On Measuring The Welfare Cost of Business Cycles[1]

Christopher Otrok

University of Virginia

First Version: August 1998

This Version: December 1999

Abstract: Lucas (1987) argues that the gain from eliminating aggregate fluctuations is trivial. Following Lucas, a number of researchers have altered assumptions on preferences and found that the gain from eliminating business cycles are potentially very large. However, in these exercises little discipline is placed on preference parameters. This paper estimates the welfare cost of business cycles, allowing for potential time-non-separabilities in preferences, where discipline is placed on the choice of preference parameters by requiring that the preferences be consistent with observed fluctuations in a model of business cycles. That is, a theoretical real business cycle world is constructed and the representative agent is then placed in this world. The agent responds optimally to exogenous shocks, given the frictions in the economy. The agent's preference parameters, along with other structural parameters, are estimated using a Bayesian procedure involving Markov Chain Monte Carlo methods. Two main results emerge from the paper. First, the form for the time-non-separability estimated in this paper is very different than the forms suggested and used elsewhere in the literature. Second, the welfare cost of business cycles is close to Lucas's estimate.

"Economic instability at the level we have experienced since the Second World War is a minor problem"
"It is worth re-emphasizing that these calculations rest on assumptions about preferences only, and not about any particular mechanism--equilibrium or disequilibrium--assumed to generate business cycles"

Robert E. Lucas Jr., 1987

"The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions"

Robert E. Lucas Jr., 1981

I   Introduction

Since 1930 much government policy has been aimed at reducing business cycle fluctuations, yet economists still are uncertain whether the magnitude of the potential gains from such a reduction are worth the effort. Lucas (1987) calculates that the welfare gain from eliminating *all* consumption fluctuations is less than one-hundreth of one percent of consumption when preferences are logarithmic. Clearly, in this estimate there is little room to increase welfare by designing policies to reduce or eliminate business cycles. In other calculations of the welfare cost of consumption volatility, most researchers have followed Lucas's (1987) approach. It is a simple procedure that only requires specifying a form for preferences and a time series representation for the fluctuating consumption process. Expected utility of the fluctuating consumption process is calculated and equated to the utility associated with a deterministic (non-fluctuating) consumption process. The welfare cost is the percent of consumption that the agent would pay, at every date and state, to move from the fluctuating economy to the non-fluctuating economy.

As noted by several authors, Lucas's result is sensitive to the assumption about the functional form for preferences and the exogenous process driving consumption. Lucas assumed standard constant relative risk aversion (CRRA) preferences and a trend-stationary iid consumption process. Other researchers have re-examined Lucas's welfare cost calculation under alternative assumptions for preferences and the consumption process. For example, Dolmas (1998), Obstfeld (1994), Pemberton (1995) and Tallarini (1998) estimate the potential welfare gains using non-expected utility and find gains ranging from one to twenty percent of consumption, depending on parameter choices. Van Wincoop (1994) finds that the gains from eliminating consumption fluctuations with habit formation preferences can be even larger, ranging from 1 to 25 percent. Campbell and Cochrane (1995) and Campbell (1998) argue that consumers with habit formation preferences consider consumption fluctuations as seriously as changes in growth, i.e., the cost is very large. Campbell argues more generally that any form of preferences that generates an equity premium as large as is observed in the U.S. implies a large welfare cost of fluctuations.[2]

---

[2] Atkeson and Phelan (1994) argue that it is not necessarily the case that a model that delivers large equity premia leads to a large gain from eliminating aggregate fluctuations. They construct an incomplete markets model in which

Should we believe any of these 'back-of-the-envelope' welfare cost estimates? One reason to be skeptical is that little discipline has been placed on the preferences. In Otrok (1999) I show that it is trivial to make the welfare cost of business cycles as large as one wants by simply choosing an appropriate form for preferences. Apparently reasonable settings of the parameters of time-non-separable preferences can produce welfare costs of U.S. consumption fluctuations as high as 40% of total consumption. Further, it is difficult to argue about such settings on a priori grounds since there is no straightforward economic interpretation of preference parameters in time-non-separable preferences.[3] What is needed is a method to discipline the choice of preference parameters. The approach taken in this paper is to use business cycle theory, in the spirit of Kydland and Prescott (1982) and King, Plosser and Rebelo (1988) [KPR], to determine an appropriate form for time-non-separable preferences. The discipline is that the preferences should be able to reproduce the observed consumption process, along with other aggregate variables, in a model of business cycles.

The approach in this paper to measuring welfare cost is closest to that of Tallarini (1998). Tallarini augments the standard KPR model with state-non-separable preferences, and calculates the welfare cost of business cycles (along with asset prices). His point is that he can raise the welfare cost of business cycles (and raise equity premium) by increasing risk aversion, holding constant intertemporal substitution, yet the implications for the second-moment properties of aggregate variables remain virtually unchanged from the predictions of the standard model. Here I take the discipline a step further in requiring that the model deliver fluctuations that mimic the type of persistent fluctuations we observe in the data, as the standard KPR model that Tallarini uses is known to miss many of these important features of the U.S. data. With non-separable preferences this discipline is of critical importance since welfare depends on not only the overall volatility of consumption, but also on the temporal nature of that volatility. Further, the choice of preference parameters with time-non-separable preferences dictates which types of volatility will be more costly (see Otrok 1999). The focus of this paper is *time*-non-separabilities in preferences, yet since welfare cost with *state*-non-separabilities preferences also depends on the temporal nature of consumption fluctuations, the approach developed here is applicable to the state-non-separable class of preferences as well.

As is well known from the work of Cogley and Nason (1995a) and Watson (1993) the standard business cycle model does not generate 'business cycle' fluctuations per se. The prototypical model is unable to generate what we observe in the data--volatility concentrated at business cycle frequencies.

---

eliminating aggregate fluctuations does not result in large welfare gains. However, in a complete markets, representative agent, preference-based asset pricing model it is the case that high equity premia and high welfare cost of fluctuations are tightly linked.

[3] In contrast, with time-separable preferences the only free parameter has the interpretation of measuring risk aversion, a parameter about which we can discuss what is reasonable and what is not.

However, an extensive literature has developed in recent years that adds realistic frictions to the standard model to explain the types of fluctuations we observe in the data. The approach here is to use this work to build a model of business cycles, and then ask what types of preferences would be consistent with aggregate fluctuations. Specifically, I construct a theoretical world using frictions that have been established elsewhere in the real business cycle literature to be useful in explaining features of the aggregate data. The representative agent is then placed in this world, and responds optimally to exogenous shocks, given the frictions in the economy. The agents' preference parameters, along with other structural parameters, are estimated using a full-information Bayesian procedure involving Markov Chain Monte Carlo methods. The resulting welfare cost associated with this form is calculated. Two main results emerge. First, the form for the time-non-separability estimated in this paper is very different than the forms suggested and used elsewhere. Second, the welfare cost of business cycles is close to Lucas's estimate.

It should be noted that this paper is not the first to embed time-non-separable preferences in consumption into a RBC model. Habit formation preferences have been used in business cycle models by Boldrin, Christiano and Fisher (1995), Lettau and Uhlig (1997) and Jermann (1998).[4] Lettau and Uhlig (1997) find that habit reduces consumption volatility to a counterfactually small fraction of output volatility. Boldrin, Christiano and Fisher (1995) and Jermann (1998) focus on the ability of the RBC model to generate realistic asset pricing implications; e.g., to match the equity premium in the data. However, the models used by these authors cannot match the persistent fluctuations in growth rates observed in the data. The approach in these papers is to impose a parameterization of preferences in the model, then add frictions and calibrate the remaining model parameters to match the overall volatility of the data for some aggregate variables. In contrast, the approach here is first to find a form for preferences that is consistent with aggregate fluctuations, and only then to ask substantive questions of the model.

II Time-Non-Separable Preferences and the Welfare Cost of Business Cycles

The economic interpretation of preferences that are time-non-separable [TNS] is that the utility that an agent derives from a consumption bundle today depends not only upon the characteristics of that bundle, but also upon the consumption bundles consumed in previous periods. That is, the ordering of consumption bundles matters for lifetime utility, not just the discounted sum of those bundles. The effect of consumption in one period on utility in future periods depends on whether consumption bundles are complements, substitutes or some combination, across time periods. For example, in one popular formulation, known as habit formation, consumption in previous periods reduces the utility of a

---

[4] In the RBC literature non-separable preferences in leisure were introduced by Kydland and Prescott (1982). The focus in this paper is on non-separabilities in consumption.

consumption bundle in the current period because the agent has grown accustomed to the level of consumption in the previous period. In this case consumption levels in different periods are complementary.

