CES Technology and Business Cycle Fluctuations*

Cristiano Cantore
University of Surrey

Paul Levine
University of Surrey

Bo Yang
University of Surrey and London Metropolitan University

October 22, 2010

Abstract

This paper contributes to a rapidly rising literature that brings the CES specification of the production function into the analysis of business cycle fluctuations. The main result is to confirm decisively the importance of CES rather than Cobb-Douglas (CD) production functions. Indeed in a marginal likelihood race assuming equal prior model probabilities, CES beats the CD production function with posterior model probabilities $0.999988 : 0.00012$ in favour of the CES. The reason for this result is that movements of factor shares with the CES specification help substantially to fit the data better. The marginal likelihood improvement is matched by the ability of the CES model to get closer to the data in terms of second moments, especially the volatilities of output, consumption and the real wage, and the autocorrelation functions for inflation and the nominal interest rate. A comparison with a DSGE-VAR further confirms the ability of the CES model to reduce model misspecification. Using US data, we estimate by Bayesian-Maximum-Likelihood methods a elasticity of substitution of elasticity well below unity at 0.36, a value broadly in line with the literature using other methods of estimation. The main message then for DSGE models is that we should dismiss once and for all the use of CD for business cycle analysis.

JEL Classification: C11, C52, D24, E32.
Keywords: CES production function, DSGE model, Bayesion estimation, DSGE-VAR

*To be presented at the MONFISPOL Conference, Nov 4 – 5, London Metropolitan University. We acknowledge financial support from the EU Framework Programme 7 and from the ESRC, project no. RES-062-23-2451.

Contents

1 Introduction

This paper aims to extend the DSGE model developed by Christiano *et al.* (2005) and Smets and Wouters (2007) to allow for a richer and more data coherent specification of the production side of the economy. The idea is to enrich what has become the workhorse DSGE model by substituting the usual Cobb-Douglas production assumption with a more general CES production function which allows for cyclical variations in factor shares, estimation of the capital/labour elasticity of substitution and biased technical change.

The CES production function has been used extensively in many area of economics since the middle of the previous century (Solow (1956) and Arrow *et al.* (1961)). Thanks to La Grandville (1989), who introduced the concept of *normalization*, it has been extensively used in growth theory. Indeed La Grandville (1989) showed that it was possible to obtain a perpetual growth in income per-capita, even without any technical progress. Furthermore factor substitution and the bias in technical change feature an important role in many other areas of economics[1] but, until recently have been largely disregarded in business cycle analysis. On the empirical side León-Ledesma *et al.* (2010) show that normalization improves empirical identification.[2]

The concepts of biased technical change and imperfect factor substitutability between factors of production has been introduced in business cycle analysis by Cantore *et al.* (010a). They show that the introduction of a *normalized* CES production function into an otherwise standard RBC and/or NK DSGE model significantly changes the response of hours worked to a technology shock under both price-setting mechanisms and that such response might change as well within each model depending on the parameters related to the production process (developing a *threshold* rule for the "impact" of a technology shock on hours worked). They also show how the introduction of biased technical change and imperfect substitutability allow movement in factor shares which appear to fluctuate at business cycle frequencies in the data but are *theoretically* constant under the Cobb-

[1] The value of the substitution elasticity has been linked to differences in international factor returns and convergence (e.g., Klump and Preissler (2000), Mankiw (1995)); movements in income shares (Blanchard (1997), Caballero and Hammour (1998), Jones (2003)); the effectiveness of employment creation policies (Rowthorn (1999)), etc. The nature of technical change, on the other hand, matters for characterizing the welfare consequences of new technologies (Marquetti (2003)); labour-market inequality and skills premia (Acemoglu (2002)); the evolution of factor income shares (Kennedy (1964), Acemoglu (2003)) etc.

[2] They show that using a normalized approach permits to overcome the 'impossibility theorem' stated by Diamond *et al.* (1978) and simultaneously identify the elasticity of substitution and biased technical change.

Douglas specification. Indeed there is mounting evidence in the literature [3] that whilst constant factor shares might be a good approximation for growth models where the time span considered is very long, at business cycle frequencies those shares are not constant. This is clearly showed in Figure 1 for the US data used to estimate our model.

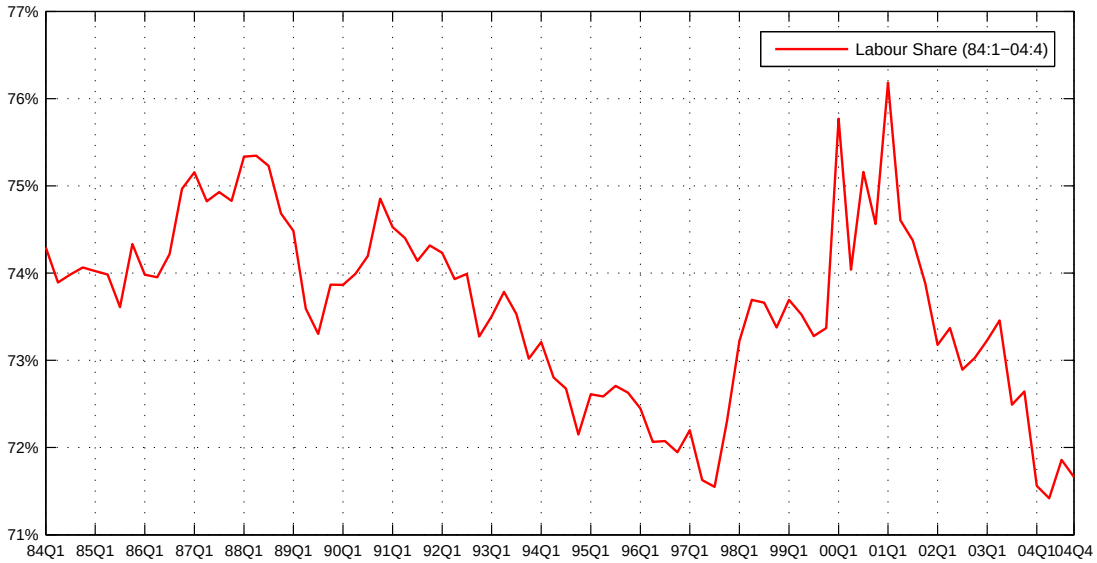


Figure 1: US Labour Share (Source: Department of Labor, U.S. Bureau of Labor Statistics)

Furthermore Cantore *et al.* (010b) test empirically the model(s) developed by Cantore *et al.* (010a) using rolling-windows Bayesian techniques in order to check if the documented time-varying relation between hours worked, productivity and output (see Fernald (2007) and Galí and Gambetti (2009) among others) can be explained using the *threshold* rule. They are also the first to show that introducing a CES production function in an estimated RBC model significantly improves the fit of the model with respect to the Cobb-Douglas case.

Apart from Cantore *et al.* (010a) and Cantore *et al.* (010b) usually DSGE models continue using the Cobb-Douglas assumption even if the empirical evidence provided through the years has now definitely ruled out the possibility of unitary elasticity of substitution (see among others Antràs (2004), Klump *et al.* (2007), Chirinko (2008) and León-Ledesma *et al.* (2010)).

[3] See for example Blanchard (1997), Jones (2003, 2005), McAdam and Willman (2008) and Santeulália-Llopis and Ríos Rull (2009).

In this paper we show that by introducing a CES production function in a medium-scale DSGE model in the spirit of Christiano *et al.* (2005) and Smets and Wouters (2007) makes it possible to exploit the movements of factor shares we observe in the data to improve significantly the performance of the model. Our main result is that in terms of model posterior probabilities, impulse responses, second moments and autocorrelations, the assumption of a CES technology significantly improves the model fit. Using US data, we estimate by Bayesian-Maximum-Likelihood (BML) methods a elasticity of substitution of elasticity between the capital/labour ratio and the wage rate/capital cost ratio to be 0.36, a value broadly in line with the literature using other methods of estimation.[4]

The rest of the paper is organized as follows. Section 2 describes the model with particular attention paid to the normalization of the CES production function. Section 3 sets out the BML estimation of the model. Section 4 examines the ability of the model to capture the main characteristics of the actual data as described by second moments and the impulse response functions of an estimated "DSGE-VAR" hybrid. Section 5 compares the variance decomposition of the structural shocks for the CES and Cobb-Douglas formulations. Section 6 concludes the paper.

2 The Augmented SW Model

Here we present, concisely the augmented SW model with a wholesale and a retail sector, Calvo prices and wages, CES production function, adjustment costs of investment and variable capital utilization. Figure 2 illustrates the model structure. The model equilibrium conditions are presented in non-linear form. The novel feature is the introduction of a CES production function in the wholesale sector, instead of the usual Cobb-Douglas form. This generalization then allows for the identification of both labour-augmenting and capital-augmenting technology shocks. As in Smets and Wouters (2007) we use a household utility function compatible with a balance growth path in the steady state, but we adopt a more standard functional form used in the RBC literature. However we do not adopt Kimball aggregators for final output and composite labour.[5] Again as in their

[4]See, for example, Table 2 in Rowthorn (1999), Chirinko (2008) and León-Ledesma *et al.* (2010)).

[5]The motivation for generalizing Dixit-Stiglitz aggregators is to bring estimates of price and wage contract lengths into line with micro-econometric evidence. In fact our estimates for US data are compatible with the simpler Dixit-Stiglitz formulation.

paper we introduce a monopolistic trade-union that allows households to supply homogeneous labour. Then as long as preference shocks are symmetric households are identical in equilibrium and the complete market assumption is no longer required for aggregation. The supply-side of the economy consists of competitive retail sector producing final output and a monopolistically competitive wholesale sector producing differentiated goods using the usual inputs of capital and work effort. Households consume a bundle of differentiated commodities, supply labour and capital to the production sector, save and own the monopolistically competitive firms in the goods sector. Capital producers provide the capital inputs into the wholesale sector.[6]

We set out the model first without specifying the form of the utility and production functions in order to obtain a flexible framework in which it will be easy to stick different functional forms.

The sequencing of decisions is as follows[7]

1. Each household supplies homogeneous labour at a price $W_{h,t}$ to a trade-union. Households choose their consumption, savings and labour supply given aggregate consumption (determining external habit). In equilibrium all household decisions are identical.

