Dynamic probabilities of restrictions in state space models: An application to the New Keynesian Phillips Curve

Gary Koop
Department of Economics
University of Strathclyde
Scotland
Gary.Koop@strath.ac.uk

Roberto Leon-Gonzalez
National Graduate Institute for Policy Studies
Japan
rlg@grips.ac.jp

Rodney W. Strachan
School of Economics
University of Queensland
Australia
r.strachan@uq.edu.au

November 2007
[Draft: Please do not cite without permission]

Abstract: Empirical macroeconomists are increasingly using models (e.g. regressions or Vector Autoregressions) where the parameters vary over time. State space methods are frequently used to specify the evolution of parameters in such models. In any application, there are typically restrictions on the parameters, having either economic or econometric significance, that a researcher might be interested in. This motivates the question of how to calculate the probability that a restriction holds at a point in time (without assuming the restriction holds at any other point in time). This paper develops methods to answer this question. In particular, the principle of the Savage-Dickey density ratio is used to obtain the time-varying posterior probabilities of restrictions. We use our methods in a macroeconomic application involving the Phillips curve. Macroeconomists are interested in whether the long-run Phillips curve is vertical. This is a restriction for which we can calculate the posterior probability using our methods. Using U.S. data, the probability that this restriction holds tends to be fairly high, but decreases slightly over time (apart from a slight peak in the late 1970s). We also calculate the probability that another restriction, that the NAIRU is not identified, holds. The probability that it holds fluctuates over time with most evidence in favor of the restriction occurring after 1990.

Key Words: Bayesian, state space model, Savage-Dickey density ratio, time varying parameter model.

JEL Classification: C11, C32, E52

1 Introduction

Many recent papers, particularly in the field of macroeconomics, have worked with extensions of regressions or Vector autoregressions (VARs) where the parameters can change over time. With a wide variety of data sets, there is strong empirical evidence that such time-varying parameter (TVP) models are necessary to capture data properties of economic relevance (e.g. the evolution of transmission mechanisms or the processes generating the exogenous shocks). The TVP-VARs of (among many others) Cogley and Sargent (2001, 2005), Koop, Leon-Gonzalez and Strachan (2007) and Primiceri (2005) have the form:

$$\begin{aligned} y_t &= Z_t \alpha_t + \varepsilon_t \\ \alpha_{t+1} &= \alpha_t + \eta_t, \end{aligned}$$

where y_t is a vector of dependent variables and Z_t contains appropriate lags of the dependent variable and deterministic terms.

TVP-VARs are a popular and powerful tool in modern macroeconomic research, however they suffer from some drawbacks. Firstly, TVP-VARs are reduced form models which do not impose any restrictions on the coefficients. However, economic theory often suggests restrictions on α_t and, in traditional VAR applications such restrictions are often tested or imposed. Secondly, TVP-VARs can be over-parameterized, rendering it hard to obtain precise inference. Imposing restrictions on TVP-VARs can reduce such over-parameterization problems. These considerations suggest that developing methods for calculating the probability of restrictions on the coefficients in state space models such as the TVP-VAR is of interest. This is what we do in this paper using Bayesian methods. Since α_t varies over time, the probability of a restriction holding can also vary over time. This motivates our terminology "dynamic posterior probabilities" of a restriction holding.

In this paper, we develop methods for calculating such dynamic posterior probabilities for linear restrictions on state variables such as α_t using output from standard simulation algorithms for state space models (see, e.g., Durbin and Koopman, 2002) and the principle of the Savage-Dickey density ratio (SDDR), see Verdinelli and Wasserman (1995). Although our focus is on macroeconomic applications, we stress that these methods are of more general use for state space modeling.

There is a huge variety of types of restrictions that might be of interest in macroeconomics. In the empirical part of this paper, we consider an application to the New Keynesian Phillips curve and show how the underlying theory implies certain restrictions. Another example is King, Plosser, Stock and Watson (1991). This used the balanced growth hypothesis to motivate restrictions on the relationships between income, investment and consumption. Another example is Lettau and Ludvigson (2004), which investigates the relationship among asset wealth, income and consumption. As discussed in that paper and further investigated in Koop, Potter and Strachan (2007), an empirically-motivated restriction (a weak exogeneity restriction) is important in obtaining a key result. These papers all use models where the parameters are constant over time and, thus, they can simply calculate the probability that a restriction holds using familiar statistical methods. However, to our knowledge there is little work which focusses on calculating the probability that restrictions hold in TVP models. Given this lack and the empirical importance of TVP models, it is natural that we consider this problem and develop methods of obtaining the posterior probability that a restriction holds at a particular point of time, without requiring the restriction to be imposed at any other time.

It is also important to stress that we are not developing methods for, e.g., the recursive testing of a restriction. The dynamic posterior probabilities of restrictions we obtain are conditional upon the full sample and, in this sense, allow for more efficient inference. That is, if we write the restriction being tested as $A\alpha_t = \alpha^*$ for some known matrix A and vector α^*, we calculate $\Pr(A\alpha_t = \alpha^*|y)$ where $y = (y_1', .., y_T')'$ denotes the full sample. So we are computing the probability that the restriction holds at time t, but does not hold elsewhere, and this probability is conditional upon all information available. We are not calculating $\Pr(A\alpha_t = \alpha^*|y_1, .., y_t)$ as would be done in a recursive approach (although, as noted below, it would be easy to modify our methods to calculate this latter probability). Thus, our methods address questions of the form: "What is the probability that a restriction/theory holds at time t, given all the information in our data set?".

The structure of the paper is as follows. In Section 2 we present the state space model and associated posterior simulation methods based on the Kalman filter. In Section 3 we present the basic ideas behind the Savage-Dickey density ratio and demonstrate how this, combined with posterior simulator output, can be used to calculate the dynamic posterior probabilities of any linear restriction on the states. Section 4 uses these methods in an application relating to the New Keynesian Phillips Curve. Section 5 concludes.