A more precise way to characterize TNS preferences is to note that an agent with such preferences cares not only about the overall amount of consumption volatility, but also about the temporal nature of that volatility. For example, an agent may have a stronger aversion to high frequency fluctuations then low frequency fluctuations?[5] We can use this intuition to dramatically increase the welfare cost of U.S. business cycles by observing that U.S. consumption volatility is mostly a low frequency phenomena, and then we can write down a utility function that has an heightened aversion to this type of low frequency volatility. For example, we can use Heaton's (1995)[6] form for time-non-separable preferences and choose parameters such that the agent dislikes low frequency volatility- and we would find that the welfare cost of U.S. consumption fluctuations associated with these preferences is 40 percent![7] Is this estimate believable? No, clearly the calculation was rigged to deliver such a welfare cost. But this 'rigged' calculation is in the same spirit as both the welfare cost and equity premium calculations presently made in the literature. The starting point was specifying the consumption process to characterize the U.S. consumption process. Next, preferences were specified which, by design, had an aversion to the observed consumption process. An alternative approach to measuring welfare cost is pursued in the next section, where I build a model that links preferences to the observed consumption process.

III   Building a Model of Business Cycles

A key feature of welfare calculations with TNS preferences is the temporal nature of the consumption process; therefore it is important that the model exhibits the persistent fluctuations we observe in the data. Watson (1993) shows that the standard business cycle model is not able to generate fluctuations with hump shaped spectra for aggregate growth rates that we observe in the U.S. economy. Cogley and Nason (1995a), Rotemberg and Woodford (1996), and Wen (1996) show that the standard model lacks an internal mechanism to propagate shocks, which leads to trivial dynamics. There has been, however, a growing body of research aimed at improving the model's ability to match the data. The model presented here combines a number of these features in one place, in order to build a model that will be capable of matching the data on a number of dimensions. The model has two sectors, each with its own capital stock that cannot be transferred to the other sector. Additional features are time-to-build capital, an aggregate dynamic complementarity and variable rates of capacity utilization. To this model I add time-

---

[5] In Otrok (1999) I develop spectral utility and spectral welfare cost functions to illustrate this preference formally.
[6] Details of these preferences will be given in section III.
[7] The welfare cost was estimated using the U.S. consumption process, following Otrok (1999). Details are omitted here since the example is only illustrative.

non-separable preferences. I begin by explaining the mechanics of each feature, along with a brief summary of why the feature was originally proposed. The reader is referred to the original work for a more thorough explanation and justification of these features.

Preferences

Many ways of modeling non-separabilities in preferences have been suggested.[8] The interest here is in a parametric form for the time-non-separability that is both flexible in the sense that it encompasses many forms for the non-separability, yet parsimonious in that the number of parameters to be estimated is relatively small. Such a form has been suggested by Heaton (1995), and will be used in this paper for preferences over consumption. Preferences over leisure are standard CRRA preferences and enter separably in the utility function.[9] The preferences are defined as:

$$(1) \qquad U(s) = E \left\{ \sum_{t=0}^{\infty} \beta^t \frac{s_t^{1-\gamma}}{1-\gamma} \right\} + \varphi \frac{(l_t)^{1-\eta}}{1-\eta}, \quad \gamma, \eta > 0$$

where l_t denotes leisure and s_t is defined by:

$$s_t = s_t^d - \alpha(1-\theta)s_t^h, \qquad 0 < \alpha < 1,$$

where s_t^d is the 'durability stock' and governs the amount of substitutability in consumption across time and s_t^h is the 'habit stock' and governs the amount of consumption complementarity across time.[10] These processes evolve according to:

$$s_t^d = \sum_{j=0}^{\infty} \delta^j c_{t-j}, \qquad 0 < \delta < 1,$$

$$s_t^h = \sum_{j=0}^{\infty} \theta^j s_{t-1-j}^d, \qquad 0 < \theta < 1.$$

With δ and α equal to 0 we have CRRA time-separable utility. With $\delta = 0$ and $\theta = 0$ we have the popular one-lag formulation of 'habit preferences'. With $\alpha = 0$ preferences would be a form known as 'durability', used by Eichenbaum, Hansen and Singleton (1988). The powerful feature of these preferences, from the

---

[8] Different functional forms have been suggested by Constantinides (1990), Eichenbaum Hansen and Singleton (1988), Heaton (1995), Abel (1990, 1999), Campbell and Cochrane (1997), Boldrin, Christiano and Fisher (1997), Abel (1998) and Carroll, Overland and Weil (1997).

[9] Heaton's (1995) study did not use leisure.

[10] The definition of s_t is a restriction on the parameters in the more general function $s_t = \sum_{j=0}^{\infty} a_j c_{t-j} \equiv A(L)C_t$, where L is the lag operator. Heaton's restriction decreases the number of parameters that need to be estimated, while still allowing for the possibility of consumption complementarity or substitutability at long horizons.

viewpoint of this paper, is that with appropriately chosen parameter values, the preferences can deliver an heightened aversion to either low or high frequency volatility.

The total number of available hours is normalized to 1, so the agent faces the time constraint: $n_t + l_t = 1$, where n_t is hours worked and l_t is leisure hours. Also, in the empirical work below, the parameter φ is chosen so that steady-state labor hours represent 30 percent of total time.

Production

Two goods are produced in the economy, an investment good and a consumption good; each is produced in a separate sector. The consumption good is produced by many identical firms with technology given by

$$(2) \qquad c_t = A_t (\phi_{c,t} k_{c,t})^{1-\nu} n_{c,t}^{\nu} Y_{t-1}^{\mu c}$$

where A_t is an aggregate technology shock, $\phi_{c,t}$ is the rate of capacity utilization, $k_{c,t}$ is the (consumption) capital stock, $n_{c,t}$ is labor in the consumption sector, and Y_{t-1} is the last period's aggregate production of both the consumption and investment goods. Lower-case variables denote firm-level variables, and upper case variables denotes aggregate variables. Each firm chooses a capital stock, rate of utilization, amount of labor, and treats the shock and aggregate output as exogenous.

Investment goods are produced in a separate sector, with a similar technology:

$$(3) \qquad i_t = A_{i,t} A_t (\phi_{i,t} k_{i,t})^{1-\nu} n_{i,t}^{\nu} Y_{t-1}^{\mu i}$$

where $A_{i,t}$ is a technology shock that is specific to the investment sector.

The production of the consumption and investment goods in separate sectors follows the recent work on multi-sector models by Huffman and Wynne (1998), DeJong, Ingram, Wen and Whiteman (1998), Benhabib, Perli and Sakellaris (1997) and others. These researchers have established that this type of production technology contributes to the model's ability to match the data. The two-sector model is particularly important in modeling consumption fluctuations: With concave preferences there is an incentive for the agent to smooth consumption, and with the one-sector economy, ample means to do so. A two-sector economy adds a realistic friction that frustrates the agent's ability to smooth consumption, and potentially results in more realistic predictions for consumption volatility. A friction on the movement of the capital stock is required in the model to prevent smoothing across sectors. In fact, without some type of friction, the two-sector model aggregates to a one-sector model (see Huffman and Wynne 1998). Here I require that the capital stocks be sector specific with no movement of capital across

sectors. The idea of sector-specific irreversible capital is appealing, since empirically it appears that it is difficult to transform capital between sectors. [11]

The second important feature of the production technology is the aggregate dynamic complementarity. Cooper and Johri (1996) study the contribution of contemporaneous and dynamic complementarities to the propagation mechanism of the one sector RBC model. They find that dynamic complementarities significantly enhance the propagation of technology shocks in RBC models. The complementarity can be thought of as capturing learning-by-doing spillover effects, where the spillovers across sectors occur with a lag. Cooper and Haltiwanger (1996) provide an empirical study of aggregate dynamic complementarities. They examine microeconomic evidence of sectoral output along with the implications for aggregate data and find evidence in favor of dynamic complementarities.[12]

The third feature of the production technology is the variable rate of capacity utilization. This feature was introduced into RBC models by Greenwood, Hurcowitz and Huffman (1988). Variable rates of utilization have been added to help the standard RBC model to match the cyclical properties of US business cycles by Burnside and Eichenbaum (1995), DeJong, Ingram, Wen and Whiteman (1998), and Wen (1998a,b). Additionally, as argued by King and Rebelo (1998), variables rates of capacity utilization overcome one of the most prominent objections to the RBC paradigm: the reliance on technology shocks. With variable utilization the standard deviation of the shocks is much smaller (variable rates of utilization amplify shocks), and the model predicts little 'technical regress'.

Time-To-Build Capital

The importance of time-to-build capital in generating aggregate fluctuations dates back to the work of Kydland and Prescott (1982). Capital that takes multiple periods to build creates a persistent demand for investment, once a new investment project has been started. This demand can potentially drive persistent fluctuations in output. However as shown by Rouwenhorst (1991), the version of time-to-build used by Kydland and Prescott cannot match the cyclical behavior of US investment, nor does it have a significant impact on output persistence. Wen (1998) remedies this deficiency by introducing a version of time-to-build where investments in adjacent periods are not perfect complements. The form is Cobb-Douglas as follows:

---

[11] Huffman and Wynne (1998) develop a more complete and rigorous argument for sector-specific capital. In their work they introduce a cost of adjusting capital.