2. Capital producing firms that at time convert final output into new capital which is sold on to intermediate firms.

3. A monopolistic trade-union differentiates the labour and sells type $N_t(j)$ at a price $W_t(j)$ to a labour packer in a sequence of Calvo staggered wage contracts. In equilibrium all households make identical choices of total consumption, savings, investment and labour supply.

[6]There are other differences with Smets and Wouters (2007): (i) Our price and wage mark-up shocks follow an AR(1) process instead of the ARMA process chosen by SW; (ii) in SW the government spending shock is assumed to follow an autoregressive process which is also affected by the productivity shock; (iii) we have a preference shock instead of the risk-premium shock. Chari *et al.* (2009) criticized the risk premium shock arguing that has little interpretation and in unlikely to be invariant to monetary policy. We prefer our somewhat simpler set-up and we expect none of the differences to affect the main focus of the paper which is on the comparison between CD and CES production functions.

[7]Sequencing matters for the monopolistic trade-unions and intermediate firms who anticipate and exploit the downward-sloping demand for labour and goods respectively. Different set-ups with identical equilibria are common in the literature. Monopolistic prices can be transferred to the retail sector. When it comes to introducing financial frictions, for example, as in Gertler and Karadi (2009) the introduction of separate capital producers as in our set-up is convenient, but not essential in the SW model without such frictions.

4. The competitive labour packer forms a composite labour service according to a constant returns CES technology $N_t = \left(\int_0^1 N_t(j)^{(\zeta-1)/\zeta} dj\right)^{\zeta/(\zeta-1)}$ and sells onto the intermediate firm.

5. Each intermediate monopolistic firm f using composite labour and capital rented from purchased from capital producers to produce a differentiated intermediate good which is sold onto final goods firm at price $P_t(f)$ in a sequence of Calvo staggered price contracts.

6. Competitive final goods firms use a continuum of intermediate goods according to another constant returns CES technology to produce aggregate output $Y_t = \left(\int_0^1 Y_t(f)^{(\mu-1)/\mu} df\right)^{\mu/(\mu-1)}$.

We now solve the model by backward induction starting with the production of final goods.

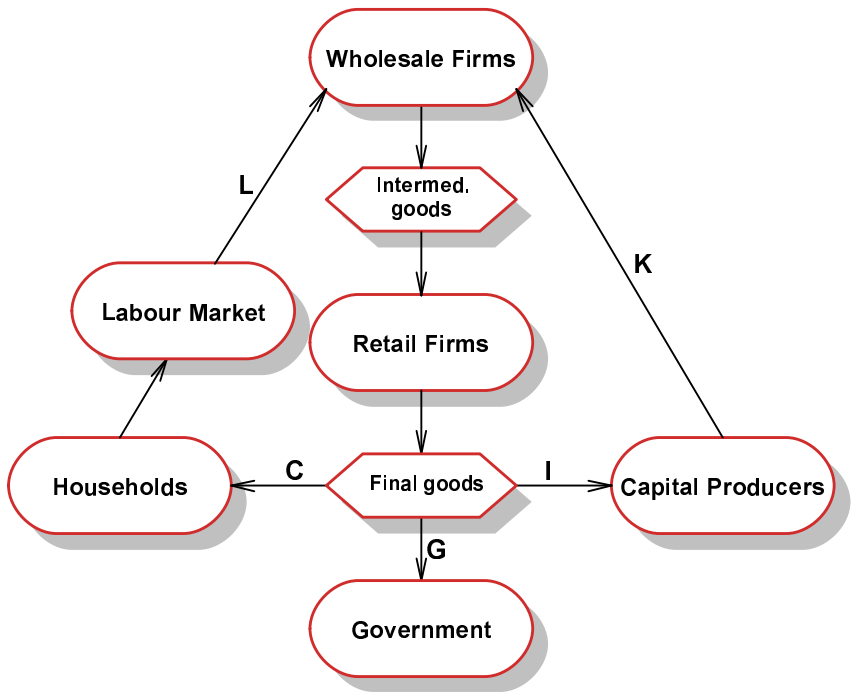


Figure 2: Model Structure

2.1 Final Goods

Each final goods firms minimizes the cost $\int_0^1 P_t(f)Y_t(f)df$ of producing the final output $Y_t = \left(\int_0^1 Y_t(f)^{(\zeta-1)/\zeta}df\right)^{\zeta/(\zeta-1)}$. This leads to the standard result for the Dixit-Stigliz aggregator

$$Y_t(f) = \left(\frac{P_t(f)}{P_t}\right)^{-\zeta} Y_t \tag{1}$$

$$P_t = \left[\int_0^1 P_t(f)^{1-\zeta}df\right]^{\frac{1}{1-\zeta}} \tag{2}$$

$$P_tY_t = \int_0^1 P_t(f)Y_t(f)df \tag{3}$$

where P_t is an aggregate price index. Note that (1) and (3) imply (2).

2.2 Intermediate Firms

In the intermediate goods sector each good f is produced by a single firm f using composite labour and capital with a technology:

$$Y_t(f) = (1-c)F(ZK_t, ZN_t, N_t, U_tK_t) \tag{4}$$

where c are fixed costs of production and U_t allows for variable capital utilization. The parameter c is pinned down by a free-entry condition that drives profits in the steady state to zero. Given that at this stage we do not specify the form of the production function we allow for all the possible specification of technology shocks. Calling ZK_t capital-augmenting and ZN_t labour-augmenting we are in the case of Hicks neutrality if $ZK_t = ZN_t > 0$, Solow neutrality if $ZK_t > 0$ and $ZN_t = 0$ and Harrod neutrality in the case of $ZK_t = 0$ and $ZN_t > 0$. Then minimizing costs $P_tRR_t^KU_t(f)K_t(f) + W_tN_t(f)$ leads to

$$\frac{W_t}{P_t} \equiv MPL_t = MC_t(f)F_{N,t} \tag{5}$$

$$RR_t^K \equiv MPK_t = MC_t(f)F_{K,t} \tag{6}$$

where MPL_t and MPK_t are the marginal products of labour and capital respectively, RR_t^K is the real cost of capital. As usual the firm's cost minimizing real marginal costs

$(MC_t(f))$ is given by the Lagrange multiplier related to the production function constraint.

Pricing by the firm follows the standard Calvo framework supplemented with indexation. At each period there is a probability of $1-\xi_p$ that the price is set optimally.[8] The optimal price derives from maximizing discounted profits. For those firms and workers unable to reset, prices are indexed to last period's aggregate inflation, with indexation parameter γ_p. With indexation parameter $\gamma_p \geq 0$, this implies that successive prices with no re-optimization are given by $P_t^0(f),\ P_t^0(f)\left(\frac{P_t}{P_{t-1}}\right)^{\gamma_p},\ P_t^0(f)\left(\frac{P_{t+1}}{P_{t-1}}\right)^{\gamma_p}, \ldots$. For each intermediate producer f the objective is at time t to choose $\{P_t^0(f)\}$ to maximize discounted profits

$$E_t \sum_{k=0}^{\infty} \xi_p^k \Lambda_{t,t+k} Y_{t+k}(f) \left[P_t^0(f) \left(\frac{P_{t+k-1}}{P_{t-1}} \right)^{\gamma} - P_{t+k} MC_{t+k} \right] \tag{7}$$

subject to $Y_{t+k}(f) = \left(\frac{P_t^0(f)}{P_{t+k}} \left(\frac{P_{t+k-1}}{P_{t-1}} \right)^{\gamma_p} \right)^{-\zeta} Y_{t+k}$ (from (1)), where $\Lambda_{t,t+k} \equiv \beta \frac{\Lambda_{C,t+k}/P_{t+k}}{\Lambda_{C,t}/P_t}$, is the nominal stochastic discount factor over the interval $[t, t+k]$ and ζ is the elasticity of substitution across intermediate goods. Since firms are atomistic, the aggregate price index and the discount factor are given in their calculations.

This leads to the following first-order condition:

$$E_t \sum_{k=0}^{\infty} \xi_p^k \Lambda_{t,t+k} Y_{t+k}(f) \left[P_t^0(f) \left(\frac{P_{t+k-1}}{P_{t-1}} \right)^{\gamma_p} - MS_{p,t} P_{t+k} \mathrm{MC}_{t+k} \right] = 0 \tag{8}$$

where we introduced, as usual in the literature, a time varying mark-up of prices over marginal costs $MS_{p,t} = \frac{\zeta}{(\zeta-1)} eP_t$ with eP_t being the price mark-up shock. Then by the law of large numbers the evolution of the price index is given by

$$P_{t+1}^{1-\zeta} \quad = \quad \xi_p \left(P_t \left(\frac{P_t}{P_{t-1}} \right)^{\gamma_p} \right)^{1-\zeta} + (1-\xi_p)(P_{t+1}^0(f))^{1-\zeta} \tag{9}$$

2.3 Labour Packer

As with final goods firms, the labour packer minimizes the cost $\int_0^1 W_t(j) N_t(j) dj$ of producing the composite labour service $N_t = \left(\int_0^1 N_t(j)^{(\mu-1)/\mu} dj \right)^{\mu/(\mu-1)}$. This leads to the

[8] Thus we can interpret $\frac{1}{1-\xi_p}$ as the average duration for which prices are left unchanged.

standard result for the Dixit-Stigliz aggregator

$$N_t(j) = \left(\frac{W_t(j)}{W_t}\right)^{-\mu} N_t \tag{10}$$

$$W_t = \left[\int_0^1 W_t(j)^{1-\mu} dj\right]^{\frac{1}{1-\mu}} \tag{11}$$

$$W_t N_t = \int_0^1 W_t(j) N_t(j) dj \tag{12}$$

where W_t is an aggregate wage index. Note that (10) and (12) imply (11).

