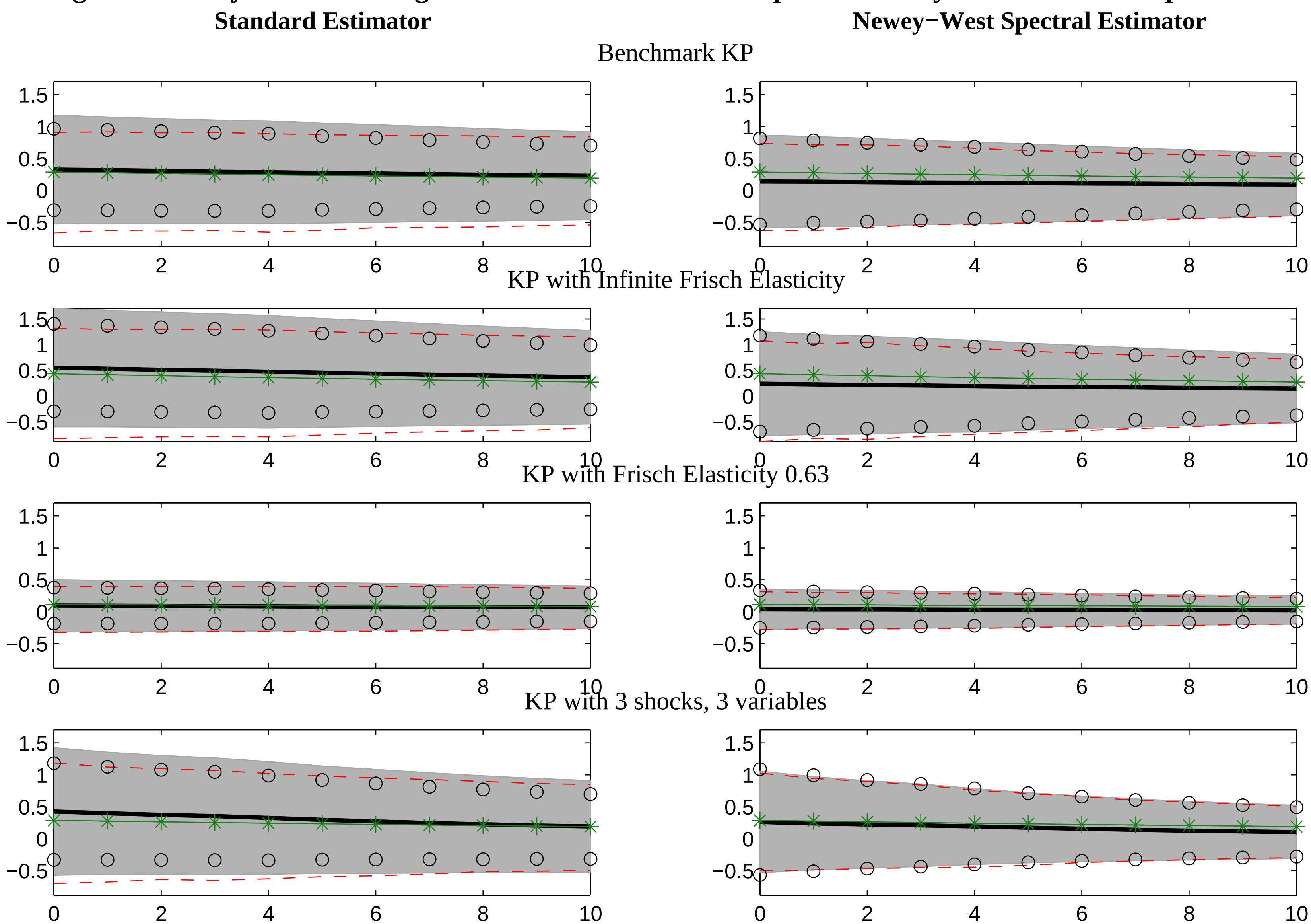
Figure 2: Analysis of Short−Run Identification Assumption



Note: hours response, in percent terms, to a 1.2 (KP) or 0.6 (CKM) percent innovation in technology, Z_t.

Solid line − mean response, Gray area − mean response plus/minus two standard errors,
Starred line − true response, Dashed line − 95.5 percent probability interval of responses,
Circles − average value of econometrician estimated plus/minus two standard errors.