2 Basic Posterior Results for the State Space Model

We begin by defining our model (which is slightly more general than the one discussed in the introduction). This is a standard state space model, although we focus on the TVP interpretation of it. Let y_t for $t = 1, .., T$ denote a vector of observations on p dependent variables and $y = (y_1', y_2', \ldots, y_T')'$ be the $Tp \times 1$ vector of all the observations on all dependent variables. The measurement and state equations in the state space model are given by

$$
\begin{aligned}
y_t &= Z_t \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H_t) & (1)\\
\alpha_{t+1} &= T_t \alpha_t + \eta_t \qquad \eta_t \sim N(0, Q_t),\ t = 1, \ldots, T, \text{ and} & (2)\\
\alpha_1 &\sim N(a_1, P_1). & (3)
\end{aligned}
$$

We further assume ε_t and η_s are independent for all t and s. The vector Z_t is of dimension $1 \times m$ such that α_t is a $m \times 1$ vector, T_t is an $m \times m$ matrix and η_t is an $n \times 1$ vector.

Many popular macroeconomic models fit in this framework. A TVP regression model (such as we use in our empirical section), has $p = 1$ and Z_t containing observations on explanatory variables and their lags and lagged dependent variables. The TVP-VARs of Cogley and Sargent (2005) and Primiceri (2005) have $T_t = I$ and Z_t containing an intercept and suitable lags of the dependent variables and H_t taking a particular multivariate stochastic volatility form. Cogley and Sargent (2001) adopts the same form with the additional assumption that H_t is a constant. Adding regression effects is a trivial extension that we will not consider to keep the notation uncluttered. That is, adding a term of the form $X_t\beta$ to (1) can easily be done. It is also worth noting that, in most macroeconomic applications it is important to allow for stochastic volatility in ε_t. In our empirical application, we do allow for stochastic volatility. However, since our theoretical derivations relate to α_t, we will simply leave H_t as the time-varying measurement error variance, unspecified.

Posterior simulation of state space models can be done using several algorithms (e.g. in this paper we use the algorithm of Durbin and Koopman, 2002). This can be used as a block in a Markov Chain Monte Carlo (MCMC) algorithm for carrying out posterior simulation in the state space model. That is, conditional on H_t, T_t, Q_t, a_1 and P_1, the algorithm of Durbin and Koopman can be used to draw α_t for $t = 1, .., T$. Conditional on

α_t, posterior draws of H_t, T_t, Q_t, a_1 and P_1 can be taken (although often some of these are set to pre-selected values such as $T_t = I$). Given conditionally conjugate priors, the formulae for the posterior draws of H_t, T_t, Q_t, a_1 and P_1 are standard (see, e.g., Koop, 2003, pages 196-197) so we will not discuss them here (see the appendix for details) and will not explicitly list these conditioning arguments in our subsequent discussion of the conditional posterior of α_t.

For this model, the posterior distribution of the vector $\alpha = (\alpha_1', .., \alpha_T')'$ is Normal and can be pinned down by its mean and variance. Standard Kalman filtering and smoothing methods can be used to obtain these (see, e.g., Durbin and Koopman, 2001, chapter 4). To fix notation, we describe the relevant steps of the filter and smoother here.

The Kalman filter is a sequence of recursive operations running from $t = 1$ to $t = T$ that simplify and speed algorithms for drawing from α_t. The Kalman filter provides $E(\alpha_t|y_1, .., y_t) = a_t$ and variance $var(\alpha_t|y_1, .., y_t) = P_t$ by evaluating:

$$\begin{aligned} v_t &= y_t - Z_t a_t \qquad F_t = Z_t P_t Z_t' + H_t \\ K_t &= T_t P_t Z_t' F_t^{-1} \qquad L_t = T_t - K_t Z_t \\ a_{t+1} &= T_t a_t + K_t v_t \qquad P_{t+1} = T_t P_t L_t' + Q_t. \end{aligned}$$

As a digression, in the introduction, we distinguished our question of interest: "What is the probability that a restriction/theory holds at time t, given all the information in our data set?" from one that could be answered by recursive estimation techniques ("What is the probability that a restriction/theory holds at time t, given the information available at time t?"). If one were interested in the latter question, we could use $E(\alpha_t|y_1, .., y_t)$ and $var(\alpha_t|y_1, .., y_t)$ along with the SDDR (described in the next section) to obtain recursive dynamic posterior probabilities of restrictions. That is, we could compute $\Pr(A\alpha_t = \alpha^*|y_1, .., y_t)$ using Kalman filter output and the SDDR.

Our objective, however, is to compute the posterior probability that the restrictions hold at time t given the full sample: $\Pr(A\alpha_t = \alpha^*|y)$. For this we require knowledge of $\widehat{\alpha}_t = E(\alpha_t|y)$ and $V_t = var(\alpha_t|y)$. To obtain these, we must run the state smoother which is a series of recursions that run in reverse order from the Kalman filter (i.e. $t = T$ to $t = 1$). These recursions are:

$$
\begin{aligned}
r_{t-1} &= Z_t' F_t^{-1} v_t + L_t' r_t \qquad N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \\
\widehat{\alpha}_t &= a_t + P_t r_{t-1} \qquad V_t = P_t - P_t N_{t-1} P_t
\end{aligned}
$$

for $t = T, \ldots, 1$.

We now have described the posterior for the state space model (conditional on other model parameters) and, in particular, the posterior means and variances, $\widehat{\alpha}_t = E(\alpha_t|y)$ and $V_t = var(\alpha_t|y)$. We now turn to the main focus of this paper: using the Savage-Dickey density ratio to compute dynamic posterior probabilities of restrictions on α_t.