2.4 Trade-Unions

Wage setting by the trade-union again follows the standard Calvo framework supplemented with indexation. At each period there is a probability $1-\xi_w$ that the wage is set optimally. The optimal wage derives from maximizing discounted profits. For those trade-unions unable to reset, wages are indexed to last period's aggregate inflation, with wage indexation parameter γ_w. Then as for price contracts the wage rate trajectory with no re-optimization is given by $W_t^0(j)$, $W_t^0(j)\left(\frac{P_t}{P_{t-1}}\right)^{\gamma_w}$, $W_t^0(j)\left(\frac{P_{t+1}}{P_{t-1}}\right)^{\gamma_w}, \cdots$. The trade union that buys homogeneous labour at a price $W_{h,t}$ and converts it into a differentiated labour service of type j. The trade union time t then chooses $W_t^0(j)$ to maximize

$$E_t \sum_{k=0}^{\infty} \xi_w^k \Lambda_{t,t+k} N_{t+k}(j) \left[W_t^0(j)\left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma_w} - W_{h,t+k}\right] \tag{13}$$

where $N_t(j)$ is given by (10) so that $N_{t+k}(j) = \left(\frac{W_t^0(j)}{W_{t+k}}\left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma_w}\right)^{-\eta} N_{t+k}$ and η is the elasticity of substitution across labour varieties. By analogy with (8) this leads to the following first-order condition

$$E_t \sum_{k=0}^{\infty} \xi_w^k \Lambda_{t,t+k} N_{t+k}(j) \Big[W_t^0(j)\left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma_w} - MS_{w,t} W_{h,t+k}\Big] = 0 \tag{14}$$

where $MS_{w,t} = \frac{\eta}{(\eta-1)} eW_t$ is the time varying wage mark-up with eW_t being the wage mark-up shock. Then by the law of large numbers the evolution of the wage index is given

by

$$W_{t+1}^{1-\eta} = \xi_w \left(W_t \frac{\left(\frac{P_t}{P_{t-1}}\right)^{\gamma_w}}{\frac{P_{t+1}}{P_t}} \right)^{1-\eta} + (1-\xi_w)(W_{t+1}^0(j))^{1-\eta} \tag{15}$$

2.5 Capital Producers

Capital producing firms convert I_t of output into $(1 - S(X_t))I_t$ of new capital sold at a real price Q_t. They then maximize expected discounted profits

$$E_t \sum_{k=0}^{\infty} D_{t,t+k} \left[Q_{t+k} ZI_{t+k} (1 - S\left(I_{t+k}/I_{t+k-1}\right)) I_{t+k} - I_{t+k} \right]$$

where $D_{t,t+k} \equiv \beta \frac{\Lambda_{C,t+k}}{\Lambda_{C,t}}$ is the *real* stochastic discount factor over the interval $[t, t+k]$. This results in the first-order condition

$$Q_t ZI_t (1 - S(X_t) - X_t S'(X_t)) + E_t \left[D_{t,t+1}\, Q_{t+1} ZI_{t+1} S'(X_{t+1}) \frac{I_{t+1}^2}{I_t^2} \right] = 1 \tag{16}$$

Capital accumulation is given by

$$K_{t+1} = (1-\delta)K_t + (1 - S(X_t))I_t ZI_t; \tag{17}$$

where δ is the depreciation rate, ZI_t is the investment specific shock, $X_t = \frac{I_t}{I_{t-1}}$ and $S()$ satisfies S', $S'' \geq 0$; $S(1+g) = S'(1+g) = 0$.

Demand for capital by firms must satisfy

$$E_t[1 + R_{t+1}] = \frac{E_t \left[F_{K,t} + (1-\delta) Q_{t+1} \right]}{Q_t} \tag{18}$$

In (18) the rhs is the gross return to holding a unit of capital in from t to $t+1$. The lhs is the gross return from holding bonds, the opportunity cost of capital. We complete this set-up with the functional form

$$S(X) = \phi_X (X_t - (1+g))^2 \tag{19}$$

where g is the balanced growth rate.

Owners of physical capital can control the intensity at which capital is utilized in production. As in Christiano *et al.* (2005) and Smets and Wouters (2007) we assume that using the stock of capital with intensity U_t produces a cost of $a(U_t)K_t$ units of the composite final good. The functional form is chosen consistent with the literature:

$$a(U_t) = \gamma_1(U_t - 1) + \frac{\gamma_2}{2}(U_t - 1)^2 \tag{20}$$

and satisfies $a(1) = 0$ and $a'(1), a''(1) > 0$. Note that $\frac{\gamma_1}{\gamma_2} = \frac{1-\phi}{\phi}$ in the Smets and Wouters (2007) set-up. In order to compare results we will estimate ϕ.

2.6 The Household Problem

The Household problem is standard and can be summarized by:

$$\begin{aligned}
\text{Utility} : \Lambda_t &= \Lambda(C_t, L_t) && (21)\\
\text{Euler} : \Lambda_{C,t} &= \beta E_t\left[(1 + R_{t+1})\Lambda_{C,t+1}\right] && (22)\\
\text{Labour Supply} : \quad \frac{\Lambda_{h,t}}{\Lambda_{C,t}} &= -MRS_t \equiv -\frac{W_{h,t}}{P_t} && (23)\\
\text{Leisure} : L_t &\equiv 1 - N_t && (24)
\end{aligned}$$

For later use it is useful to write the Euler consumption equation as

$$1 = E_t\left[(1 + R_{t+1})D_{t,t+1}\right] \tag{25}$$

2.7 Monetary Authority, Aggregation and Equilibrium

Nominal and real interest rates are related by the Fischer equation

$$E_t[1 + R_{t+1}] = E_t\left[\frac{1 + R_{n,t}}{\Pi_{t+1}}\right] \tag{26}$$

where the nominal interest rate is a policy variable, typically given by a simple Taylor-type rule:

$$\log\left(\frac{1 + R_{n,t}}{1 + R_n}\right) = \alpha_R \log\left(\frac{1 + R_{n,t-1}}{1 + R_n}\right) + \alpha_\pi \log\left(\frac{\Pi_t}{\Pi}\right) + \alpha_Y \log\left(\frac{\frac{Y_t}{Y}}{\frac{Yf_t}{Yf}}\right) + \epsilon_{m,t} \tag{27}$$

where, following Smets and Wouters (2007) we define the output gap has the deviation between the output in the economy with nominal rigidities and the output that would prevail in a flexi-price and wage economy. [9]

The resource constraint must take into account relative price dispersion across varieties and wage dispersion across firms. By writing $Y_t(f)^W = F(Z_t, N_t(j), U_t K_{t-1})$.[10] At firm level supply must equal demand:

$$(1-c)F(Z_t, \left(\frac{W_t(j)}{W_t}\right)^{-\mu} N_t, U_t K_{t-1}) = (C_t + I_t + G_t + a(U_t)K_{t-1}) \left(\frac{P_t(m)}{P_t}\right)^{-\zeta} \tag{28}$$

Integrating across all firms, taking into account that the capital-labour ratio is common across firms and that the wholesale sector is separated from the retail sector we obtain

$$(1-c)F(Z_t, \int_0^1 \left(\frac{W_t(j)}{W_t}\right)^{-\mu} dj N_t, U_t K_{t-1}) = (C_t + I_t + G_t + a(U_t)K_{t-1}) \int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\zeta} df \tag{29}$$

where the price dispersion is given by $\Delta_{p,t} = \int_0^1 \left(\frac{P_t(f)}{P_t}\right)^{-\zeta} df$ and wage dispersion is given by $\Delta_{w,t} = \int_0^1 \left(\frac{W_t(j)}{W_t}\right)^{-\mu} dj$. As shown in Appendix B:

$$\Delta_{p,t} = \xi \tilde{\Pi}_t^{\zeta} \Delta_{p,t-1} + (1-\xi) \left(\frac{P_t^0}{P_t}\right)^{-\zeta} \tag{30}$$

$$\Delta_{w,t} = \xi_w \tilde{\Pi}_{w,t}^{\mu} \Delta_{w,t-1} + (1-\xi_w) \left(\frac{W_t^0}{W_t}\right)^{-\mu} \tag{31}$$

$$\tag{32}$$

Then (28) takes the form:

$$Y_t = (1-c)\frac{Y_t^W}{\Delta_{p,t}\Delta_{w,t}} = C_t + I_t + G_t + a(U_t)K_{t-1} \tag{33}$$

2.8 Representation of Price-Wage Dynamics as Difference Equations

In order to set up the model in DYNARE we need to represent the price and wage dynamics as difference equations. Both sides of the foc for pricing (8) and wage (14), are of the form considered in Appendix A. Using the Lemma, inflation dynamics are given by:

[9] Yf is the output of an RBC model with habits, adjustment costs to investment and variable capital utilization.

[10] Where by simplicity we call Z_t a vector containning each type of biased and un-biased technical change defined in (4).

$$\Pi_{p,t} \equiv \frac{P_t}{P_{t-1}} = \pi_t + 1 \tag{34}$$

$$\frac{P_t^0}{P_t} \equiv \frac{J_{p,t}}{H_{p,t}} \tag{35}$$

$$\tilde{\Pi}_{p,t} \equiv \frac{\Pi_{p,t}}{\Pi_{p,t-1}^{\gamma_p}} \tag{36}$$

$$MS_{p,t} \equiv \frac{\zeta}{\zeta - 1} e P_t \tag{37}$$

where we introduced a price mark-up shock $MS_{p,t}$ and by:

$$H_{p,t} - \xi_p \beta E_t[\tilde{\Pi}_{p,t+1}^{\zeta-1} H_{p,t+1}] = Y_t \Lambda_{C,t} \tag{38}$$

$$J_{p,t} - \xi_p \beta E_t[\tilde{\Pi}_{p,t+1}^{\zeta} J_{p,t+1}] = MS_{p,t} Y_t MC_t \Lambda_{C,t} \tag{39}$$

$$1 = \xi_p \tilde{\Pi}_{p,t}^{\zeta-1} + (1 - \xi_p) \left(\frac{J_{p,t}}{H_{p,t}} \right)^{1-\zeta} \tag{40}$$