[12] A shortcoming of the standard two-sector model is that output of the two sectors is predicted to be negatively correlated, while in U.S. data output across sectors moves together. DeJong and Ingram (1998) and Benhabib, Perli and Sakellaris (1997) remedy the problem with a third sector, interpreted to be human capital or home production, which moves countercyclically. Huffman and Wynne (1998) remedy the problem with a cost of adjusting capital. Here I follow the suggestion of Cooper and Haltiwanger (1996) that an aggregate dynamic complementarity can explain the observed sectoral output comovement.

(4) $m_{c,t} = i^r_{c2,t} i^{1-r}_{c1,t-1}$, and

(5) $m_{i,t} = i^r_{i2,t} i^{1-r}_{i1,t-1}$.

Where $m_{c,t}$ and $m_{i,t}$ is new capital put in place in period t in each sector. To add new productive capital in period t+1, an investment project is started in period t-1 ($i_{c1,t-1}$), and additional investment ($i_{c2,t}$) is added in period t. The final investment good ($m_{c,t}$) is added to the capital stock in period t + 1. In conjunction with variable rates of capacity utilization and an employment externality, which provide an elastic supply side to the economy, Wen shows that this model for investment is capable of producing investment cycles with the same second moment properties, or spectrum, of those that we observe in the U.S.

Capital Stock law of Motion

The law of motion for the capital stock in the two sectors is:

(6) $k_{c,t+1} = (1 - \dfrac{\delta_0}{\omega} \phi^{\omega}_{c,t}) k_{c,t} + m_{c,t}$,

(7) $k_{i,t+1} = (1 - \dfrac{\delta_0}{\omega} \phi^{\omega}_{i,t}) k_{i,t} + m_{i,t}$.

The rate of depreciation depends on the level of utilization. The functional form, introduced by Greenwood, Hercowitz and Huffman (1988), captures the cost of increasing the rate of capacity utilization: the rate of depreciation also increases.

Technology Shocks

The log of the two shocks follow an AR(1) processes, as is standard in the RBC literature:

(8) $\ln(A_t) = \rho_A \ln(A_{t-1}) + \varepsilon^A_t$,

(9) $\ln(A_{i,t}) = \rho_I \ln(A_{i,t-1}) + \varepsilon^I_t$.

$$\begin{bmatrix} \varepsilon^A_t \\ \varepsilon^I_t \end{bmatrix} \sim N(0,\Sigma); \quad \text{where } \Sigma = \begin{bmatrix} \sigma^2_A & \sigma_{AI} \\ \sigma_{AI} & \sigma^2_I \end{bmatrix}$$

Model Solution

The competitive equilibrium of the model requires that the agent maximize utility and the firm maximize profits. Equilibrium allocations are found by solving the agent's optimization problem. The agent's problem is to maximize (1) subject to (2)-(9). The agent treats the shocks and the aggregate complementarity as exogenous (i.e. derivatives are not taken with respect to these variables). An approximate solution to the model is found by log-linearizing the agents' first order conditions around the steady state of the system. The system of linear expectational difference equations is then solved using the

methods described in Sims (1996) and Klein (1998). The approach in both papers is based on a QZ matrix decomposition. One advantage of this approach is that the solution algorithm is very fast, which will be important in the estimation stage since the model needs to be solved many times. The QZ approach is also useful since it is straightforward to handle technical problems related to singularities that naturally arise in multi-sector models. The presence of an externality poses no special difficulty to the solution algorithm provided that the solution is unique, i.e., that there are no indeterminacies in the model. While the presence of multiple sectors can potentially decrease the size of the externality parameter that will lead to indeterminacies (see Benhabib and Farmer 1996), the structure of this model is such that for no reasonable parameter setting will the solution contain multiple equilibria. In terms of generating multiple equilibria, the important difference between the model here and that of Benhabib and Farmer is that here the capital stock cannot be transferred across sectors, while in Benhabib and Farmer the capital stock can be transferred across sectors within the current period. It turns out that when the capital stock cannot be transferred across sectors (or there is some cost to doing so), a very large externaility parameter is needed to generate multiple equilibria.[13]

The solution to the model is written in terms of deviations from steady state values. The linearization procedure results in a first-order difference equation for the state variables:

(10) $\quad \xi_{t+1} = F\xi_t + C\varepsilon_t,$

where

$$\xi_t = \begin{bmatrix} \hat{k}_{c,t} & \hat{k}_{i,t} & \hat{i}_{c1,t-1} & \hat{i}_{i1,t-1} & \hat{Y}_{t-1} & \hat{s}_t^d & \hat{s}_t^h & \hat{A}_t & \hat{A}_{i,t} \end{bmatrix}',$$

and ^ variables indicate percent deviation from steady state. The decision rules for the choice variables are linear functions of the states:

(11) $\quad z_t = H\xi_t,$

where

$$z_t = \begin{bmatrix} \hat{c}_{c,t} & \hat{n}_{c,t} & \hat{n}_{i,t} & \hat{\phi}_{c,t} & \hat{\phi}_{i,t} & \hat{i}_{c2,t} & \hat{i}_{c1,t} & \hat{i}_{i2,t} & \hat{i}_{i1,t} \end{bmatrix}'.$$

This vector can be augmented with other variables that are combinations of the choice and state variables. For example, we could add total investment, which is the sum of the individual investment series.

IV   Model Estimation

The parameters of the model are estimated using a Bayesian procedure. There are several practical advantages to the Bayesian approach to estimation. First, there are parameter restrictions that are

---

[13] I experimented with the externality parameters to see how large they would have to be to generate multiple equilibria, and found that they would have to be significantly larger than 1, which was very far away from the

difficult to enforce using Maximum Likelihood (ML) or the Generalized Method of Moments (GMM). Second, the model above has a number of boundaries, and it is likely that some of the estimated parameters will be on or near a boundary. In this case, the asymptotic distribution theory underlying GMM and ML estimators is of questionable value. Furthermore, Christiano and den Haan (1995) have shown that in a sample the size of U.S. post war quarterly data, GMM parameter estimates of dynamic macroeconomic models often do not have the distributions implied by asymptotic theory. Model parameter restrictions, parameter estimates near boundaries, and sample size issues pose no special difficulties in the Bayesian framework.

An additional reason for employing a Bayesian estimation procedure is that economists have prior beliefs, often strong ones, about parameter values of structural macroeconomic models that should enter into the estimation. For example, we would want to rule out certain parameter values as unreasonable (e.g. a labor share parameter of 0.1 in the production function). A Bayesian procedure allows these beliefs to be incorporated formally.

An alternative approach to formal estimation is to calibrate the model. Given the apparent complexity of the estimation procedure that follows, it might seem to be easier to perform a calibration exercise. However, formal estimation and the presence of the likelihood function place an important discipline on the choice of parameters by 'choosing' the parameters using a well known, explicit set of principles. A calibration exercise would be open to the criticism that the results were contrived, and a different set a parameters should have been used.

The estimation procedure is similar to the approach developed in DeJong, Ingram and Whiteman ([DIW], 1998a,b). Begin by denoting the parameters of interest as:

$$\Gamma = [\ \gamma \quad \delta \quad \theta \quad \alpha \quad \beta \quad \eta \quad \mu_c \quad \mu_i \quad \nu \quad r \quad \delta_0 \quad \omega \quad \rho_A \quad \rho_I \quad \text{vec}(\Sigma)'\ ]',$$

and the sample of data as X_t. The posterior density of the parameters, $P(\Gamma|X_t)$, can be written as the product of the likelihood function of the data given the parameters, $L(X_t|\Gamma)$, and the prior, $\pi(\Gamma)$:

(12) $P(\Gamma|X_t) \propto L(X_t|\Gamma)\pi(\Gamma)$

The simple structure of the linear observer system (10)-(11) associated with the linearization of the model's equilibrium law of motion makes it possible to use the Kalman Filter to calculate the value of the likelihood function given the parameters. We are interested in posterior distributions of the parameters and functions of the parameters, such as model second moments. Thus, the fact that (12) is the product of the likelihood and prior complicates calculations. If we had a large sample of parameters drawn directly from the posterior in (12) it would be straightforward to find, say, the median of the posterior distribution

---

parameter values I ultimately estimated. In the estimation procedure that follows, if multiple equilibria would have arisen the computer program would have stopped.

of a parameter by simply calculating the median value of the parameters that we drew from (12). Similarly, we could calculate the median of the posterior distribution of a function of the parameters (for example, output growth volatility) by evaluating the function for each set of parameters we drew, and then calculating the median of these values.

Unfortunately we cannot directly simulate from the posterior in (12). However, we can calculate the value of the posterior given a set of parameters, and as in the work of DIW this fact will prove important in designing an algorithm to simulate from (12). DIW use importance sampling (see Geweke, 1989) to simulate draws from the posterior. Here Markov Chain Monte Carlo (MCMC) methods are used to sample from the posterior (for surveys of MCMC methods see Chib and Greenberg 1996, Geweke 1997, or Tierney 1994). The process for simulating from a distribution using MCMC methods begins with an initial arbitrary value for the parameter vector, Γ^1 (from within the support of the distribution). The height of the posterior at this value, $P(\Gamma^1|X_t)$, is then calculated. Next, a drawing for the parameter vector, $\Gamma*$, is taken from a candidate-generating density, $q(\Gamma*|\Gamma^1)$. (Specifically, this density generates a candidate vector $\Gamma*$ for the next step in the Markov Chain given the current draw Γ^1.) Next, the height of the posterior is calculated using the parameter vector $\Gamma*$, $P(\Gamma*|X_t)$. Then Γ^2 takes the value $\Gamma*$ with probability:

$$\alpha(\Gamma^1,\Gamma*) = \min\left[\frac{P(\Gamma*|X_t)}{P(\Gamma^1|X_t)}, 1\right].$$

If Γ^2 does not take the value $\Gamma*$, it is set equal to Γ^1. The process is repeated many times and given regularity conditions, the set of draws converge to a set of draws from the posterior. To prevent the arbitrary starting point from affecting the results, the first 5000 draws are discarded.