3 Calculating Dynamic Posterior Probabilities of Restrictions using the SDDR

The SDDR is a convenient way of calculating the Bayes factor comparing a restricted to an unrestricted model (call them M_R and M_U). We begin with a general statement of the SDDR before applying it to the state space model. Suppose M_U, has a parameter vector $\theta = (\omega', \psi')'$. The prior for this model is $p(\omega, \psi|M_U)$. The restricted version of the model, M_R, has $\omega = \omega_0$ where ω_0 is a vector of constants. The parameters in ψ are left unrestricted in each model. The prior for M_R is $p(\psi|M_R)$. Suppose the priors in the two models satisfy:

$$
p(\psi|\omega = \omega_0, M_U) = p(\psi|M_R), \tag{4}
$$

then the Bayes factor comparing M_R to M_U is:

$$
BF = \frac{p(\omega = \omega_0|y, M_U)}{p(\omega = \omega_0|M_U)}, \tag{5}
$$

where $p(\omega = \omega_0|y, M_U)$ and $p(\omega = \omega_0|M_U)$ are the unrestricted posterior and prior for ω evaluated at the point ω_0. Equation (5) is referred to as the SDDR. The proof of this theorem is given in many places, including Verdinelli and Wasserman (1995). Note that (4) almost never restricts the form of the prior. For instance, if (as is commonly done), conditional on ω, the same prior is used for parameters which are common to both models, then (4) is satisfied. In fact, it is an even weaker restriction than this, requiring only the same prior for common parameters to hold at one point (ω_0) in the parameter space. In this paper, we assume (4) is satisfied. In the rare cases where this condition is not reasonable, Verdinelli and Wasserman (1995) provide a similar, but slightly more complicated expression for the Bayes factor and a trivial extension of our methods is required.

Crucially, the SDDR involves only manipulations involving the posterior and prior for the unrestricted model which, in our case is the state space model described in the preceding section. Other approaches to Bayes factor calculation (e.g. calculating the marginal likelihood separately for the restricted and unrestricted models) would require estimation of the restricted model. For state space models with restrictions of the sort we consider (i.e. restrictions imposed at time t, but not necessarily at other times), we know of no available Bayesian methods for estimating the restricted model.

Here we derive the SDDR for restrictions on α_t of the form $A\alpha_t = \alpha^*$ where A is a known $q \times m$ matrix and α^* is a known $q \times 1$ vector. Our derivations are greatly simplified due to the fact that α_t and, thus, $A\alpha_t$, have priors and posteriors which, conditional on the other model parameters, are Normal.

Using the SDDR, the Bayes factor comparing the restricted to the unrestricted model is:

$$BF = \frac{p\left(A\alpha_t = \alpha^*|y\right)}{p\left(A\alpha_t = \alpha^*\right)}, \tag{6}$$

where the posterior and the prior in (6) are those for the unrestricted state space model described in the previous section.

We first discuss evaluating the posterior term in the numerator of (6). To make the notation compact, let $\phi = \{H_t, T_t, Q_t, a_1, P_1\}$ denote all the parameters (or pre-selected constants) in the state space model, other than the states themselves. Our results of the previous section tell us that, even though $p\left(\alpha_t = \alpha^*|y\right)$ does not have a convenient analytical form, $p\left(\alpha_t = \alpha^*|y, \phi\right)$ is Normal. To be precise,

$$p\left(\alpha_t|y, \phi\right) = \left(2\pi\right)^{-m/2}\left|V_t\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\alpha_t - \widehat{\alpha}_t\right)' V_t^{-1}\left(\alpha_t - \widehat{\alpha}_t\right)\right\}.$$

Using standard results for the multivariate Normal distribution, we have

$$p\left(A\alpha_t|y, \phi\right) = \left(2\pi\right)^{-m/2}\left|AV_tA'\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(A\alpha_t - A\widehat{\alpha}_t\right)'\left(AV_tA'\right)^{-1}\left(A\alpha_t - A\widehat{\alpha}_t\right)\right\},$$

and, thus,

$$p\left(A\alpha_t=\alpha^*|y,\phi\right)=(2\pi)^{-m/2}\left|AV_tA'\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\alpha^*-A\widehat{\alpha}_t\right)'\left(AV_tA'\right)^{-1}\left(\alpha^*-A\widehat{\alpha}_t\right)\right\}. \tag{7}$$

Given output from our MCMC algorithm ($\phi^{(r)}$ for $r=1,...,R$), we can approximate the numerator of (6) by:

$$\widehat{p}\left(A\alpha_t=\alpha^*|y\right)=\frac{1}{R}\Sigma_{r=1}^{R}p\left(A\alpha_t=\alpha^*|y,\phi^{(r)}\right), \tag{8}$$

using (8). As usual in MCMC algorithms, R can be chosen to ensure any desired accuracy of approximation and MCMC diagnostics used to monitor convergence.

We now turn to the prior term in the denominator of (6). We adopt a similar strategy as with the numerator, except using prior simulator output. That is, $p\left(\alpha_t=\alpha^*\right)$ does not have a convenient analytical form, but $p\left(\alpha_t=\alpha^*|\phi\right)$ is Normal. Thus, if we have output from a prior simulator ($\phi^{(s)}$ for $s=1,...,S$), we can approximate the denominator of (6) by:

$$\widehat{p}\left(A\alpha_t=\alpha^*\right)=\frac{1}{S}\Sigma_{s=1}^{S}p\left(A\alpha_t=\alpha^*|\phi^{(s)}\right). \tag{9}$$

To derive $p\left(A\alpha_t=\alpha^*|\phi^{(s)}\right)$, we use (2) and (3). Firstly, we rewrite (3) as $\alpha_1=a_1+P_1^{1/2}z$ where $z\sim N\left(0,I_m\right)$. The expressions in (2) and (3) define the hierarchical prior for the states. Using these, we obtain

$$\begin{aligned}\alpha_t &= \Pi_{i=1}^{t-1}T_i\alpha_1+\Sigma_{j=1}^{t-1}\Pi_{i=j+1}^{t-1}T_i\eta_j\\ &= \Pi_{i=1}^{t-1}T_ia_1+\Pi_{i=1}^{t-1}T_iP_1^{1/2}z+\Sigma_{j=1}^{t-1}\Pi_{i=j+1}^{t-1}T_i\eta_j.\end{aligned}$$

Since α_t is a linear function of Normally distributed independent random variables, α_t is Normally distributed with prior mean

$$\underline{\alpha}_t=E\left(\alpha_t\right)=\Pi_{i=1}^{t-1}T_ia_1$$

and prior variance

$$var\left(\alpha_t\right)=\underline{V}_t=\Sigma_{j=1}^{t-1}\Pi_{i=j+1}^{t}T_iQ_j\left(\Pi_{i=j+1}^{t}T_i\right)'+\Pi_{i=1}^{t-1}T_iP_1\left(\Pi_{i=1}^{t-1}T_i\right)'.$$