For staggered wage setting, symmetrically, wage dynamics are given by defining:

$$\Pi_{w,t} \equiv \frac{W_t}{W_{t-1}} \Pi_{p,t} \tag{41}$$

$$\frac{W_t^0}{P_t} \equiv \frac{J_{w,t}}{H_{w,t}} \tag{42}$$

$$MS_{w,t} \equiv \frac{\mu}{\mu - 1} e W_t \tag{43}$$

where we introduced a wage mark-up shock $MS_{w,t}$. Aggregate wage dynamics are then given by

$$H_{w,t} - \xi_w \beta E_t \left[\Pi_{w,t+1}^{\mu} \left(\frac{\Pi_{p,t+1}}{\Pi_{p,t}^{\gamma_w}} \right)^{\mu-1} \right] H_{w,t+1} = N_t \, \Lambda_{C,t} \tag{44}$$

$$J_{w,t} - \xi_w \beta E_t \left[\Pi_{w,t+1}^{\mu} \left(\frac{\Pi_{p,t+1}}{\Pi_{p,t}^{\gamma_w}} \right)^{\mu} \right] = -MS_{w,t} N_t MU_t^N \tag{45}$$

$$\xi_w \left[\Pi_{w,t} \frac{\Pi_{p,t}}{\Pi_{p,t-1}^{\gamma_w}} \right]^{\mu-1} + (1 - \xi_w) \left(\frac{\frac{J_{w,t}}{H_{w,t}}}{\frac{W_t}{P_t}} \right)^{1-\mu} = 1 \tag{46}$$

2.9 The Normalized CES Production Function

The production function is assumed to be CES as in Cantore *et al.* (010a) which nests Cobb-Douglas as a special case and admits the possibility of neutral and non-neutral technical change. Here we adopt the 're- parametrization' procedure described in Cantore and Levine (2010) in order to normalize the CES production function:

$$\begin{aligned} Y_t &= \left[\alpha_k(ZK_tK_t)^{\psi} + \alpha_n(ZN_tN_t)^{\psi}\right]^{\frac{1}{\psi}} \; ; \psi \neq 0 \;\&\; \alpha_k + \alpha_h \neq 1 \\ &= (ZK_tK_t)^{\alpha_k}(ZN_tN_t)^{\alpha_n} \; ; \psi \to 0 \;\&\; \alpha_k + \alpha_h = 1 \end{aligned} \tag{47}$$

where Y_t, K_t, N_t are output, capital and labour inputs respectively at time t and ψ is the substitution parameter and α_k and α_n are sometimes referred as distribution parameters. As explained earlier, the terms ZK_t and ZN_t capture capital-augmenting and labour-augmenting technical progress respectively. Calling σ the elasticity of substitution between capital and labour,[11] with $\sigma \in (0, +\infty)$ and $\psi = \frac{\sigma-1}{\sigma}$ then $\psi \in (-\infty, 1)$. When $\sigma = 0 \Rightarrow \psi = -\infty$ we have the Leontief case, whilst when $\sigma = 1 \Rightarrow \psi = 0$ (47) collapses to the usual Cobb-Douglas case.

From the outset a comment on dimensions would be useful. Technology parameters are not measures of efficiency as they depend on the units of output and inputs (i.e., is not dimensionless[12]) and the problem of normalization arises because unless $\psi \to 0$, α_n and α_k in (47) are not shares and in fact are not dimensionless.

[11] The elasticity of substitution for the case of perfect competition, where all the product is used to remunerate factor of productions, is defined as the elasticity of the capital/labour ratio with respect to the wage/capital rental ratio. Then calling W the wage and $R + \delta$ the rental rate of capital we can define the elasticity as follows:

$$\sigma = \frac{d\frac{K}{L}\frac{L}{K}}{d\frac{W}{R+\delta}\frac{R+\delta}{W}}.$$

See La Grandville (2009) for a more detailed discussion.

[12] For example for the Cobb-Douglas production function in the steady state, $Y = K^{\alpha}(AN)^{1-\alpha}$, by dimensional homogeneity, the dimensions of A are (output per period)$^{\frac{1}{1-\alpha}}$ / ((person hours per period)$\times$ (machine hours per period)$^{\frac{\alpha}{1-\alpha}}$). For some this poses a fundamental problem with the notion of a production function - see Barnett (2004). Units can be chosen so that when $N = 1$ and $K = 1$, then $Y = 1$ implying $A = 1$. For the equilibrium to be independent of the choice of units, it follows that it must be independent of the steady state value A. This is readily demonstrated in what follows.

Marginal products of labour and capital are respectively

$$F_{N,t} = \frac{Y_t}{N_t}\left[\frac{\alpha_n (ZN_tN_t)^{\psi}}{\alpha_k ZK_tK_t^{\psi} + \alpha_n N_t^{\psi}}\right] = \alpha_n ZN_t^{\psi}\left(\frac{Y_t}{N_t}\right)^{1-\psi} \tag{48}$$

$$F_{K,t} = \frac{Y_t}{K_t}\left[\frac{\alpha_k ZK_tK_t^{\psi}}{\alpha_k ZK_tK_t^{\psi} + \alpha_n (ZN_tN_t)^{\psi}}\right] = \alpha_k ZK_t^{\psi}\left(\frac{Y_t}{K_t}\right)^{1-\psi} \tag{49}$$

The equilibrium of real variables depends on parameters defining the RBC core of the model ϱ, σ_c, δ, ψ, α_k and α_n, and those defining the NK features. Of the former, ϱ, ψ and σ_c are dimensionless, δ depends on the unit of time, but unless $\psi = 0$ and the technology is Cobb-Douglas, α_k and α_n depend on the units chosen for factor inputs, namely machine units per period and labour units per period. To see this rewrite the wholesale firm's foc (5) and (6) in terms of factor shares

$$\frac{W_tN_t}{P_t^W Y_t} = \alpha_n ZN_t^{\psi}\left(\frac{Y_t}{N_t}\right)^{-\psi} \tag{50}$$

$$\frac{(R_t+\delta)K_t}{P_t^W Y_t} = \alpha_k ZK_t^{\psi}\left(\frac{Y_t}{K_t}\right)^{-\psi} \tag{51}$$

where $P_t^W \equiv MC_tP_t$ is the price of wholesale output. Then we have

$$\frac{W_tN_t}{(R_t+\delta)} = \frac{\alpha_n}{\alpha_k}\left(\frac{ZK_tK_t}{ZN_tN_t}\right)^{-\psi} \tag{52}$$

Thus α_n (α_k) can be interpreted as the share of labour (capital) iff $\psi = 0$ and the production function is Cobb-Douglas. Otherwise the dimensions of α_k and α_n depend on those for $\left(\frac{ZK_tK_t}{ZN_tN_t}\right)^{\psi}$ which could be for example, (effective machine hours per effective person hours)$^{\psi}$. In our aggregate production functions we choose to avoid specifying unit of capital, labour and output.[13] It is impossible to interpret and therefore to calibrate or estimate these 'share' parameters.

There are two ways to resolve this problem; 're-parameterize' the dimensional parameters α_k and α_n so that they are expressed in terms of dimensionless ones all parameters to be estimated or calibrated (see Cantore and Levine (2010)), or 'normalize' the production

[13] Unlike the physical sciences where particular units are explicitly chosen so dimension-dependent parameters pose no problems. For example the fundamental constants such as the speed of light is expressed in terms of metres per second; Newtons constant of gravitation has units cubic metres per (kilogram $\times$ second^2) etc.

function in terms of deviations from a steady state. We consider these in turn.

2.9.1 Re-parametrization of α_n and α_k

On the balanced-growth path (bgp) consumption, output, investment, capital stock, the real wage and government spending are growing at a common growth rate g driven by exogenous labour-augmenting technical change $\overline{ZN}_{t+1} = (1+g)\overline{ZN}_t$, but labour input N is constant.[14] As is well-known a bgp requires either Cobb-Douglas technology or that technical change must be driven solely by the labour-augmenting variety (see, for example, Jones (2005)). Then $ZK_t = ZK$ must also be constant along the bgp.

On the bgp let capital share and wage shares in the wholesale sector be α and $1-\alpha$ respectively. Then using (50) and (51) we obtain our *re-parameterization* of α_n and α_k:

$$\alpha_k = \alpha \left(\frac{\bar{Y}_t}{ZK\bar{K}_t} \right)^{\psi} \tag{53}$$

$$\alpha_n = (1-\alpha) \left(\frac{\bar{Y}_t}{\overline{ZN}_t N} \right)^{\psi} \tag{54}$$

Note that $\alpha_k = \alpha$ and $\alpha_n = 1-\alpha$ at $\psi = 0$, the Cobb-Douglas case. [15] This completes the stationarized equilibrium defined in terms of *dimensionless* RBC core parameters ϱ, σ_c, ψ,π and δ which depends on the unit of time, plus NK parameters. In (53) and (54) dimensional parameters are expressed in terms of other endogenous variables Y, N and K which themselves are functions of $\theta \equiv [\sigma, \psi, \pi, \delta, \cdots]$. Therefore $\alpha_n = \alpha_n(\theta)$, and $\alpha_k = \alpha_k(\theta)$ which expresses why we refer to this procedure as reparameterization.

There is one more normalization issue: the *choice of units* at some point say $t = 0$ on the steady state bgp. We use for simplicity $\bar{Y}_0 = \overline{ZN}_0 = ZK = 1$ but, as it is straightforward to show that having expressed the model in terms of dimensionless parameters through re-parameterization makes the steady state ratios of the endogenous variables of the model independent of this choice.

[14] If output, consumption etc are defined in per capita terms then N can be considers as the proportion of the available time at work and is therefore both stationary and dimensionless.

[15] And as argued before if $\pi \in (0,1)$ $\alpha_k + \alpha_n = 1$ iff $\psi = 0$.