An important ingredient in the MCMC algorithm is the candidate-generating density, $q(\Gamma*|\Gamma^1)$. There are two important issues when specifying this density. First, the density should be easy to simulate from (e.g. a normal distribution). Second, the density should be symmetric: that is $q(\Gamma^2|\Gamma^1) = q(\Gamma^1|\Gamma^2)$. Chib and Greenberg (1996) show that if $q(\Gamma*|\Gamma^1)$ is symmetric then the Markov Chain will be reversible, which can be verified by checking that the following condition holds: $P(\Gamma^1|X_t)q(\Gamma^1|\Gamma^2)\alpha(\Gamma^1,\Gamma^2) = P(\Gamma^2|X_t)q(\Gamma^2|\Gamma^1)\alpha(\Gamma^2,\Gamma^1)$. If the Markov chain is reversible then the invariant distribution of the chain is $P(\Gamma|X_t)$. That is, with a sufficiently long chain we will obtain a set of draws from the posterior of interest. Here the candidate-generating density is multivariate normal, with mean equal to the parameters from the last draw, and a diagonal covariance matrix with diagonal elements that are small relative to the dispersion in the support of the parameter distribution. This is a random-walk chain, which is particularly useful when we have little information on the shape of the distribution we are simulating from, as is the

case here.[14] Note that the multivariate normal distribution is both easy to simulate from and satisfies the symmetry requirement.

The MCMC algorithm relies on the ability to evaluate (12) for a given set of parameters. However, since the model is stochastically singular there are nontrivial implications for only two observable time series. When evaluating the likelihood, instead of the observation equation in (11) we use an alternative equation, which is the decision rule for these two observables, here consumption and total investment:

(11') $\Lambda_t = H\xi_t,$

where $\Lambda_t = \begin{bmatrix} C_{c,t} & I_t \end{bmatrix}'$, and I_t is total investment. Obtaining the value of the posterior simply requires multiplying the value of the likelihood (obtained from the Kalman filter) by the prior (which is evaluated at the current draw of the parameter vector). The data used in the estimation are quarterly observations from 1948:1 to 1997:4 on real nondurable and services consumption and non-residential private investment. Both time series were Hodrick-Prescott (HP) filtered before estimation with a smoothing parameter of 1600. The HP filtered data is compared directly with the output of the model. The model output is not HP filtered following the arguments presented in Cogley and Nason (1995b).

The model has predictions for many other time series, each of which could be used to estimate the model in place of the consumption and investment data. It seems natural to use the consumption data since I am measuring the welfare cost of consumption volatility. Furthermore, all other studies of the welfare cost use consumption data, and for comparability consumption should be used here. The case for using investment data is not as strong as the case for consumption data. However, investment is a standard variable of interest to business cycle researchers so it is a natural selection. Moreover, I have estimated the model with other time series in place of the consumption and investment data to illustrate the insensitivity of the results to these changes; results of these robustness checks are reported below.

When using MCMC methods it is necessary to check for convergence. Two issues are relevant. First, the length of the chain matters. A chain must be sufficiently long to ensure convergence. In practice, convergence is assessed by running the chain multiple times, and checking to see if the results are the same, or are affected by chain length. Here, a chain length of 100,000 was used, and repeated runs resulted in the same parameter estimates. The second issue is the sensitivity to starting conditions. If the posterior is multi-modal and approaching disjointedness, then different starting points will lead to different results. This problem is checked by starting the chain from different points from within the

---

[14] If we had a reasonable idea of the shape of the posterior then we might be able to use that information to specify a candidate-generating distribution that was 'similar' to the posterior. The additional information could result in a chain that converged more rapidly (e.g. require fewer simulations). This type of procedure would be an independence Markov Chain, which is similar to the importance-sampling methods used by DIW.

support of the posterior and verifying that the results are the same.[15] Computationally, one run of a Markov Chain of length of 100,000 took approximately 13 hours on a 300 Mhz pentium-pro desktop computer.

Priors

Prior beliefs over the parameters of the model are imposed on the estimation. These priors are specified as independent Normal variables. The priors are normally distributed with means, where possible, set equal to values that these parameters take elsewhere in the RBC literature. The prior standard deviation is set to reflect the uncertainty over this value. For example, we have much more prior information about the labor share parameter in the Cobb-Douglas production function than we do for the parameters of the utility function.

The priors for the preference non-separability parameters, α, θ, and δ were $N(0.5, 0.4^2)$, truncated at 0 and 1. The large standard deviation reflects the lack of prior information about these parameter values. The prior is disperse enough that it has negligible impact on the parameter estimates. The priors for the power parameters in the utility function, γ and η, are $N(1.0, .15^2)$, truncated at 0. The mean of 1 reflects the common assumption of logarithmic utility in the RBC literature. The discount factor prior was $N(0.99, 0.02^2)$, truncated at 0 and 1. The mean of 0.99 implies a mean riskless interest rate of 4%, a common assumption in the RBC literature. On the production side, the prior for the labor share parameter is $N(0.58, .05^2)$, again truncated at 0 and 1. The mean reflects the parameter value often used in calibration exercises in the RBC literature. The dynamic complementarity parameters, μ_c and μ_i, have a $N(0.075, 0.15^2)$ prior, truncated at 0 and 0.5. The dispersion reflects the considerable uncertainty about this parameter while the mean reflects the belief that while there are complimentarities, they do not need to be very large to provide a significant propagation mechanism to the model. The upper bound of this truncation reflects a belief that the model has a determinate solution. With dynamic complementarities larger than 0.5, indeterminancies potentially arise. The prior for the investment share parameter in the time-to-build function was $N(0.5, 0.3^2)$, truncated at 0 and 1. This prior is disperse enough to have a negligible impact on the results, reflecting the minimal prior information we have on this parameter. The mean was simply set at the middle of the support for parameter. The prior for the autoregressive parameter of both of the technology shocks is $N(0.8, 0.3^2)$, truncated at -1 and 1. The mean is less than the commonly used parameter in the RBC literature; this reflects the argument in Cooper and Haltiwanger that dynamic complementarities decrease the reliance of RBC models on the serial correlation in shocks

---

[15] The analogous problem that arises with classical maximum likelihood estimation is determining whether the max that is obtained is a global or local max. The difference here is that we want to *sample* from all areas of the posterior since we are interested in the shape of the posterior.

to generate persistent fluctuations. Again, there is considerable uncertainty over this parameter, which is reflected in the variance of the prior. The prior for the innovation variance of the two shocks was N(0.006, 0.02^2), truncated at 0. Again, this prior is diffuse enough that it has negligible effect on the parameter estimates. The prior for the covariance of the shock innovations, N(0, 0.4^2), is also diffuse enough to have no impact on the results. Finally, the priors for the depreciation parameters, δ_0 and ω, were N(0.05, 0.02^2), truncated at 0, and N(1.5, $.25^2$), truncated at 1. These prior means, in conjunction with the prior means of other model parameters, would result in a steady state level of capacity utilization of 80 percent, as we observe in the U.S. That is, although there is no explicit prior over the steady state level of capacity utilization, the implicit prior over the steady state level of capacity utilization is consistent with US data.

A Loss Function

The output of the estimation procedure is a posterior distribution for the parameter values. In order to investigate the welfare cost of business cycles it is helpful to focus on a single set of parameters. Some common choices for selecting a set of parameters from the posterior distribution are the median, mean, and the mode. These would correspond to absolute, quadratic and zero-one loss functions. Here I choose the parameter vector using a loss function that penalizes parameters that result in a poor spectral fit for the model. As argued in section II, the spectral nature of the consumption process can have a large impact on welfare cost calculations with TNS preferences. It is appropriate then, to place additional weight in choosing parameter values that deliver the type of spectra we observe in the data.

In order to place a weight on the spectral fit of the model in the loss function it is necessary to have a measure of how close the model and data spectra are to each other. Watson (1993) has proposed such a metric for evaluating the fit of a dynamic model. Watson calculates the minimal amount of stochastic error that needs to be added to model variables so that the autocovariances of model variables are equal to the autocovariances of the data. The error that needs to be added is a stochastic process, and hence, using spectral methods, we can calculate the spectrum of the error and compare this to the spectrum of the data and the model. The ratio of the error spectrum to the data spectrum at each frequency can be interpreted as a 1-R^2 measure in a standard regression model. To get a summary statistic to evaluate the fit at all frequencies we could integrate the error spectrum and divide this by the integral of the data spectrum. This summary statistic has an interpretation as a 1-R^2 measure for the fit of the model at all frequencies.