Thus, the prior (conditional on ϕ) can be written as:

$$p\left(\alpha_t|\phi\right)=(2\pi)^{-m/2}\left|\underline{V}_t\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\alpha_t-\underline{\alpha}_t\right)'\underline{V}_t^{-1}\left(\alpha_t-\underline{\alpha}_t\right)\right\}. \tag{10}$$

Thus, the necessary term in (9) is:

$$p\left(A\alpha_t=\alpha^*|\phi\right)=(2\pi)^{-m/2}\left|A\underline{V}_tA'\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\alpha^*-A\underline{\alpha}_t\right)'\left(A\underline{V}_tA'\right)^{-1}\left(\alpha^*-A\underline{\alpha}_t\right)\right\}. \tag{11}$$

In summary, given output from a posterior simulator and (7) we can obtain the numerator of the SDDR given in (6). Output from a prior simulator and (11) can be used to approximate the denominator. Given the Bayes factor comparing the restricted and unrestricted models, the posterior probability of the restriction can be calculated in the standard way.

Finally, a simple extension of our methods is worth noting. In this paper, we focus on restrictions at a single point in time (i.e. $p\left(A\alpha_t=\alpha^*|y\right)$). The derivations in this section have been based on the fact that the simulation smoother directly provides us with the mean and variance of the Normal marginal distribution $p\left(\alpha_t|y,\phi\right)$. However, there might be some cases where the reader is interested in restrictions involving more than one point in time. Suppose for instance that interest centers on $p\left(\alpha_t=\alpha_{t+1}=\alpha^*|y,\phi\right)$. To evaluated the SDDR for this restriction, the joint p.d.f. $p\left(\alpha_t,\alpha_{t+1}|y,\phi\right)$ must be evaluated at the restriction. The simulation smoother does not directly provide us with this p.d.f., but relevant formulae can developed in a straightforward manner. That is, since $p\left(\alpha|y,\phi\right)$ is Normal any relevant conditional or marginal will be Normal. Since $p\left(\alpha|y,\phi\right)$ can be evaluated in one of several standard ways (see, for instance, Carter and Kohn, 1994, Lemma 2.1) any relevant conditional or marginal can be evaluated using methods which involve the Kalman filter and simulation smoother in a slightly more complicated way than we have done here.

4 Dynamic Evidence on the Phillips Curve

There has been a resurgence of interest in the Phillips curve. Influential examples of this work include King and Watson (1994), Staiger, Stock and Watson (1997) and Sargent, Williams and Zha (2006). Of particular importance in these studies are the questions of whether there is a trade-off between unemployment and inflation and, if not, what is the non-accelerating inflation rate of unemployment (NAIRU). Sargent, Williams and Zha (2006) argue that, through the 1970s, US monetary authorities observed evidence that suggested there was a trade-off between unemployment and inflation. In the following decade this trade-off vanished and so the monetary authorities cut inflation as they expected no cost in terms of unemployment. Thus, we have a restriction of policy interest (i.e. that the long run Phillips curve is vertical) and other authors have presented evidence that the support for this restriction is changing over time (although this evidence involves looking at point estimates and credible intervals for functions of the time-varying parameters). This is an ideal setup for using our new approach to calculating the dynamic posterior probability of this restriction.[1]

Staiger, Stock and Watson (1997) use a time varying model with a vertical Phillips curve imposed at all points in time. This specification coincides with the expectations augmented Phillips curve and permits estimation of the NAIRU. Since they use a time varying model, they are able to report time varying estimates of the NAIRU. With this model, there exist points in the parameter space at which the NAIRU is not unidentified. This provides us with a second restriction for which we can calculate dynamic posterior probabilities.

We use a state space model of the same general form as Staiger, Stock and Watson (1997), although we do not always impose a vertical long run Phillips curve. We use data from 1953Q1 through 2006Q2 on the unemployment rate (seasonally adjusted civilian unemployment rate, all workers over age 16) and inflation rate (the annual percentage change in a chain-weighted GDP price index).[2]

If there were a trade-off between inflation and unemployment, we would expect peaks in unemployment to coincide or lead troughs in inflation. Figure (1) shows the actual behavior of these series. Peaks in unemployment seem to be followed by troughs in inflation. However, the strength of this relationship seems to vary over time. Through the 1970s, the peaks in unemployment appear almost coincident with peaks in inflation. It is likely that the relationship between the level of the two series has changed over time. At least it appears that there is not a consistently negative relationship between the two. Considering the relationship between the change in inflation and the level of unemployment (not plotted here), however, there does seem to be a negative relationship. This would be the case if the long run Phillips curve were vertical. But again, this relationship between the change in inflation and the level of unemployment appears to have changed - at least during the 1970s.

As discussed in King and Watson (1994) and in Sargent, Williams and Zha (2006), there are different ways of specifying the relationship between unemployment (u_t) and inflation (π_t). We use two specifications. In the first of these (which does not impose the vertical Phillips curve), the inflation rate (π_t) is the dependent variable and, thus, in terms of our notation for the state space model: $y_t = \pi_t$. In the second (which does impose the

[1] It might seem a little incongruous that a long run relation might hold at one point in time and not another. It is probably more informative to think of the long run relation as an equilibrium towards which the model attracts the variables at that point in time.

[2] The data were obtained from the Federal Reserve Bank of St. Louis website, http://research.stlouisfed.org/fred2/.