2.9.2 The Production Function in Deviation Form

This simply bypasses the need to retain α_k and α_n and writes the dynamic production function in deviation form about its steady state as

$$\frac{Y_t}{\bar{Y}_t} = \left[\frac{\alpha_k(ZK_tK_t)^\psi + \alpha_n(ZN_tN_t)^\psi}{\alpha_k(ZK\bar{K}_t)^\psi + \alpha_n(\overline{ZN}_tN)^\psi}\right]^{\frac{1}{\psi}} = \left[\frac{\alpha_k\left(\frac{ZK_tK_t}{ZK\bar{K}_t}\right)^\psi}{\alpha_k + \alpha_n\left(\frac{\overline{ZN}_tN}{ZK\bar{K}_t}\right)^\psi} + \frac{\alpha_n\left(\frac{ZN_tN_t}{\overline{ZN}_tN_t}\right)^\psi}{\alpha_k\left(\frac{ZK\bar{K}_t}{\overline{ZN}_tN}\right)^\psi + \alpha_n}\right]^{\frac{1}{\psi}}$$

Then from (53) and (54) we can write this simply as

$$\frac{Y_t}{\bar{Y}_t} = \left[(1-\alpha)\left(\frac{K_t}{\bar{K}_t}\right)^\psi + \alpha\left(\frac{ZN_tN_t}{\bar{A}_tN}\right)^\psi\right]^{\frac{1}{\psi}} \tag{55}$$

as in Cantore *et al.* (010a). The steady-state normalization now consists of $\overline{ZN}_0 = \bar{Y}_0 = ZK = 1$[16] and is characterized entirely by fixed shares of consumption, investment and government spending and by labour supply as a proportion of available time, all dimensionless quantities apart from the unit of time.

Using either of these two approaches, as showed by Cantore and Levine (2010), the steady state ratios of the endogenous variables and the dynamics of the model are not affected by the starting values of output and the two source of shocks ($\bar{Y}_0, \overline{ZN}_0$, ZK) which only represent choice of units. Crucially, this implies also that changing σ does not change our steady state ratios and factor shares, impulse response functions are directly comparable, and parameter values are consistent with their economic interpretation.

2.10 Utility Function

The household utility function is chosen to be compatible with a balanced-growth steady state and allows external habit-formation:

$$\Lambda_t = \frac{eB_t((C_t - \chi C_{t-1})^{(1-\varrho)}(1-N_t)^\varrho)^{1-\sigma_c} - 1}{1-\sigma_c} \tag{56}$$

$$\Lambda_{C,t} = eB_t(1-\varrho)(C_t - \chi C_{t-1})^{(1-\varrho)(1-\sigma_c)-1}((1-N_t)^{\varrho(1-\sigma_c)}) \tag{57}$$

$$\Lambda_{N,t} = -eB_t\varrho(C_t - \chi C_{t-1})^{(1-\varrho)(1-\sigma_c)}(1-N_t)^{\varrho(1-\sigma_c)-1} \tag{58}$$

[16] Which is almost identical to the one used in Cantore *et al.* (010a) although they normalize as well hours worked to 1 using the accounting identity $\bar{Y} = (\bar{R}+\delta)\bar{K} + \bar{W}\bar{N}$.

Where χ represents the habit formation parameter and eB_t a preference shock.

2.11 Shocks

To close the model we need to specify the law of motion of the shocks

$$\log ZK_t - \log ZK = \rho_{ZK}(\log ZK_{t-1} - \log ZK) + \epsilon_{ZK,t} \quad (59)$$

$$\log ZN_t - \log \overline{ZN}_t = \rho_{ZN}(\log ZN_{t-1} - \log \overline{ZN}_t) + \epsilon_{ZN,t} \quad (60)$$

$$\log ZI_t - \log ZI = \rho_{ZI}(\log ZI_{t-1} - \log ZI) + \epsilon_{ZI,t} \quad (61)$$

$$\log G_t - \log \overline{G}_t = \rho_G(G_{t-1} - \overline{G}_t) + \epsilon_{G,t} \quad (62)$$

$$\log eP_t - \log eP = \rho_P(eP_{t-1} - eP) + \epsilon_{P,t} \quad (63)$$

$$\log eW_t - \log eW = \rho_W(eW_{t-1} - eW) + \epsilon_{W,t} \quad (64)$$

$$\log eB_t - \log eB = \rho_W(eB_{t-1} - eB) + \epsilon_{B,t} \quad (65)$$

In total the model has these 7 AR(1) shocks plus the shock to the monetary policy rule.

3 Estimation

We estimate the linear version of the model around zero steady state inflation by Bayesian methods using DYNARE. We use the same data set as in Smets and Wouters (2007) in first difference at quarterly frequency. Namely, these observable variables are the log differences of real GDP, real consumption, real investment and real wage, log hours worked, the log difference of the GDP deflator and the federal funds rate. As in Smets and Wouters (2007), hours worked are derived from the index of average hours for the non-farm business sector and we divide hourly compensation from the same sector by the GDP price deflator to obtain the real wage. All series are seasonally adjusted. All data are taken from the FRED Database available through the Federal Reserve Bank of St.Louis and the US Bureau of Labour Statistics. The sample period is 1984:1-2004:4. A full description of the data used can be found in Smets and Wouters (2007).

The corresponding measurement equations for the 7 observables are:

$$dy = Y_t - Y_{t-1} + ctrend \tag{66}$$

$$dc = C_t - C_{t-1} + ctrend \tag{67}$$

$$di = I_t - I_{t-1} + ctrend \tag{68}$$

$$dw = W_t - W_{t-1} + ctrend \tag{69}$$

$$\Pi_{obs} = \Pi_t + conspie \tag{70}$$

$$R_{obs} = Rn_t + consrn \tag{71}$$

$$h_{obs} = N_t + conslab \tag{72}$$

Where we introduce a common trend to the real variables and a specific one to inflation, nominal interest rate and hours worked.

3.1 Bayesian Methods

Bayesian estimation entails obtaining the posterior distribution of the model's parameters, say θ, conditional on the data. Using the Bayes' theorem, the posterior distribution is obtained as:

$$p(\theta|Y^T) = \frac{L(Y^T|\theta)p(\theta)}{\int L(Y^T|\theta)p(\theta)d\theta} \tag{73}$$

where $p(\theta)$ denotes the prior density of the parameter vector θ, $L(Y^T|\theta/)$ is the likelihood of the sample Y^T with T observations (evaluated with the Kalman filter) and $\int L(Y^T|\theta)p(\theta)d\theta$ is the marginal likelihood. Since there is no closed form analytical expression for the posterior, this must be simulated.

One of the main advantages of adopting a Bayesian approach is that it facilitates a formal comparison of different models through their posterior marginal likelihoods, computed using the Geweke (1999) modified harmonic-mean estimator. For a given model $m_i \in M$ and common data set, the marginal likelihood is obtained by integrating out vector θ,

$$L\left(Y^T|m_i\right) = \int_{\Theta} L\left(Y^T|\theta, m_i\right) p\left(\theta|m_i\right) d\theta \tag{74}$$

where $p_i\left(\theta|m_i\right)$ is the prior density for model m_i, and $L\left(Y^T|m_i\right)$ is the data density for model m_i given parameter vector θ. To compare models (say, m_i and m_j) we calculate

the posterior odds ratio which is the ratio of their posterior model probabilities (or Bayes Factor when the prior odds ratio, $\frac{p(m_i)}{p(m_j)}$, is set to unity):

$$PO_{i,j} = \frac{p(m_i|Y^T)}{p(m_j|Y^T)} = \frac{L(Y^T|m_i)p(m_i)}{L(Y^T|m_j)p(m_j)} \tag{75}$$

$$BF_{i,j} = \frac{L(Y^T|m_i)}{L(Y^T|m_j)} = \frac{\exp(LL(Y^T|m_i))}{\exp(LL(Y^T|m_j))} \tag{76}$$

in terms of the log-likelihood. Components (75) and (76) provide a framework for comparing alternative and potentially misspecified models based on their marginal likelihood. Such comparisons are important in the assessment of rival models, as the model which attains the highest odds outperforms its rivals and is therefore favoured.

Given Bayes factors, we can easily compute the model probabilities $p_1, p_2, \cdots, p_n$ for n models. Since $\sum_{i=1}^{n} p_i = 1$ we have that $\frac{1}{p_1} = \sum_{i=2}^{n} BF_{i,1}$, from which p_1 is obtained. Then $p_i = p_1 BF(i,1)$ gives the remaining model probabilities.

3.2 Likelihood Comparison of Models

We compare 4 different model specifications in order to see if the introduction of factor substitutability and/or the biased technical change improves the fit of the estimation. In the first row of Table (1) we present the likelihood density of the model with the CD production function where only the labour-augmenting technology shock is present. In the second row we introduce the CES and calibrate the elasticity of substitution to 0.4, following the literature as in Cantore *et al.* (010a), and introduce the capital-augmenting shock whilst in rows 3 and 4 we estimate σ in a model with and without the latter shock. Strictly speaking a meaningful likelihood comparison that provides information about σ is only possible between row 1 and 3 (where we can compare like for like).

Table 1 reveals that Models with the CES production function clearly outperforms its CD counterpart with a posterior probability of 100%. This suggests that incorporating a CES production function offers substantial improvements in terms of the model fitness to the data in the US economy. The differences in log marginal likelihood are substantial. For example, the log marginal likelihood difference between the first two specifications is 10.59 corresponding to a posterior Bayes Factor of 33735. As suggested by Kass and Raftery (1995), the posterior Bayes Factor needs to be at least $e^3 \approx 20$ for there to be a

positive evidence favouring one model over the other.

Model Notation	σ	Technology shocks	Log data density	Difference with CD
CD	1	ZL	-469.13	0
–	calibrated=0.4	ZK & ZL	-458.54	10.58
CES1	estimated=0.33	ZL	-459.23	9.89
CES2	estimated=0.36	ZK & ZL	-460.24	8.88

Table 1: Marginal Likelihood comparison between CD and CES specifications

3.3 Estimation Results

The joint posterior distribution of the estimated parameters is obtained in two steps. First, the posterior mode and the Hessian matrix are obtained via standard numerical optimization routines. The Hessian matrix is then used in the Metropolis-Hastings (MH) algorithm to generate a sample from the posterior distribution. Two parallel chains are used in the Monte-Carlo Markov Chain Metropolis-Hastings (MCMC-MH) algorithm. Thus, 250,000 random draws (though the first 30% 'burn-in' observations are discarded) from the posterior density are obtained via the MCMC-MH algorithm, with the variance-covariance matrix of the perturbation term in the algorithm being adjusted in order to obtain reasonable acceptance rates (between 20%-40%).

Estimation results from posteriors maximization are presented in Appendix C. We used the same priors as Smets and Wouters (2007) for common parameters whereas we used a loose prior for the elasticity of substitution between capital and labour in order to see if the data are informative about the value of this parameter. A few structural parameters are kept fixed in the estimation procedure, in accordance with the usual practice in the literature (see Table 2). This is done so that the calibrated parameters reflect steady state values of the observed variables.