Here I incorporate Watson's 1-R^2 measure of fit into a loss function. The loss function is:

(13) $l(\Gamma, \Gamma^*) = (\Gamma - \Gamma^*)'(I_{17} \times W(\Gamma))(\Gamma - \Gamma^*).$

where I_{17} is an identity matrix of dimension 17, and $W(\Gamma)$ is the (scalar) product of four Normal density functions evaluated at the $1\text{-}R^2$ measure of fit, for output, consumption, investment and labor growth. A $N(0, 0.03^2)$ density is used for the fit of output, consumption and investment growth.[16] A more disperse distribution, $N(0, 0.13^2)$, is used for the labor growth fit since we know a priori that this model will not have the ability to match labor volatility very well.[17, 18]

To minimize expected loss find:

(14) $\hat{\Gamma}^* = \arg\min_{\Gamma^*} \int l(\Gamma, \Gamma^*) L(X_t | \Gamma) \pi(\Gamma) d\Gamma$

Substituting in the loss function (13) and taking derivatives we have:

(15) $\hat{\Gamma}^* = \int \Gamma W(\Gamma) L(X_t | \Gamma) \pi(\Gamma) d\Gamma$,

where $\hat{\Gamma}^*$ is the set of parameters that minimizes expected loss. The expectation in (15) is calculated using the MCMC algorithm described above. The mean of the distribution from this simulation minimizes the loss function.[19] In the empirical results that follow, I report moments of the posterior distribution for the parameters and the implied variances and covariances of the model. The minimum expected-loss parameter vector will also be reported.

V   Parameter Estimates and Properties of The Model

Table 1 gives the medians and 5 percent and 95 percent quantiles of the posterior distribution of the parameters, along with the parameters that minimize the loss function in (15). The effect of the loss function on the choice of parameters varies by parameter. For the leisure utility curvature parameter the value that minimizes the loss function is almost the median of the posterior, while for the Cobb-Douglas labor share parameter the value that minimizes expected loss is outside of the posterior 90% coverage interval.

As can been seen from this table the parameter estimates appear to be relatively tight for the production parameters but with more uncertainty over the preference parameters. The estimate of the labor share parameter (ν) is somewhat higher than the 0.58 prior mean (and the value commonly used in the RBC literature following the assumptions of KPR). For comparison, Christiano and Eichenbaum

---

[16] Data on output is the sum of the consumption and investment data used in the estimation.

[17] The average weekly labor hours time series used here was not collected by the Bureau of Labor Statistics after 1995, therefore the data spectra are estimated using 1948:1-1995:1 data.

[18] That is $W(\Gamma) = \left(\prod_{j=o,c,i} \frac{1}{\sqrt{2\pi} \times 0.03} \left[-\frac{1}{2 \times 0.03^2} [(1-R_j^2) - 0]^2 \right] \right) \frac{1}{\sqrt{2\pi} \times 0.13} \left[-\frac{1}{2 \times 0.13^2} [(1-R_l^2) - 0]^2 \right]$, where

$(1-R^2)$ is the Watson measure of fit and o, c, i, and l stand for output, consumption, investment and labor growth.

(1992), who used Generalized Methods of Moments, estimate this parameter to be 0.64, McGrattan (1994) uses Maximum Likelihood and finds a value of 0.60, while DIW (1998a) estimate the parameter to be 0.71. Of interest, but perhaps not surprising, the serial correlation in the shocks is much less than generally assumed. It is standard in the literature to assume values for the ρ parameters in the range of 0.95-1.00; here they are much lower, implying that the model has a reasonably strong internal mechanism to propagate shocks. The estimates here are similar and somewhat smaller than the 0.80 posterior means reported in DeJong and Ingram (1998). The variance of the shocks is significantly smaller than models that do not have variable rates of utilization, but on the same order of magnitude of models that do. The preference parameter estimates will be contrasted with Heaton's estimates after some predictions of the model have been established.

Table 1: Quantiles of Parameter Posterior Distributions

Parameter	5%	Median	95%	Loss Fnc	Parameter	5%	Median	95%	Loss Fnc
n_o	0.1316	0.2359	0.3526	0.2825	$var(e_I)$	3.9E-05	5.1E-05	6.7E-05	3.9E-05
m_I	0.0052	0.0395	0.1018	0.0266	$cov(e_A, e_I)$	-4.6E-05	-2.4E-05	-8.1E-06	-7E-06
n	0.5532	0.5916	0.6252	0.6749	b	0.9878	0.9882	0.9887	0.9876
r	0.6189	0.6581	0.6889	0.6877	q	0.0178	0.1299	0.3039	0.1146
d_0	0.0440	0.0478	0.0519	0.0501	a	0.2618	0.4460	0.6113	0.2145
w	1.5418	1.5513	1.5626	1.5651	d	0.0223	0.1533	0.3428	0.0844
r_I	0.4859	0.5512	0.6109	0.5915	g	0.5363	0.7228	0.9471	0.5503
r_A	0.5841	0.6415	0.7051	0.7413	h	0.8895	1.1206	1.3280	1.1223
$var(e_A)$	3.69E-05	5.886E-05	8.76E-05	2.00E-05					

The second moments of the model variables, in both levels and growth rates, along with analogous data moments are given in Tables 2 and 3. The model is estimated using HP filtered log consumption and log investment data so Table 2 naturally compares the moments of the log-levels from the model with the moments of HP-filtered (logged) data. Table 3 compares predictions for growth rates-the first differences of the log-levels. Since the model log-levels correspond to HP filtered data the growth rates in Table 3 use the first difference of HP filtered logged U.S. data. The tables contain 5%, median and 95% quantiles of the posterior distributions for these moments, along with the moments associated with the parameters chosen with the loss function. The second moments seem to be consistent with the

---

[19] An alternative interpretation of the parameter vector that results from evaluating (15) is that it is the mean of a posterior distribution with a data-dependent prior on the spectra of growth rates, i.e. the posterior distribution is now $L(X_t \mid \Gamma)\pi(\Gamma)W(\Gamma)$.

data with the exception of labor volatility, which is too small relative to the data. This failure is common to business cycle models.[20]

Table 2: Second Moments of Data and Model (Levels)

	Data		Model (quantiles of Posterior Distribution)							
	Std Dev	R. Std Dev	Std Dev				R. Std Dev			
			5%	Median	95%	Loss Fnc	5%	Median	95%	Loss Fnc
Output	0.014	1.000	0.015	0.018	0.021	0.018	1.000	1.000	1.000	1.000
Consumption	0.009	0.659	0.009	0.010	0.012	0.013	0.523	0.570	0.621	0.686
Investment	0.050	3.562	0.043	0.049	0.057	0.050	2.520	2.745	2.987	2.742
Labor	0.015	1.070	0.007	0.009	0.010	0.008	0.435	0.485	0.548	0.441
Cons. Labor	(NA)	(NA)	0.003	0.005	0.006	0.003	0.194	0.258	0.351	0.139
Inv Labor	(NA)	(NA)	0.036	0.041	0.048	0.042	2.071	2.309	2.582	2.314

Table 3: Second Moments of Data and Model (Growth)

	Data		Model (quantiles of Posterior Distribution)							
	Std Dev	R. Std Dev	Std Dev				R. Std Dev			
			5%	Median	95%	Loss Fnc	5%	Median	95%	Loss Fnc
Output	0.007	1.000	0.008	0.009	0.010	0.007	1.000	1.000	1.000	1.000
Consumption	0.005	0.759	0.005	0.006	0.006	0.005	0.564	0.622	0.682	0.717
Investment	0.027	3.800	0.025	0.027	0.030	0.026	2.749	2.948	3.181	3.422
Labor	0.009	1.241	0.005	0.006	0.008	0.004	0.571	0.664	0.793	0.579
Cons. Labor	(NA)	(NA)	0.002	0.003	0.005	0.001	0.229	0.369	0.554	0.169
Inv Labor	(NA)	(NA)	0.023	0.025	0.027	0.024	2.449	2.696	2.975	3.218

The correlations in the tables in the Appendix indicate that the comovement of variables is reasonably close to U.S. data. One feature of multi-sector models is that often consumption and investment move (counterfactually) in opposite directions (see Huffman and Wynne 1998). Here the two sectors move together due to the aggregate complementarity (see Cooper and Haltiwanger 1996).[21] Overall, the correlations compare favorably with the correlations presented for other RBC models. Here, most of the data correlations lie in or close to the posterior distribution implied by the model.

---

[20] A literature has evolved to increase the volatility of labor hours in the standard model, Hansen (1985) and Rogerson (1986) suggest indivisible labor, DeJong and Ingram (1998) add a human capital sector, Benhabib, Rogerson and Wright (1991) add home production and Wen (1998) uses time-non-separabilities in leisure preferences in conjunction with an employment externality, all of which result in a better fit for labor volatility.