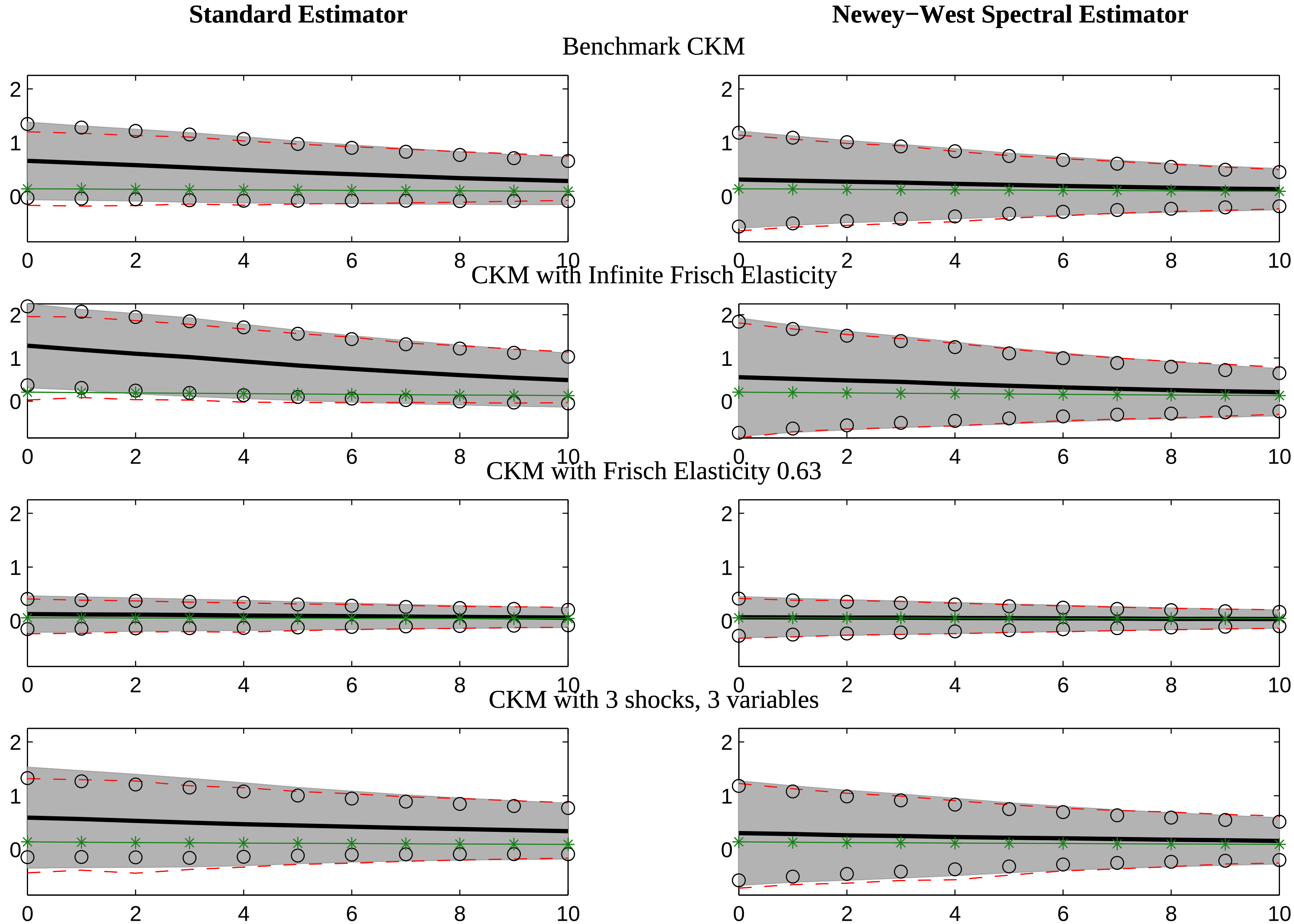
Figure 3: Analysis of the Long−Run Identification Assumption with Kydland−Prescott Specification



Note: hours response, in percent terms, to a 1.2 percent innovation in technology, Z_t.

Solid line − mean response, Gray area − mean response plus/minus two standard errors,
Starred line − true response, Dashed line − 95.5 percent probability interval of responses,
Circles − average value of econometrician estimated plus/minus two standard errors.