First we focus on the posterior estimates obtained using the most general CES model, CES2. As shown in Tables 3 and 4, the point estimates under the CES assumption are tight and plausible. In particular, focusing on the parameters characterizing the degree of price stickiness and the existence of real rigidities, we find that the price indexation parameters are estimated to be smaller than assumed in the prior distribution (in line with those reported by Smets and Wouters (2007)). The estimates of the $\gamma's$ imply that inflation is intrinsically not very persistent in the CES model specifications. The posterior mean

Calibrated parameter	Symbol	Value
Discount factor	β	0.99
Depreciation rate	δ	0.025
Growth rate	g	$\frac{\delta}{4}$
Substitution elasticity of goods	ζ	7
Substitution elasticity of labour	μ	7
Variable capital utilization	γ_1	$\frac{1}{\beta} + \delta - 1$
Fixed cost	c	$1 - MC = 0.1429$
Implied steady state relationship		
Government expenditure-output ratio	g_y	0.2
Consumption-output ratio	c_y	0.6408
Investment-output ratio	i_y	$1 - g_y - i_y$

Table 2: Calibrated Parameters

estimates for the Calvo parameters, ξ_p and ξ_w, imply an average price contract duration of around 2.56 quarters and an average wage contract duration of around 2.04 quarters, respectively. These results are in general consistent with the findings from empirical works on the DSGE modelling in the US economy. It is interesting to note that the risk-aversion parameter (σ_c) is estimated to be greater than assumed in the prior distribution, indicating that the inter-temporal elasticity of substitution (proportional to $1/\sigma_c$) is estimated to be about 0.47 in the US, which is plausible as suggested in much of RBC literature. As expected, the policy rule estimates imply a fairly strong response (α_π) to expected inflation by the US Fed Reserve and the degree of interest rate smoothing (α_r) is fairly strong.

Figure 4 in Appendix C plots the prior and posterior distributions for the above CES model. The location and the shape of the posterior distributions are largely independent of the priors we have selected since priors are broadly less informative. Most of the posterior distributions are roughly symmetric implying that the mean and median coincide. According to Figure 4, there is little information in the data for some parameters where prior and posterior overlap. Perhaps the most notable finding comes from the estimation of the parameter σ - our key parameter in the CES setting. As a result of assuming a very diffuse prior with large standard deviation, we find that the data is very informative about this parameter (as clearly shown in the figure, curves do not overlap each other and are very different) and the point estimate of σ in Table 4 is close to the plausible values. This further provides strong evidence to support the empirical importance of the CES assumption.

We now turn on the comparisons between paremeters estimates under CD and CES specification. Parameter posteriors that are quantitatively different[17] from the estimation using a Cobb-Douglas specification are underlined in Tables 5 and 6.

Starting with the parameters related to the exogenous shocks (Table 5) we notice that the estimated standard deviations of the newly introduced capital-augmenting technology shock is very small but, probably because of its introduction, the standard deviation of the investment specific shock reduces significantly (from 3.67 in the CD specification to 2.45 in the CES case). The two shocks are clearly related and it is very likely that when ZK is absent ZI is capturing "capital-biased" technological progress. We also notice that the estimated standard deviation of the price mark-up shock is 0.07 higher under the CES specification whereas the standard deviation of the preference shock is 0.1 lower. The autoregressive parameters of the exogenous shocks are not affected significantly by the CES choice.

Posterior estimation of the investment adjustment costs parameter (ϕ^X) reduces by 1.20 points when we estimate the model under CES showing once again how introducing factor-biased technical change affects significantly the estimation of 'investment-related' parameters. The parameters of the utility function also appear to be affected by the choice of the production function (σ_c increases by 0.13, ϱ reduces by 0.23 and χ reduces by 0.11). Regarding the parameters associated with sticky prices and wages only the probability of no price-adjustment (ξ_p) changes significantly, decreasing from 0.72 to 0.61. Monetary policy weights (except the weight on inflation which increases slightly), real and nominal trends estimations are not affected by the introduction of factor substitutability and biased technical change.

4 Model Validation

After having shown the model estimates and the assessment of relative model fit to its other rivals with different restrictions, we use them to investigate a number of key macroeconomic issues in the US. The model favoured in the space of competing models may still be poor (potentially misspecified) in capturing the important dynamics in the data. To further evaluate the absolute performance of one particular model against data, it is nec-

[17] Difference in posteriors up to 0.05 were not considered quantitatively relevant here.

essary to compare the model's implied characteristics with those of the actual data (and an identified VAR model).

In this section, we address the following questions: (i) can the models capture the underlying characteristics of the actual data? (ii) what are the impacts of the structural shocks on the main macroeconomic time series?

4.1 Standard Moment Criteria

Summary statistics such as first and second moments have been standard as means of validating models in the literature on DSGE models, especially in the RBC tradition. As the Bayes factors (or posterior model odds) are used to assess the relative fit amongst a number of competing models, the question of comparing the moments is whether the models correctly predict population moments, such as the variables' volatility or their correlation, i.e. to assess the absolute fit of a model to macroeconomic data.

To assess the contributions of assuming different specifications of production function in our estimated models, we compute some selected second moments and present the results in this subsection. Table 3 presents the (unconditional) second moments implied by the above estimations and compares with those in the actual data. In particular, we compute these model-implied statistics by solving the models at the posterior means obtained from estimation. The results of the model's second moments are compared with the second moments in the actual data to evaluate the models' empirical performance.

			Standard Deviation					
Model	Output	Consumption	Investment	Real wage	Inflation	Interest rate	Employment	Labour share
Data	0.59	0.53	1.80	0.60	0.25	0.64	2.47	3.82
Model CD	0.93	0.66	2.15	0.65	**0.37**	0.43	5.56	0
Model CES1	**0.82**	**0.56**	**1.76**	**0.59**	0.47	**0.54**	**5.51**	**1.86**
Model CES2	**0.82**	**0.56**	1.78	**0.59**	0.46	0.53	5.58	1.83
			Cross-correlation with Output					
Data	1.00	0.60	0.63	-0.10	-0.14	0.14	-0.22	0.03
Model CD	1.00	**0.51**	0.80	0.33	**-0.20**	-0.30	0.13	-
Model CES1	1.00	0.43	**0.73**	**-0.05**	0.08	**-0.07**	**0.10**	**-0.31**
Model CES2	1.00	0.44	0.74	-0.04	0.09	**-0.07**	**0.10**	**-0.31**
			Autocorrelations (Order=1)					
Data	0.34	-0.03	0.49	-0.04	0.75	0.96	0.97	0.12
Model CD	0.46	0.40	0.68	0.42	0.72	0.91	**0.98**	-
Model CES1	**0.29**	**0.27**	**0.53**	**0.36**	0.77	**0.94**	**0.98**	**0.97**
Model CES2	**0.29**	0.28	0.54	**0.36**	**0.76**	**0.94**	**0.98**	**0.97**

Table 3: Selected Second Moments of the Model Variants

In terms of the standard deviations, Model CD generates relatively high volatility compared to the actual data (except for the interest rate and the CD production assumes constant labour share). Overall, the estimated models are able to reproduce broadly acceptable volatility for the main variables of the DSGE model and all model variants can successful replicate the stylized fact in the business cycle research that investment is more volatile than output whereas consumption is less volatile. In line with the Bayesian model comparison, the NK models with CES technology fit the data better in terms of implied volatility, getting closer to the data in this dimension (we highlight the 'best' model (performance) in bold). Note that both of our CES models clearly outperform the CD model in capturing the volatilities of all variables except for inflation and does extremely well at matching the consumption, investment and real wage volatilities in the data. Furthermore by not imposing a constant labour share as in the CD model we are capable of capturing about half the volatility observed in the data. As suggested by the likelihood comparison, the differences in generating the moments between the CES specification with only the shock ZK and the CES with both ZK and ZL shocks are qualitatively very small.

Table 3 also reports the cross-correlations of the eight observable variables vis-a-vis output. All models perform successfully in generating the positive contemporaneous correlations of consumption and investment observed in the data. The data report that the real wage is countercyclical. Both our CES models, perform successfully in generating the negative contemporaneous real wage-output correlation observed in the data, suggesting that assuming CES production helps to capture this labour market dimension. The highlighted numbers in this category together with the evidence above show that the feature of CES in the model is particularly important in characterizing the investment and real wage dynamics. However, as evidence from the implied volatilities confirms, the main shortcoming of all the models, including the preferred ones, is the difficulty at replicating the cross-correlations of output with employment and mimicking the volatility observed in the employment data. This is not a very surprising result because there are no labour market frictions assumed in all the models under investigation. All models fail to predict the positive correlation between output and interest rate and CES models have problem in replicating the negative contemporaneous cross-relation between inflation and output.

This is consistent with the work of Smets and Wouters (2003) as they find that the implied cross-correlations with the interest rate and inflation are not fully satisfactory. Nevertheless, looking at the highlighted figures, the results in general show that, in the models where the CES specification is present, cross-correlations of endogenous variables are generally closer to those in the actual data. This further strengthens the empirical relevance of the CES assumption.

All models appear to match well the autocorrelations (order=1) of the endogenous variables output, investment, inflation, interest rate and employment. Using the CES model, output is less autocorrelated at order 1, whilst investment and inflation are more autocorrelated than those in the data.

To summarize this sub-section, overall Bayesian Maximum-likelihood based methods suggest that the ability of the model's second moments to fit those of the data generally match the outcome of the likelihood race. The two CES models deliver a better fit to the actual data for most of the second moment features in Table 3. However, as noted above, the differences in the second moments of the two competing CES variants are very small.

4.2 Autocorrelation Functions

We have so far considered autocorrelation only up to order 1. To further illustrate how the estimated models capture the data statistics and persistence in particular, we now plot the autocorrelations up to order 10 of the actual data and those of the endogenous variables generated by the model variants in Figure 3.