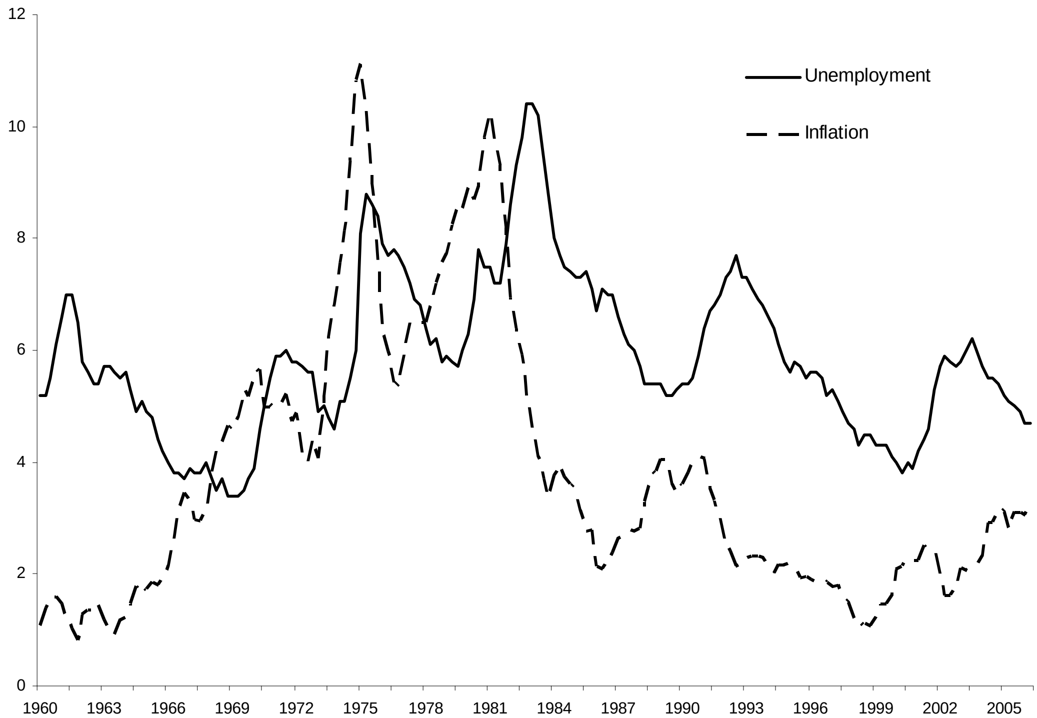

Figure 1: Inflation and Unemployment levels.

vertical Phillips curve and, thus, allows for estimation of the NAIRU), the change in inflation is the dependent variable ($y_t = \Delta\pi_t$). In the first specification the explanatory variables are $Z_t = (1, \pi_{t-1}, \pi_{t-2}, u_t, u_{t-1}, u_{t-2})$. In the second. $Z_t = (1, -\Delta\pi_{t-1}, u_t, u_{t-1}, u_{t-2})$. Note that we are always using two lags of both variables, the same choice as Staiger, Stock and Watson (1997) and Sargent, Williams and Zha (2006).

To further explain these two specifications, ignore for now the time variation in parameters. Our first specification is an unrestricted one where inflation depends on lags of itself, current unemployment and lags of unemployment:

$$\pi_t = \beta_1 + \beta_2\pi_{t-1} + \beta_3\pi_{t-2} + \beta_4 u_t + \beta_5 u_{t-1} + \beta_6 u_{t-2} + \varepsilon_t. \tag{12}$$

A vertical long run Phillips curve is implied by the restriction that the sum of the coefficients on the lags of inflation sum to one: $\beta_2 + \beta_3 = 1$. This is the first restriction of interest.

If we impose this restriction, we obtain:

$$\Delta\pi_t = \beta_1 - \beta_3 \Delta\pi_{t-1} + \beta_4 u_t + \beta_5 u_{t-1} + \beta_6 u_{t-2} + \varepsilon_t. \tag{13}$$

Model (13) is the second specification we use in our empirical work. To see how it relates to the issue of the identification of NAIRU, we can rewrite it as:

$$\Delta\pi_t = -\beta_3 \Delta\pi_{t-1} + \beta_4 (u_t - \underline{u}) + \beta_5 (u_{t-1} - \underline{u}) + \beta_6 (u_{t-2} - \underline{u}) + \varepsilon_t, \tag{14}$$

where $\underline{u}$ is the NAIRU. That is, (14) embeds the idea that it is deviations of unemployment from its natural rate which trigger inflation. The relationship between the coefficients in (13) and (14) is $\beta_1 = -\underline{u}(\beta_4 + \beta_5 + \beta_6)$. From this, it can be seen that the NAIRU can be estimated from (13) provided $\beta_4 + \beta_5 + \beta_6 \neq 0$. The NAIRU is not identified if $\beta_4 + \beta_5 + \beta_6 = 0$ which is our second restriction of interest.

In terms of the notation used for our state space model, we extend (12) to allow for time variation in coefficients by putting t subscripts on the coefficients and letting: $\alpha_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \beta_{4,t}, \beta_{5,t}, \beta_{6,t})'$ and calculate the dynamic posterior probability that $\beta_{2,t} + \beta_{3,t} = 1$. We extend (13) to allow for time variation in coefficients by setting: $\alpha_t = (\beta_{1,t}, \beta_{3,t}, \beta_{4,t}, \beta_{5,t}, \beta_{6,t})'$ and calculate the dynamic posterior probability that $\beta_{4,t} + \beta_{5,t} + \beta_{6,t} = 0$. Since these are both linear restrictions on the states, our SDDR approach is directly applicable.

Before presenting empirical results, we must extend our model in a manner that is empirically important (but not of great relevance for either the economic or econometric theory discussed in this paper). It is well documented that the variance of inflation has changed markedly over time and, thus, stochastic volatility must be added to the model. We do this in a standard way and provide exact details in the appendix. In terms of the state space model in (1) and (2), the stochastic volatility assumption specifies H_t. We further assume $T_t = I$ and $Q_t = Q$. Details on priors and posterior simulation for the unknown parameters are given in the appendix. Suffice it to note here that we use a training sample prior of the sort commonly-used by empirical macroeconomist (see, e.g., Cogley and Sargent, 2001, 2005, Koop, Leon-Gonzalez and Strachan, 2007 and Primiceri, 2005).

We begin with our first specification, (12). Before presenting evidence on whether the long run Phillips curve is vertical, we present information on the measurement error standard deviation. Figure 2 presents the posterior median of this error standard deviation (based on a standard stochastic volatility model) along with the 16^{th} and 84^{th} percentiles. The general pattern resembles that found in earlier studies using TVP models that allow both regression coefficients and error variances to vary (e.g., Primiceri, 2005, and Koop, Leon-Gonzalez and Strachan, 2007). There is a clear increase in volatility around the mid 1970s and another increase around 1980. Another feature that is in common with Koop, Leon-Gonzalez and Strachan (2007), which uses more recent data than Primiceri (2005), is the slight upward trend in volatility after 2000.