[21] Huffman and Wynne (1998) rectify this problem with adjustment costs for moving capital across sectors and adjusting the capital stock. The sector-specific capital stock here can be viewed as an infinite adjustment cost. The comovement of consumption and investment then, is probably due to both this feature and the aggregate dynamic complementarity. Another way to get the consumption and investment sectors to move together is too add a third sector that is countercyclical as in Benhabib, Perli and Sakellaris (1998) and DeJong and Ingram (1998).

The spectra of the model and data growth rates, along with the stochastic error needed to rectify the two are shown in Figure 1. These spectra were calculated using the parameters that minimized the loss function. Again, with the exception of labor hours, the model seems to be generally consistent with this aspect of the data. The error that needs to be added to the output spectrum is 9% of the total data spectrum. For consumption the error is 14%, and for investment the error is 4%. The error in the labor sector is 70%. This is a vast improvement over the standard model, which has error on the order of magnitude of 43% for output, 53% for consumption, 40% for investment and 73% for labor.[22]

In summary, this model is a reasonably good model of business cycles. It is does not match the data on all dimensions, but it reflects the latest research in modeling business cycles. In doing so, it inherits the imperfections of these models.

Before moving to the welfare cost calculations it is useful to contrast the estimates for the parameters of the utility function here with Heaton's (1995) estimates. Heaton estimated the parameters using consumption and asset return data in a model with an exogenously specified consumption process. His parameter estimates were: $\gamma = 1.762$, $\delta = 0.934$, $\theta = 0.921$, and $\alpha = 0.672$. His estimates fall well outside the 5% and 95% quantiles of the posterior estimated here. Heaton's estimates imply an aversion to low frequency volatility.[23] For the estimates here, the agent dislikes high frequency volatility relative to low frequency volatility--the opposite of Heaton's estimates. The estimate for δ in table 1, which governs the amount of substitutability in consumption across periods, is close to zero. It appears then, that consumption is complementary across time periods, and this effect only lasts for a few periods. Another important difference in the estimates in table 1 relative to previous estimates of the non-separability parameter is that the degree of complementarity is much lower (e.g., Constantinides 1990, Heaton 1995, Jermann 1998, Abel 1990). In other words, while there is an aversion to high frequency volatility with the parameter estimates here, that aversion does not appear to be extreme. However, given a choice, if this agent is faced with a fluctuating consumption process she would still prefer that those fluctuations be concentrated at low frequencies. The strength, in quantitative terms, of the preference for low frequency and aversion to high frequency volatility is addressed in the next section.

---

[22] This is not to imply that this is the first or the only model that is an improvement over the standard model using Watson's measure of fit. For example in Wen (1998) the Watson error for Output, Consumption, Investment and labor growth is 0.09, 0.75, 0.09, 0.28, in Perli and Sakellaris (1998) the error is 0.17, 0.52, 0.55 and 0.46.

[23] Heaton's parameter estimates imply a welfare cost of 0.25%, the 40% welfare cost reported in section II used Heaton's parameters except α was set equal to 0.93.

Figure 1: Model and Data Spectra

a: Output Growth

b: Consumption Growth

c: Investment Growth

d: Labor Growth
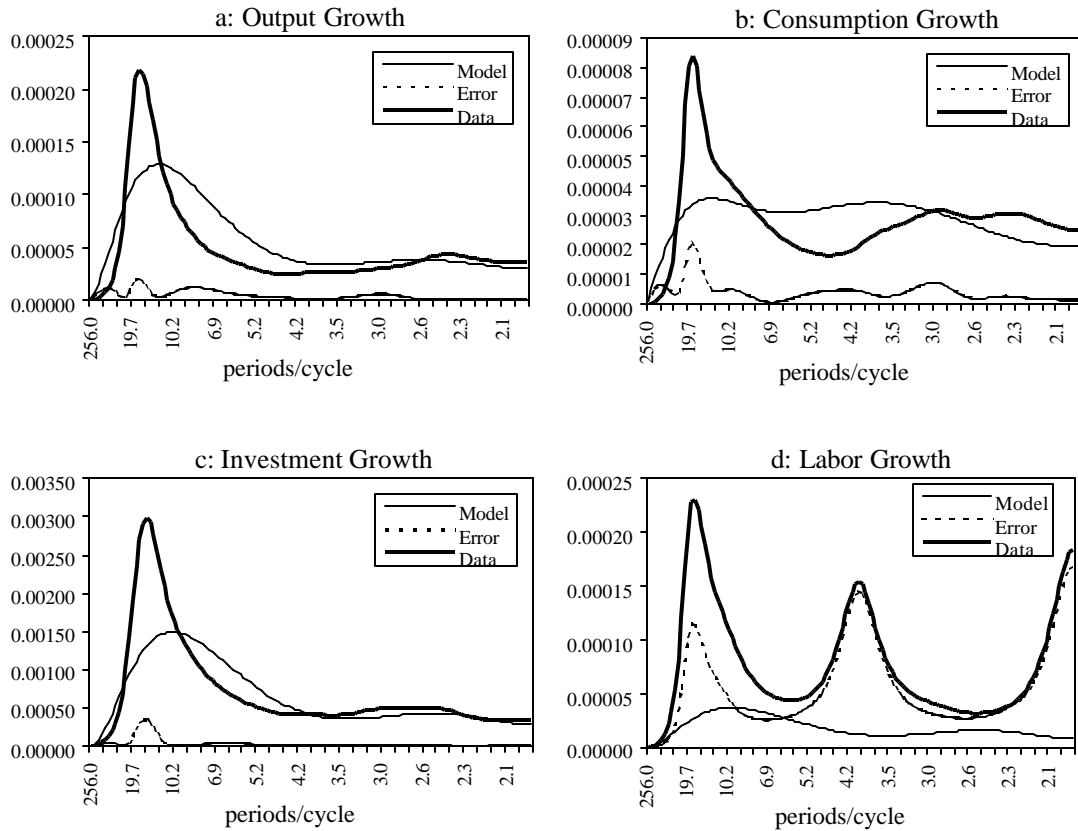
VI   Welfare Cost Estimates

The welfare cost of consumption volatility in the model economy can be measured by comparing the expected utility of the representative agent in the fluctuating economy with expected utility in the steady state. I then calculate how much consumption in the steady state needs to be reduced to make the agent indifferent between the steady state and the fluctuating economy. Expected utility in the fluctuating economy can be found by simulating the economy described by (10)-(11), using the minimum expected loss parameter estimates, and calculating utility. This is repeated and the mean of utility across simulations is the expected utility of the agent in the fluctuating economy. Since utility is separable in consumption and leisure we can decompose the welfare cost into the fraction due to consumption volatility and the fraction due to leisure volatility. Since the focus here is on the cost of consumption volatility, and the model is not a good model of the labor sector, the cost of leisure volatility is not reported.

The welfare cost of all consumption volatility in the U.S. is only 0.0044 percent per year, or $3.52 per person in 1997 dollars. This is similar to the estimate made by Lucas with time separable preferences, which was 0.008 percent. The figure calculated here is strikingly different than the estimates presented

elsewhere with non-separable preferences. It is not the case, however, that this result was obvious ex-ante because we have an optimizing model in which the agent can choose the consumption that he or she likes, therefore minimizing the costliness of volatility. The agent is choosing the consumption process that we observe, but this does not necessarily imply that he is happy about this choice. The agent is only responding optimally to the exogenous shocks, given the frictions in the model. The small welfare cost found here is an empirical outcome of the model parameter estimates, not a reflection that the nature of the optimizing model necessarily predicts a low welfare cost. It is easy to imagine alternative parameters or an alternative hypothetical world with frictions and preferences such that consumption volatility is both high and unavoidable--leading to a high cost of fluctuations in an optimizing model.

VII   Sensitivity of Results

One feature of the welfare calculations presented above is that the preference parameter estimates, and hence the welfare estimates, depend on the model and the assumptions used in estimation. It is worthwhile to know how the results will change with alternative models, prior views over the parameters, and data used in the estimation. Of course, when altering the model it is important to note that the model 'fit' may be worse. Table 4 reports the preference parameter estimates obtained with the loss function, model fit using Watson's metric, and welfare cost estimates for alternative models, priors and data.[24] The first column of the table indicates the alteration from the baseline model presented in section III. The first row contains the relevant information from the baseline model.

The technology of the model is the key ingredient in generating persistent business cycles. The features were chosen to work in concert; however, by removing some of the parts we can get an idea of the marginal contribution of some of the features of the model. Additionally, we will learn if the preference parameter estimates depend critically on any one feature of the model. In table 4 we see that removing the dynamic complementarity results in a small decline in model fit. However, the preference parameters have changed significantly. While there appears to be a stronger amount of 'habit' (larger α), the curvature parameter is closer to linear preferences. The result is a decline in the welfare cost estimate. Removing the time-to-build feature has almost no effect on the preference parameter estimates, but results in a substantial decline in model fit. The welfare cost estimate falls again.

Next, the baseline model was estimated using investment and labor data.[25] The non-separability preference parameter estimates were left virtually unchanged while the curvature parameter increased from 0.55 in the baseline model to 1. However, the fit of the model deteriorates significantly for the

---

[24] Full details of all parameter estimates and model moments and not reported due to space constraints.