Of particular interest is that, when assuming CES production, the implied autocorrelograms fit very well the observed autocorrelation of inflation, interest rate and investment, whilst the CD model generates much less sluggishness and is less able to match the autocorrelation of inflation and interest rate observed in the data from the second lag onwards. Overall we find that, with nominal price stickiness in the models and highly correlated estimated price markup shocks, inflation persistence can be captured closely in DSGE models when CES production is assumed.

When it comes to output, all models perform well in matching the observed output persistence. However, the employment (hours worked) is more autocorrelated in all models than in the data, but now the model with the CD feature gets closer to the data for

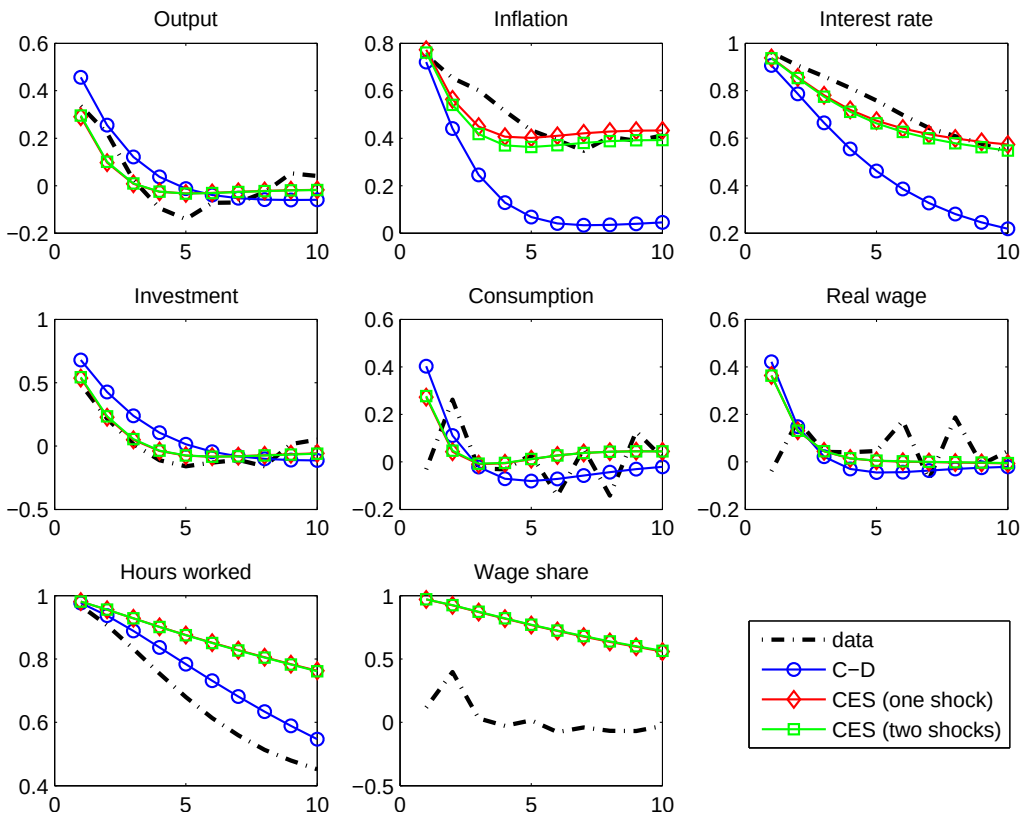


Figure 3: Autocorrelations of Observables in the Actual Data and in the Estimated Models

higher order autocorrelation. All models match reasonably well the autocorrelations of investment, consumption and real wage. To summarize, the results for higher order autocorrelations for the most part show that the DSGE models under the more general CES production function are better at capturing the main features of the US data, strengthening the argument that the assumption of CES helps to improve the model fit to data.

4.3 Comparison with a DSGE-VAR

An alternative way of validating the model performance is to follow Del Negro and Schorfheide (2004) and Del Negro *et al.* (2007) and to compare it with a hybrid model that is a combination of an unrestricted VAR and the VAR implied by the estimated DSGE model. We then go on to compare the estimated DSGE model this 'DSGE-VAR' in terms of their impulse response functions (IRFs). We also investigate the impact on IRFs of changing the production function. Since we have demonstrated that there is little

difference between the CES variants in terms of matching the data, this exercise is only performed for the best CES (one-shock) model.

The DSGE-VAR approach uses DSGE model itself to construct a prior distribution for the VAR coefficients so that DSGE-VAR estimates are tilted toward DSGE model restriction, thus identifying the shocks for the IRFs. This method constructs the DSGE prior by generating dummy observations from the DSGE model, and adding them to the actual data and leads to an estimation of the VAR based on a mixed sample of artificial and actual observations. The ratio of dummy over actual observations (called the hyper-parameter λ) controls the variance and therefore the weight of the DSGE prior relative to the sample. For extreme values of this parameter (0 or ∞) either an unrestricted VAR or the DSGE is estimated. If λ is small the prior is diffuse. When $\lambda = \infty$, we obtain a VAR approximation[18] of the log-linearized DSGE model. As λ becomes small the cross-equation restrictions impled by the DSGE model are gradually relaxed. The empirical performance of a DSGE-VAR will depend on the tightness of the DSGE prior. Details on the algorithm used to implement this DSGE-VAR are to be found in Del Negro and Schorfheide (2004) and Del Negro *et al.* (2007).

We fit our VAR to the same data set used to estimate the DSGE model. We consider a VAR with 4 lags.[19] We use a data-driven procedure to determine the tightness of prior endogenously based on the marginal data density. Our choice of the optimal λ is 1.1 and this is found by comparing different VAR models using the estimates of the marginal data density (Figure 1). In particular, we iterate over a grid that contains the values of $\lambda = [0.43; 0.8; 1; 1.1; 1.2; 1.4; 1.5; 1.6; 2; 5; 10; \infty]$, we find that $\lambda = 1.1$ has the highest posterior probability for all models. Note that 0.43 is the smallest λ value for which we have a proper prior. Overall, the DSGE-VAR(4) with $\lambda = 1.1$ has the highest posterior probability.[20] This implies that the mixed sample that is used to estimate the VAR has a higher weight on the DSGE model (artificial observations) than on the VAR (actual observations). $\hat{\lambda}$ represents how much the economic model (DSGE) is able to explain the real data. More importantly, the results from comparing across different models show that

[18]The accuracy of the approximation depends on the invertibility of the DSGE model's moving average components and on the number of autoregressive lags included (Del Negro *et al.* (2007)).

[19]This choice of the lag length maximizes the marginal data density associated with the DSGE-VAR($\hat{\lambda}$).

[20]Alternatively, one can simply find the 'optimal' $\hat{\lambda}$ by estimating the parameter λ as one of the deep parameters (see, for more details, Adjemian *et al.* (2008)).

Models CES consistently outperform Model CD and CD is strongly rejected in favour of CES when the weight on the DSGE model becomes higher(when λ tends to ∞). .

The improved performance of the CES models over the CD applies at the optimum λ, but the log of the marginal likelihood difference (LL) is only 1. This implies a Bayes factor of $\exp(1) \simeq 2.73$ favouring the CES models. Beyond the optimum the LL far more rapidly for the CD model reaching a difference close to that for the actual linearized DSGE model reported earlier. Overall the LL plots then confirm the fact and the degree to which the CES models are less misspecified.

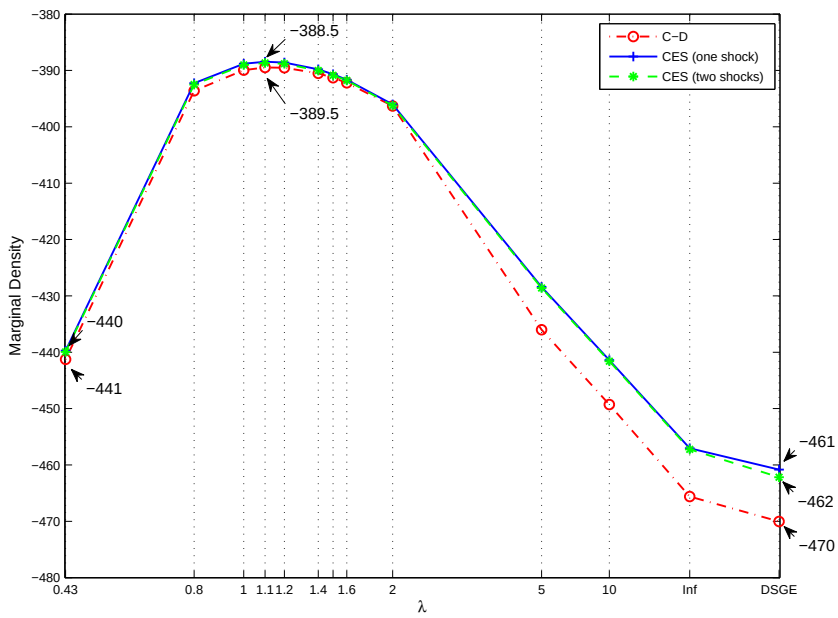


Figure 4: Marginal Likelihood as a Function of λ

Turning to the impulse responses, Figures 6–13 in Appendix C depict the mean responses corresponding to a positive one standard deviation shock. The endogenous variables of interest are the observables in the estimation and each response is for a 20 period (5 years) horizon. All DSGE impulse responses are computed simulating the vector of DSGE model parameters at the posterior mean values reported in Tables 3 and 4. The impulse responses for VAR(4) are obtained using the DSGE-VAR identification procedure using the best-fitting DSGE-VAR(4) with $\lambda = 1.1$. The surface between the black dashed lines in each panel covers the space between the first and ninth posterior deciles of the

VAR's IRFs.

Overall, we find that the sign and magnitude of the DSGE and VAR impulse responses are quite similar implying that the DSGE model mimics the DSGE-VAR model in, at least, some dimensions. Clearly the most important difference comes from fluctuations in factor shares under the CES specification. Fluctuations of shares translate as well in different IRFs of interest rate and wage in the two models. Indeed in Figure 6 (labour-augmenting shock) we can see how the response of the nominal interest rate is significantly different. Figure 8 (investment shock) turns out be very interesting, given the highlighted difference in the estimation of parameters related to investment in section 3.3. We note how both, wage and interest rate, present a more sluggish response to an investment specific shock under CES and, as a result, a quite different response of consumption and inflation. Although they have similar shapes, the IRFs under CES are quantitatively different. In particular, the shock amplifies the initial responses of some variables. The disagreement in the IRFs to this particular shock can be explained by the large estimate of the shock standard deviation reported in Table 3.