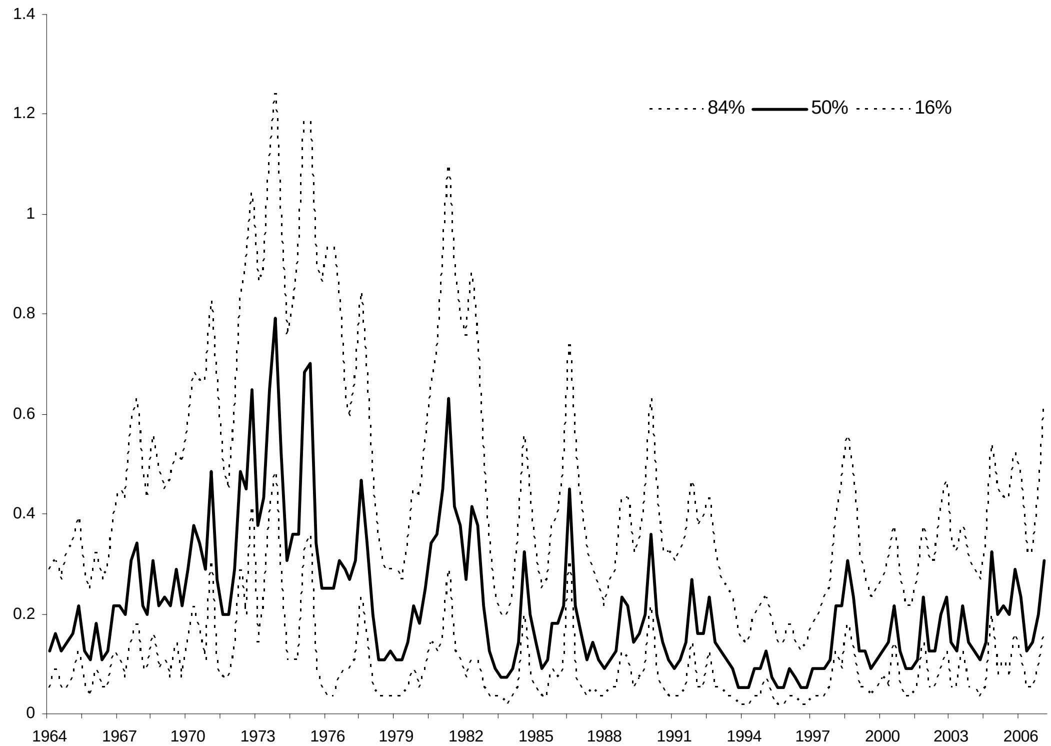

Figure 2: This figure plots the posterior median of the standard deviation (thick black line) of the error in the TVP variant of equation (12). Also plotted are the 16^{th} and 84^{th} percentiles (thin dashed lines) of the standard deviation.

We now turn to our first restriction of interest: that long run Phillips curve is vertical and, hence, that there exists a NAIRU. Figure 3 shows the dynamic posterior probability of this restriction. The probability that this restriction holds tends to gradually fall from about 80% to 65% over the sample, with the exception of a peak of about 85% near the end of the 1970s. To give some measure of the strength of these results we also report the natural log of the posterior odds ratio, $\ln(POR_t)$, and the log to the base 10 of the posterior odds ratio, $\log_{10}(POR_t)$. These posterior odds ratios are calculated assuming a prior odds ratio of one. Kass and Raftery (1995) justifies and provides some useful rules of thumb for interpreting posterior odds ratios. Note that our values of $\ln(POR_t)$ vary from 0.6 to 1.7. This is the region in which Kass and Raftery's rules of thumb suggest the evidence in support of the restriction is at worst 'Not worth more than a bare mention' and at best 'Positive'. Kass and Raftery also provides slightly different rules of thumb in terms of $\log_{10}(POR_t)$. Our values vary from 0.25 to 0.75 which their rules of thumb say vary between 'Not worth more than a bare mention' and at best 'Substantial'. In short, we are finding some weak evidence in favor of the restriction that the long run Phillips curve is vertical, but this evidence is generally declining over time.

Since the evidence (although far from compelling) is always weakly in favor of the restriction that the long run Phillips curve is vertical, we proceed to our second model which is the TVP variant of (13). Remember that this allows for the calculation of the NAIRU provided that $\beta_{4,t} + \beta_{5,t} + \beta_{6,t} \neq 0$. The 16^{th}, 50^{th} and 84^{th} percentiles of the posterior distribution for NAIRU are plotted in Figure 4 along with the actual unemployment rate. We see a clear rise in the 1970s, followed by a gradual decline over the 1980s and 1990s. This same pattern was found by Staiger, Stock and Watson (1997).

The dynamic posterior probabilities that $\beta_{4,t} + \beta_{5,t} + \beta_{6,t} = 0$ are plotted in Figure 5, along with the natural and base ten logarithms of the posterior odds ratio. The dynamic posterior probabilities that the NAIRU is not identified are much more erratic (particularly in the period from the mid 1970s through mid 1980s) than those we presented for the long run Phillips curve. Around the mid 1970s and again in the early 1980s there is evidence which approaches what Kass and Raftery's rules of thumb would say were 'Positive' and 'Substantial' evidence against the restriction. The probability that the restriction holds falls as low as 28% in 1975. However, at other periods, particularly after 1990 the evidence becomes 'Positive' for the restriction.

5 Conclusions

In this paper, we have developed methods for calculating the dynamic posterior probability of restrictions on states in a standard state space model. This is of interest to the empirical macroeconomist since TVP-VAR and

Figure 3: This figure plots the dynamic posterior probability (thick black line) of the restriction that the long run Phillips curve is vertical. The thin dashed line plots the natural log of the posterior odds ratio, $\ln(POR_t)$, and the thin unbroken line is the log to the base 10 of the posterior odds ratio, $\log_{10}(POR_t)$.

other time-varying parameter models are state space models and restrictions on the states are often suggested by economic theory. Our method for calculating the dynamic posterior probabilities are based on the Savage-Dickey density ratio and can be easily calculated using the output from standard MCMC algorithms for state space models.