[25] As noted previously, labor hours data is only available through 1995, so the sample period for this estimation is 1948:1-1995:1.

consumption sector. In fact the model fit for the consumption sector is worse than the KPR model. Fit in the labor sector improved, but the error remains high at 0.5. The welfare cost estimate is left unchanged.

Another issue is the treatment of the raw data before estimation. Here the seasonally adjusted data have been detrended using an HP filter. An alternative filtering procedure is to use a high-pass filter to eliminate the trend or low frequency movements.[26] The consumption and investment data were detrended using Baxter and King's (1993) band-pass filters, and the estimation exercise repeated. From Table 4 we see that this has no affect on the results.[27]

Table 4 also contains estimates of the model with alternative prior beliefs. The model was estimated with flat priors on all of the parameters. Results are reported with and without the loss function. For the case without the loss function, the parameters given in the table are the median of the posterior. In both cases, the estimate of the degree of 'habit' increased, though the model fit in the consumption sector deteriorated significantly without the loss function. Without the loss function the model fit in the labor sector deteriorates significantly. The welfare cost estimate remains trivially small. The utility function curvature prior in the baseline model is centered around 1 with a rather tight dispersion. When I increase the mean to 5 and loosen the dispersion to 1, the model parameter estimates are similar to the baseline model, and the welfare cost estimate declines.

Table 4: Alternative Model Estimates

	Preference Parmeters				Spectral Fit				Welfare
	q	a	d	g	Out	Con	Inv	Lab	Cost
Baseline Model	0.10	0.21	0.07	0.55	0.09	0.14	0.04	0.72	0.0044
No Time-To-Build	0.02	0.21	0.01	0.43	0.20	0.22	0.11	0.77	0.0023
No Complementarity	0.16	0.70	0.16	0.14	0.12	0.15	0.07	0.72	0.0018
Estimate w/ Labor & Inv. data	0.12	0.21	0.03	1.03	0.27	0.59	0.06	0.50	0.0039
Estimate w/ BP Filtered data	0.06	0.27	0.03	0.50	0.15	0.22	0.08	0.70	0.0040
Estimate w/ Flat Prior (loss fnc)	0.13	0.44	0.12	0.45	0.08	0.12	0.03	0.90	0.0038
Estimate w/ Flat Prior (no loss fnc)	0.19	0.57	0.10	0.62	0.17	0.38	0.04	0.72	0.0058
Estimate w/ alt. Curvature Prior	0.15	0.31	0.07	0.42	0.10	0.13	0.04	0.73	0.0028

---

[26] Of course, an HP filter can be viewed as a type of high-pass filter. The high-pass filter used here eliminates fluctuations that are of 200-period cycles and lower.

[27] A caveat with using seasonally adjusted data is that if the seasonal adjustment procedure employed by the Bureau of Economic Analysis is flawed, and alters the spectral characteristics of the consumption at frequencies other than seasonal frequencies, then the welfare cost estimates may also be flawed.

VIII   Alternative Views of the Welfare Cost of Business Cycles

The results presented above do not appear to be dramatically affected by assumptions on the technology, data or prior beliefs. However, an important caveat is that larger departures from framework above may lead to larger welfare cost estimates. First, the results are based on a complete markets rational expectations framework. One could relax the utility assumption and assume alternative forms for preferences. One prominent alternative for preferences is prospect theory. With prospect theory, people are more sensitive to losses than gains. Benartzi and Thaler (1995) show that these preferences can rationalize the observed equity premium. These preferences may potentially lead to large welfare costs associated with the losses experienced in recessions, losses that agents in the expected utility framework do not seem sensitive to.

Second, fluctuations may be very costly if they reduce the growth rate of consumption. That is, if in the process of eliminating fluctuations we raise the trend rate of growth, then the gains from eliminating fluctuations would be very large. The gains would be large with both time-separable and time-non-separable preferences. Ramey and Ramey (1995) provide empirical evidence of a link between growth and volatility. They argue that the uncertainty inherent in fluctuations lowers the mean rate of growth.

Third, many aspects of business cycles have been left out of this model. In particular, the labor/leisure implications are largely ignored. Recessions may force people who are laid off to accept jobs for which their human capital is not well matched, which may have significant long-term welfare consequences. Finally, only the cost of volatility in non-durables and services consumption is included in the analysis. There may be a cost to the volatility in durable goods consumption. Each of these issues could be investigated in future work.

Finally, even if one does not find credible the potentially large gains from the alternatives listed above, the results here should not be taken to imply that current efforts to smooth aggregate fluctuations are not worthwhile. The estimates provided here, and elsewhere, are aimed at measuring the gains from removing the residual risk. The gains to moving from a regime in which there are no efforts to smooth aggregate volatility to the current regime may be very large.

IX   Conclusion

That people may have temporally-dependent preferences carries much intuitive appeal. Through introspection we can argue that, yes, we do care about the types of consumption volatility that we face. It does not seem reasonable that someone faced with choosing a consumption process with high-frequency volatility or one with equal variance but with volatility concentrated at low frequencies, would be indifferent to that choice. Yet this is what standard time-separable preferences imply. The idea that people

are not indifferent has spawned a large literature trying to develop the quantitative implications of temporally-dependent preferences. A problem, however, is the choice of the free parameters in the utility function. How do we choose these parameters so that results that we obtain are not merely due to a proliferation of 'free' parameters?

The approach in this paper is to estimate preferences in the context of an optimizing business cycle model. The motivation was to find the form for preferences that would be capable of generating the type of consumption fluctuations we observe. There is a sizeable literature that has estimated the non-separability parameters in the agent's utility function using asset return data and an exogenously specified consumption process. The results obtained here are very different than the results from those exercises, and cast some doubt as to whether the non-separabilities estimated in these papers are reasonable. The non-separability parameter estimates in this paper would not be able to reconcile consumption fluctuations with asset return data (i.e. solve the equity premium puzzle), however, I do not view this as a shortcoming of the measure of welfare cost in this paper, but rather the result of a shortcoming in asset pricing models. The question of how non-separabilities in preferences resolve asset pricing puzzles, and whether or not this resolution is reasonable, is addressed in Otrok, Ravikumar and Whiteman (1998).

The focus of this paper has been on developing an approach to place some discipline on preference-based measures of the welfare cost of business cycles. The method was applied to a form of time-non-separable preferences that encompasses many of the most popular forms of non-separabilities. The result was an estimate for preferences that yielded a welfare cost on the same order of magnitude as that of Lucas (1987). However, this does not suggest that the welfare cost of business cycles is actually zero--and hence that we should abandon business cycle theory. Instead, it is a result that casts doubt that empirically plausible modifications to preferences alone could lead to large costs of consumption volatility.

References

Abel, Andrew B., 1990, "Asset Pricing Under Habit Formation and Catching Up With The Joneses," *American Economic Review* 80:38-42.

Abel, Andrew B., 1999, "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics* 43:3-34.

Atkeson, Andrew and Christopher Phelan, 1994, "Reconsidering the Costs of Business Cycles with Incomplete Markets," NBER Working Paper 4719.

Baxter, Marianne, and Robert G. King, 1993, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," forthcoming.

Benartzi, Shlomo and Richard H. Thaler, 1995, "Myopic Loss Aversion and The Equity Premium Puzzle," *The Quarterly Journal of Economics* 110:73-92.

Benhabib Jess, Farmer Roger E.A., 1996, "Indeterminacy and sector-specific externalities," *Journal of Monetary Economics* 37:421-443

Benhabib, Jess, Roberto Perli and Plutarchos Sakellaris 1998, "Persistence of Business Cycles in Multisector RBC Models," Working Paper, University of Pennsylvania.

Benhabib Jess, Richard Rogerson and Randall Wright, 1991, "Homework in Macroeconomics: Household Production and Aggregate Fluctuations," *Journal of Political Economy* 99:1166-1187.

Boldrin, Michele, Lawrence J. Christiano and Jonas D.M. Fisher, 1997, "Habit Persistence and Asset Returns in an Exchange Economy," *Macroeconomic Dynamics* 1, 312-332.

Boldrin, Michele, Lawrence J. Christiano and Jonas D.M. Fisher, 1995, "Asset Pricing Lessons for Modeling Business Cycles," Working paper, Northwestern University.

Burnside, Craig and Martin Eichenbaum, 1996, "Factor Hoarding and the Propagation of Business Cycle Shocks," *American Economic Review* 86:1154-1174.

Campbell, John Y., 1998, "Asset Prices, Consumption and the Business Cycle," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, forthcoming.

Campbell, John Y., and John H. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy* forthcoming.

Carroll, Christopher D., Jody Overland and David N. Weil, 1997, "Comparison Utility in a Growth Model," NBER Working Paper 6138.

Chib, Siddhartha and Edward Greenberg, 1996, "Markov Chain Monte Carlo Simulation Methods in Econometrics," *Econometric Theory* 12:409-431.