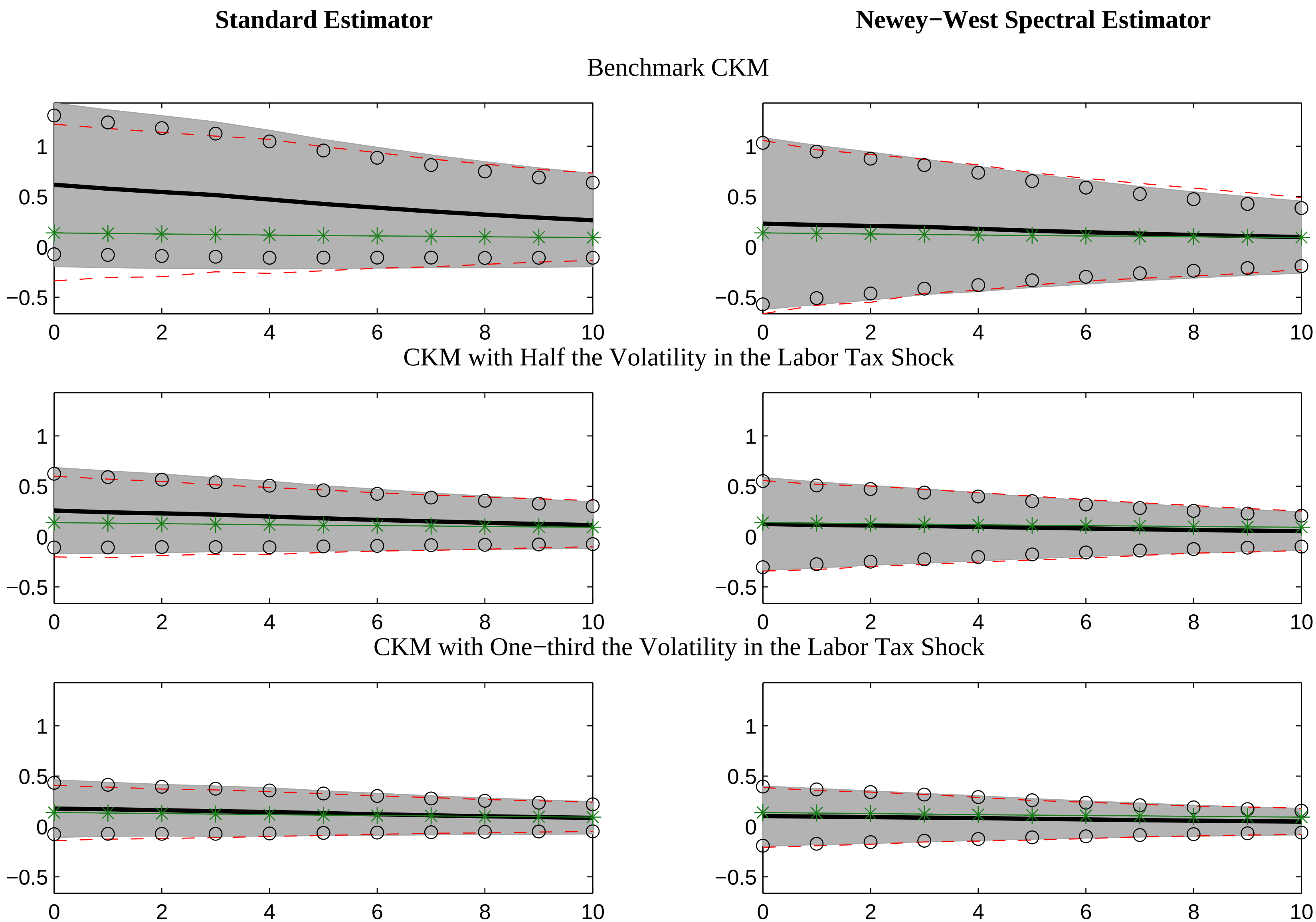
Figure 4: Analysis of the Long−Run Identification Assumption with CKM Specification



Note: hours response, in percent terms, to a 0.6 percent innovation in technology, Z_t.

Solid line − mean response, Gray area − mean response plus/minus two standard errors,
Starred line − true response, Dashed line − 95.5 percent probability interval of responses,
Circles − average value of econometrician estimated plus/minus two standard errors.

Figure 5: Analysis of the Long−Run Identification Assumption with CKM Specification



Note: hours response, in percent terms, to a 0.6 percent innovation in technology, Z_t.

Solid line − mean response, Gray area − mean response plus/minus two standard errors,
Starred line − true response, Dashed line − 95.5 percent probability interval of responses,
Circles − average value of econometrician estimated plus/minus two standard errors.

Figure 6: Ratio of Estimated to True Contemporaneous Impact of Technology on Hours (Benchmark CKM Model)