We include an empirical application involving the unemployment and inflation rates. Our methods are used to calculate the probability that the long-run Phillips curve is vertical at each point in time. The probability that this restriction holds tends to be high, but varies slightly over time. We also calculate the probability that another restriction, that the NAIRU is not identified, holds. There is less evidence for this restriction, but the probability that it holds does fluctuate over time.

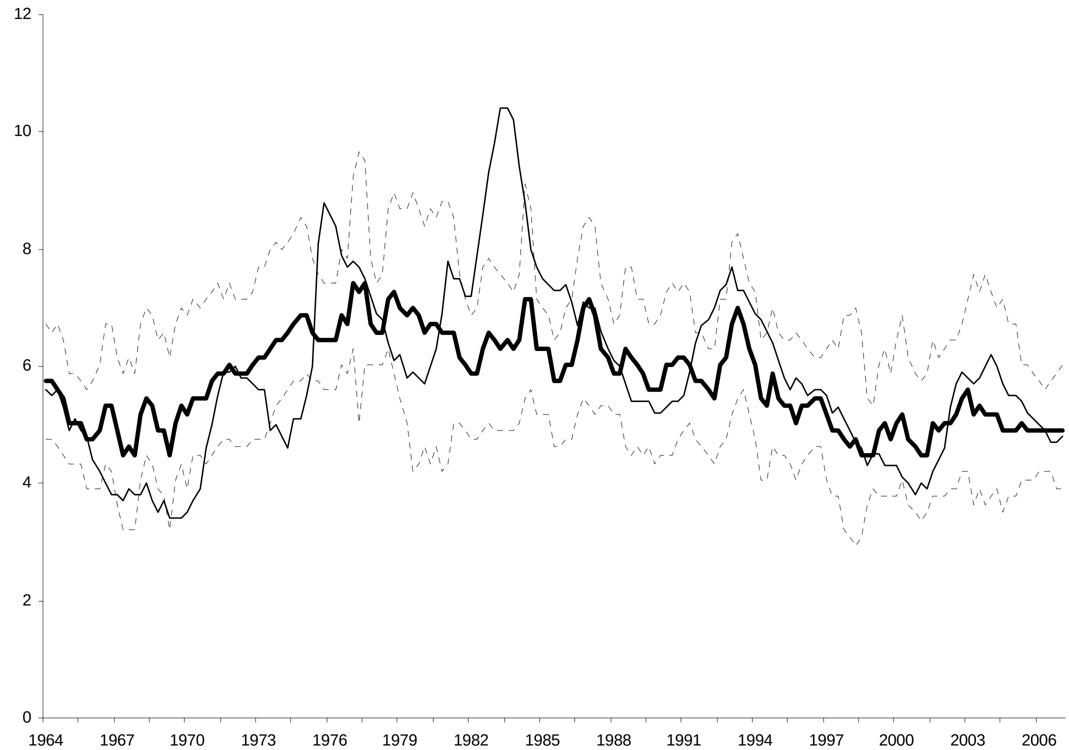

Figure 4: This figure plots the median (thick black line) and 16^{th} and 84^{th} percentiles (thin dashed line) of the posterior distribution for NAIRU. Also plotted is the actual unemployment rate (thin unbroken line).

References

Durbin, J. and Koopman, S., 2001, *Time Series Analysis by State Space Methods* (Oxford University Press, Oxford).

Durbin, J. and Koopman, S., 2002, A simple and efficient simulation smoother for state space time series analysis, *Biometrika*, 89, 603-616.

Cogley, T. and Sargent, T., 2001, Evolving post-World War II inflation dynamics, *NBER Macroeconomic Annual*, 16, 331-373.

Cogley, T. and Sargent, T., 2005, Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S, *Review of Economic Dynamics*, 8, 262-302.

Kass, R. and Raftery, A., 1995, Bayes Factors, *Journal of the American Statistical Association*, 90, 773-795.

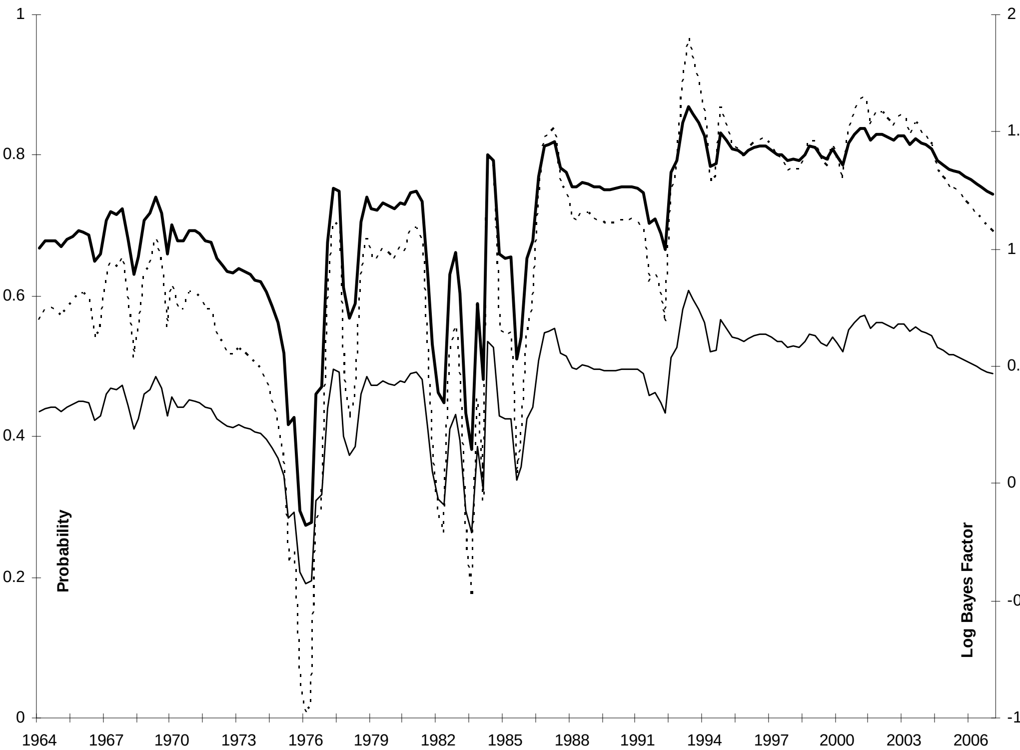

Figure 5: This figure plots the dynamic posterior probability (thick black line) that the NAIRU is not identified. The thin dashed line plots the natural log of the posterior odds ratio, $\ln(POR_t)$, and the thin unbroken line is the log to the base 10 of the posterior odds ratio, $\log_{10}(POR_t)$.