Christiano, Lawrence and Wouter Den Haan, 1995, "Small Sample Properties of GMM for Business Cycle Analysis," *Journal of Business and Economic Statistics* 14:309-327.

Cogley, Timothy and James M. Nason, 1995a, "Output Dynamics in Real-Business-Cycle Models," *American Economic Review*, 85:492-511.

Cogley, Timothy and James M. Nason, 1995b, "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research," *Journal of Economic Dynamics and Control*, 19:253-278.

Constantinides, George M., 1990, "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy* 98:519-543.

Cooper, Russell and John Haltiwanger, 1996, "Evidence on Macroeconomic Complementarities", *The Review of Economics and Statistics* 78:78-93.

Cooper, Russell and Alok Johri, 1996, "Dynamic Complementarities: A Quantitative Analysis," *Journal of Monetary Economics* 40:97-120.

DeJong, David N. and Beth F. Ingram, 1998, "The Cyclical Behavior of Skill Acquisition," Manuscript, University of Iowa.

DeJong, David N., Beth F. Ingram, Yi Wen and Charles H. Whiteman, 1997, "Cyclical Implications of the Variable Utilization of Physical and Human Capital," Manuscript, University of Iowa.

DeJong, David N., Beth F. Ingram and Charles H. Whiteman, 1998a, "Keynesian Impulses Versus Solow Residuals: Identifying Sources of Business Cycle Fluctuations," Manuscript, University of Iowa.

DeJong, David N., Beth F. Ingram and Charles H. Whiteman, 1998b, "A Bayesian Approach to Dynamic Macroeconomics," Manuscript, University of Iowa.

Dolmas, Jim, 1998, "Risk Preferences and the Welfare Cost of Business Cycles," *Review of Economic Dynamics* 1:646-676.

Eichenbaum, Martin, Lars P. Hansen and Kenneth Singleton, 1988, "A Time-Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty," *Quarterly Journal of Economics* 103:51-78.

Geweke, John, 1989, "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica* 57:1317-1339.

Geweke, John, 1997, "Posterior Simulators in Econometrics," in D. Kreps and K.F. Wallis (eds.) *Advances in Economics and Econometrics: Theory and Applications*, vol III, Cambridge: Cambridge University Press, 128-165.

Greenwood, Jeremy, Zvi Hercowitz and Gregory Huffman, 1988, "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78:402-417.

Hamilton, James, 1994, *Time Series Analysis*, Princeton University Press.

Hansen, G., 1985, "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics* 16:309-28.

Heaton, John, 1995, "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," *Econometrica* 63:681-717.

Huffman, Gregory and Mark Wynne, 1998, "The Role of Intratemporal Adjustment Costs in a Multi-Sector Economy," *Journal of Monetary Economics* Forthcoming.

Jermann, U., 1998, "Asset Pricing in Production Economies," *Journal of Monetary Economics* 41:257-276.

King, Robert G., Charles I. Plosser and Sergio T. Rebelo, 1988, "Production, Growth and Business Cycles", *Journal of Monetary Economics* 21:195-231.

King, Robert G. and Sergio T. Rebelo, 1998, "Resuscitating Real Business Cycles," *Handbook of Macroeconomics* Forthcoming.

Klein, Paul, 1998, "Using the Generalized Schur Form to Solve a System of Linear Expectational Difference Equations," forthcoming, *Journal of Economic Dynamics and Control.*

Kydland, Finn, and Edward C. Prescott, 1982, "Time To Build and Aggregate Fluctuations," *Econometrica* 50:1345-1370.

Lettau, Martin and Harald Uhlig, 1995, "Can Habit Formation be Reconciled with Business Cycle Facts?," *Review of Economic Dynamics*, forthcoming.

Lucas, Robert E., Jr., 1987, *Models of Business Cycles*, Basil Blackwell, New York.

Lucas, Robert E., Jr., 1981, *Studies in Business-Cycle Theory*, MIT Press, Cambridge, Massachusetts.

McGrattan, Ellen R., 1994, "The Macroeconomic Effects of Distortionary Taxation," *Journal of Monetary Economics* 33:573-601.

Obstfeld, Maurice, 1994, "Evaluating Risky Consumption Paths: The role of intertemporal substitutability," *European Economic Review*, 38:1471-1486.

Otrok, Christopher, 1999, "Spectral Welfare Cost Functions," forthcoming, *International Economic Review*.

Otrok, Christopher, B. Ravikumar and Charles H. Whiteman, 1998, "Habit Formation: A Resolution to the Equity Premium Puzzle?," Working Paper, University of Iowa.

Pemberton, J., 1996, "Growth Trends, Cyclical Fluctuations, and Welfare with Non-Expected Utility Preferences," *Economics Letters* 50:387-392.

Ramey, Garey and Valerie Ramey , 1995, "Cross-Country Evidence on the Link between Volatility and Growth," *American Economic Review* 85:1138-1151.

Rogerson, Richardson, 1988, "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics* 21:3-16.

Rotemberg, Julio J., and Michael Woodford, 1996, "Real-Business-Cycle Models and the Forecastable Movements in Output, Hours and Consumption," *American Economic Review* 86:71-89.

Rouwenhorst, K, 1991, "Time to Build and Aggregate Fluctuations," *Journal of Monetary Economics* 27:241-254.

Sims, Christopher A., 1996, "Solving Linear Rational Expectations Models," Manuscript, Yale University.

Tallarini, Thomas D. Jr., 1998, "Risk-Sensitive Real Business Cycles," Manuscript, Carnegie Mellon University.

Tierney, Luke, 1994, "Markov Chains for Exploring Posterior Distributions", *The Annals of Statistics* 22:1701-1762.

Van Wincoop, Eric, 1994, "Welfare Gains from International Risk Sharing," *Journal of Monetary Economics* 34:175-200.

Watson, M., 1993, "Measures of Fit for Calibrated Models," *Journal of Political Economy* 101:1011-1041.

Wen, Yi, 1996, "Where's the Beef? The Trivial Dynamics of RBC Models," Manuscript.

Wen, Yi, 1998a, "Investment Cycles," *Journal of Economic Dynamics and Control* 22:1139-1165.

Wen, Yi, 1998b, "Can A Real Business Cycle Model Pass the Watson Test?," *Journal of Monetary Economics* 42:185-203.

Appendix: Model Correlations

Table A1: Model and Data Correlation Matrix (Levels)

| | Output | | | | | Consumption | | | | | Investment | | | | | Labor | | | | | Cons. Labor | | | | |
| | Model | | | | Data | Model | | | | Data | Model | | | | Data | Model | | | | Data | Model | | | | Data |
	5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss	
Output	1.00	1.00	1.00	1.00	1.00																				
Cons.	0.72	0.78	0.84	0.90	0.89	1.00	1.00	1.00	1.00	1.00															
Invest.	0.91	0.93	0.95	0.90	0.90	0.39	0.51	0.60	0.62	0.61	1.00	1.00	1.00	1.00	1.00										
Labor	0.56	0.75	0.87	0.91	0.85	-0.06	0.22	0.44	0.63	0.75	0.79	0.91	0.97	0.99	0.78	1.00	1.00	1.00	1.00	1.00					
C. Labor	-0.91	-0.76	-0.45	0.10	NA	-0.73	-0.42	0.04	0.50	NA	-0.91	-0.80	-0.62	-0.30	NA	-0.80	-0.52	-0.17	-0.23	NA	1.00	1.00	1.00	1.00	1.00
Inv Labor	0.79	0.83	0.86	0.81	NA	0.17	0.31	0.42	0.47	NA	0.96	0.97	0.98	0.98	NA	0.91	0.96	0.99	0.97	NA	-0.89	-0.73	-0.53	0.81	NA

Table A2: Model and Data Correlation Matrix (Growth)

| | Output | | | | | Consumption | | | | | Investment | | | | | Labor | | | | | Cons. Labor | | | | |
| | Model | | | | Data | Model | | | | Data | Model | | | | Data | Model | | | | Data | Model | | | | Data |
	5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss		5%	Med	95%	Loss	
Output	1.00	1.00	1.00	1.00	1.00																				
Cons.	0.59	0.67	0.73	0.72	0.83	1.00	1.00	1.00	1.00	1.00															
Invest.	0.87	0.90	0.92	0.83	0.79	0.16	0.27	0.36	0.22	0.32	1.00	1.00	1.00	1.00	1.00										
Labor	0.18	0.42	0.65	0.76	0.51	-0.57	-0.37	-0.13	0.11	0.32	0.58	0.76	0.91	0.98	0.51	1.00	1.00	1.00	1.00	1.00					
C. Labor	-0.93	-0.84	-0.72	-0.26	NA	-0.90	-0.79	-0.44	0.34	NA	-0.86	-0.62	-0.44	-0.65	NA	-0.60	0.04	0.42	-0.58	NA	1.00	1.00	1.00	1.00	1.00
Inv Labor	0.64	0.71	0.77	0.71	NA	-0.15	-0.05	0.06	0.03	NA	0.92	0.95	0.97	0.98	NA	0.84	0.92	0.98	0.99	NA	-0.75	-0.35	-0.10	0.71	NA