Kim, S., Shephard, N. and Chib, S., 1998, Stochastic volatility: likelihood inference and comparison with ARCH models, *Review of Economic Studies*, 65, 361-93.

King, R., Plosser, C., Stock, J. and Watson, M., 1991, Stochastic trends and economic fluctuations, *The American Economic Review*, 81, 819-840.

King, R. and Watson, M., 1994, The post-war U.S. Phillips curve: a revisionist econometric history, *Carnegie-Rochester Conference Series on Public Policy*, 41, 157-219.

Koop, G., 2003, *Bayesian Econometrics* (Wiley, Chichester).

Koop, G., Leon-Gonzalez, R. and Strachan R.W., 2007, On the evolution of monetary policy, working paper available at http://personal.strath.ac.uk/gary.koop/koop_leongonzalez_strachan_kls5.pdf.

Koop, G., Potter, S. and Strachan, R,.W., 2007, Re-examining the consumption-wealth relationship: The role of model uncertainty, *Journal of Money, Credit and Banking*, forthcoming.

Lettau, M. and Ludvigson, S., 2004, Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption, *American Economic Review*, 94, 276-299.

Primiceri, G., 2005, Time varying structural vector autoregressions and monetary policy, *Review of Economic Studies*, 72, 821-852.

Sargent, T., Williams, N. and Zha, T., 2006, Shocks and government beliefs: The rise and fall of American inflation, *American Economic Review*, 96, 1193-1224.

Staiger, D., Stock, J. and Watson, M., 1997, The NAIRU, unemployment and monetary policy, *Journal of Economic Perspectives,* 11, 33-49.

Verdinelli, I. and Wasserman, L., 1995, Computing Bayes Factors using a generalization of the Savage-Dickey density ratio, *Journal of the American Statistical Association*, 90, 614-618.

Appendix A

In our empirical section, we use a special case of the state space model with $T_t = I$ and, thus, no additional details need to be provided about T_t. In terms of the parameters in the state space model, this leaves Q_t, H_t, a_1 and P_1.

We use a training sample prior that is the same as that used in Primiceri (2005) and Koop, Leon-Gonzalez and Strachan (2007) and is very similar to that used by Cogley and Sargent (2001, 2005), except that we average over different training samples. To be specific, we use a training sample prior with τ_0 quarters of data to choose many of the key prior hyperparameters. That is, we first run a time-invariant regression using the first τ_0 observations to produce OLS estimates of the regression coefficients, $\widehat{\beta}$, and the error variance, $\widehat{\sigma}^2$. We also obtain OLS estimates of the variance-covariance matrix of $\widehat{\beta}$ which we label $\widehat{V}_\beta$. We use these quantities as prior hyperparameters as described below. To make sure our inferences are robust to choice of training sample, we average our results over $\tau_0 = 10, 11, 12, \ldots, 30$.

For the initial conditions in our state equation, we use:

$$\alpha_1 \sim N\left(\widehat{\beta}, 4\widehat{V}_\beta\right),$$

which defines a_1 and P_1.

We set $Q_t = Q$ and use a conditionally conjugate Wishart prior for Q^{-1}. That is,

$$Q^{-1} \sim W\left(\underline{\nu}_Q, \underline{Q}^{-1}\right)$$

and the relevant posterior conditional used in the MCMC algorithm is:

$$Q^{-1}|Data \sim W\left(\overline{\nu}_Q, \overline{Q}^{-1}\right)$$

where

$$\overline{\nu}_Q = T + \underline{\nu}_Q$$

and

$$\overline{Q}^{-1} = \left[\underline{Q} + \sum_{t=1}^{T} \left(\alpha_{t+1} - \alpha_t\right)\left(\alpha_{t+1} - \alpha_t\right)'\right]^{-1}.$$

We set $\underline{\nu}_Q = 40$ and $\underline{Q} = 0.0001\widehat{V}_\alpha$.

Finally, a standard stochastic volatility model is used to specify H_t. In particular, if $h_t = \ln\left(H_t\right)$ is the log-volatility then we use:

$$h_{t+1} = h_t + e_t,$$

where e_t is $N\left(0, \sigma_h^2\right)$ and independent over t and of ε_t and η_t. Conditional on α and the other model parameters, we can use any standard algorithm for posterior simulation of the log-volatilities. In our empirical work, we use the algorithm of Kim, Shephard and Chib (1998). As was done by Primiceri (2005), for the initial condition, we take as prior:

$$\log\left(h_1\right) \sim N\left(\log\left(\widehat{\sigma}^2\right), I_3\right).$$

Finally, we use a Gamma prior for $\frac{1}{\sigma_h^2}$:

$$\frac{1}{\sigma_h^2} \sim G\left(\underline{h}, \underline{\nu}_h\right),$$

where $G\left(\underline{h}, \underline{\nu}_h\right)$ denotes the Gamma distribution with mean $\underline{h}$ and degrees of freedom $\underline{\nu}_h$. The posterior for $\frac{1}{\sigma_h^2}$ (conditional on the states) is also Gamma:

$$\frac{1}{\sigma_h^2}|Data \sim G\left(\overline{h}, \overline{\nu}_h\right)$$

where

$$\overline{\nu}_h = T + \underline{\nu}_h$$

and

$$\overline{h} = \frac{\overline{\nu}_h}{\underline{\nu}_h \underline{h}^{-1} + \sum_{t=1}^{T} \left(h_{t+1} - h_t\right)^2}.$$

We set $\underline{\nu}_h = 4$ and $\underline{h}^{-1} = 0.0001 I_3$.