If we look closely at the responses to monetary policy shock (Figure 10), we find that both in the models and in the data, consumption, investment and output display a hump-shaped response to the policy shock. For this shock the IRFs from both DSGE models are in agreement with the data and when comparing the performance from CES and CD the difference is quantitatively small. Turning to the IRFs to productivity shock (labour-augmenting), using both the CES and CD settings is able to provide reasonable responses. More importantly, if we compare each response from each DSGE model with that from the VAR model, we find that overall the discrepancy between VAR and DSGE is relatively smaller under the CES production assumption. This suggests that the DSGE model misspecification is larger with the CD production than with the CES. If we study carefully the responses to the other shocks, we can generally find the similar conclusion that CES helps reduce the discrepancy although the IRFs to the investment-specific shock are the exception. In Figure 10 we can see how the reaction of wage and interest rate are once again very different after a government spending shock. Indeed it the introduction of CES specification increases the magnitude of the responses a lot. As a result of those differences in the dynamics of factor prices we notice how investment, consumption and

inflation also present an increase in the amplification of the government spending shock. For price and wage mark-up shock (Figures 12 and 13) we notice non-negligible differences in the responses of interest rate, inflation and hours worked.

To sum up, there also exists some evidence from IRFs in favour of the CES assumption in DSGE models, but the evidence from the IRFs is not as strong as that obtained by comparing the moments and the marginal likelihood comparison amongst models which more clearly reject the CD specification.

5 Variance Decomposition of Business Cycle Fluctuations

This section investigates the contribution of each of the structural shocks to the forecast error variance of the observable variables in the models, i.e. the underlying sources of fluctuations, at various horizons. The results are based on the models' posterior distribution reported in Tables 3 and 4. The results are summarized in Figures 14 and table 6 in Appendix D.

In the short run, within a year (t=1,4), movements in real GDP are primarily driven by the exogenous government spending shock and supply-side shocks (with the dominant influence of over 70%). For instance, most of the unexpected output fluctuations are mainly explained by the government spending shock (around 30%) and the two mark-up shocks (around 20% from each shock). Within the one-year horizon, the government spending shock dominate, accounting for the biggest part of the output forecast error variance.

Not surprisingly, in the medium to long run the supply shocks and the exogenous spending shock together continue to dominate, but the contribution of government spending shock to output variability become smaller from medium to long run and the wage mark-up shock explain a bigger part of the long-run variations in output. This is especially the case when the model adopts the CES function form. In contrast, the monetary policy shock and preference shock have little impact on for output variability, regardless of forecast horizon. Based on the estimation sample, the investment shock and labour-augmenting productivity shock are found to be moderate factors behind both short-run and longer-run movements in output (account for around 11%-13% and around 7%-8%, respectively). In terms of determining the main driving forces of output, we compare all

three specifications under investigation and find similar and consistent results. Results from the CES model with two technology shocks show that, in line with its estimated standard deviation, the capital-augmenting technology shock offers a qualitatively very small impact to the output fluctuations.

Under the estimated interest rate rule we find that the monetary policy shock is by far the most determinant influence to the nominal interest rate in the short run (1 quarter), which explains around over 50% of its variance under the assumption of CES technology. The second largest component is the investment shock. In fact, this shock starts to dominate from the medium to long run and the contributions of policy shock declines quite sharply toward the longer horizon. However, models with CES specification tells a slightly different story. They show that the main driving factor in the long-run development of nominal interest rate is wage mark-up shock. As expected, the preference shock explains a big part of consumption variation in the short-medium run, whilst the investment shock contributes the largest fraction of investment movements in the short run (within a year). In terms of explaining the consumption and investment fluctuations, we do not find notable differences whether CD or CES is assumed.

Interestingly, the CD model suggests that the shock that explains most of inflation variance in the short run is the price mark-up shock but the investment shock becomes more influential from medium to long run (around 25%). In contrast, our estimated CES models show that inflation fluctuations are mostly affected by the investment shock in the short and medium runs but the the main driving factor becomes the wage mark-up shock, dominates the investment shock, in the long run. There are only limited effects on inflation from the productivity shocks and various demand shocks. One possible reason for this is, according to Smets and Wouters (2007), that the estimated slope of the NK Phillips curve is small so that only large and persistent changes in the marginal cost will have an impact on inflation. Finally, the short to medium run contributions of the selected shocks to the forecast error variance of hours worked are broadly similar across the three models. In the long run, there are different results between the CES and CD assumptions. In particular, the model adopting the CD production suggests that the two mark-up shocks both contribute significantly to the variation in hours worked whereas, when we use the more general CES setting in our DSGE model, the wage mark-up shock clearly becomes

the completely dominant force behind the long-run movements in hour worked from the mid-80's onwards. This finding from the CES model seems to be more plausible.

Overall, the results in this exercise mainly show that, over the sample period, the supply-side shocks account for much of the medium to long-run variance which is in line with the business cycle literature and identified VAR studies in industrialized economies. The disturbances from government expenditures are also important at explaining the dynamics of macro-variables in the US economy.

6 Conclusions

This paper contributes to a rapidly rising literature that brings the CES specification of the production function into the analysis of business cycle fluctuations. The main result is to confirm decisively the importance of CES rather than CD production functions. Indeed in a marginal likelihood race our estimated best CES model with an elasticity well below unity at 0.36 beats the CD production function by a substantial log-likelihood of 10.58. Assuming equal prior model probabilities, this implies that posterior model probabilities are 0.999975 : 0.000025 in favour of the CES! The reason for this for result is that movements of factor shares with the CES specification help substantially to fit the data. The marginal likelihood improvement is matched by the ability of the CES model to get closer to the data in terms of second moments, especially the volatilities of output, consumption and the real wage, and the autocorrelation functions for inflation and the nominal interest rate. A comparison with a DSGE-VAR further confirms the ability of the CES model to reduce model misspecification. The main message then for DSGE models is that we should dismiss once and for all the use of CD for business cycle analysis.

Our CES specification allows us to introduce a capital-augmenting shock alongside the labour-augmenting variety. However we find this does not bring about an improvement in the model fit and the contribution of the a capital-augmenting shock in the variance decomposition is small. We have noted the well-known result that a bgp requires either Cobb-Douglas technology or that technical change must be driven solely by the labour-augmenting variety. This raises an obstacle to the prospect of unifying business cycle analysis with long-term endogenous growth based on CES technology. One possible way forward is to follow León-Ledesma and Satchi (2010); they provide a model of *optimal*

choice of CES production technology that results in a bgp with both labour and capital-augmenting technical change. Then CES prevails in the short-run but CD in the long-run, thus allowing a capital-augmenting technical change contribution to long-run growth. These authors provide an alternative production function with these properties. A possible line for further research would be to incorporate this into the SW-type model of this paper and to assess its empirical performance.

The area where the CES model remains a concern in terms of model misspecification is in the second moments involving wages and hours. For example both CD and CES models fail miserably in reproducing the negative correlation between output and hours; furthermore the CES model produces far too much persistence in hours. As pointed out by Rowthorn (1999), a low capital-labour substitutability is crucial for understanding unemployment persistence. This suggests that future research should also look more closely at the labour market and introduce search-match frictions and unemployment alongside CES production.

References

Acemoglu, D. (2002). Technical Change, Inequality and the Labor Market. *Journal of Economic Literature*, **40**(1), 7–72.

Acemoglu, D. (2003). Labor- and capital-augmenting technical change. *Journal of the European Economic Association*, **1**, 1–37.

Adjemian, S., Darracq-Paries, M., and Moyen, S. (2008). Towards a monetary policy evaluation framework. ECB Working Paper No. 942.

Antràs, P. (2004). Is the US Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution. *Contributions to Macroeconomics*, **4**, Article 4.

Arrow, K. J., Chenery, H. B., Minhas, B. S., and Solow, R. M. (1961). Capital-labor substitution and economic efficiency. *Review of Economics and Statistics*, **43**(3), 225–250.

Barnett, W. (2004). Dimensions and Economics: Some Problems. *Quarterly Journal of Austrian Economics*, **7**(1), 95–104.

Blanchard, O. J. (1997). The Medium Run. *Brookings Papers on Economic Activity*, **2**, 89–158.

Caballero, R. J. and Hammour, M. L. (1998). Jobless growth: Appropriablity, factor substitution and unemployment. *Carnegie-Rochester Conference Proceedings*, **48**, 51–94.

Cantore, C. and Levine, P. (2010). Getting normalization right. Mimeo.

Cantore, C., León-Ledesma, M., McAdam, P., and Willman, A. (2010a). Shocking stuff: Technology, Hours, and Factor Substitution. Mimeo.

Cantore, C., Ferroni, F., and León-Ledesma, M. (2010b). Hours worked and technology shocks: A structural interpretation for the time varying relation. Mimeo.

Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2009). New keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, **1**(1), 242–66.

Chirinko, R. S. (2008). Sigma: The Long and Short of It. *Journal of Macroeconomics*, **30**(2), 671–686.

Christiano, L., Eichenbaum, M., and Evans, C. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, **113**, 1–45.

Del Negro, M. and Schorfheide, F. (2004). Priors from general equilibrium models for vars. *International Economic Review*, **45**, 643–673.

Del Negro, M., Schorfheide, F., Smets, F., and Wouters, R. (2007). On the fit of new keynesian models. *Journal of Business and Economic Statistics*, **25**(2), 123–162.

Diamond, P. A., McFadden, D., and Rodriguez, M. (1978). Measurement of the elasticity of substitution and bias of technical change. In M. Fuss and D. McFadden, editors, *Production Economics, Vol. 2*, pages 125–147. Amsterdam and North Holland.

Fernald, J. G. (2007). Trend breaks, long-run restrictions, and contractionary technology improvements. *Journal of Monetary Economics*, **54**(8), 2467–2485.

Galí, J. and Gambetti, L. (2009). On the sources of the great moderation. *American Economic Journal: Macroeconomics*, **1**(1), 26–57.

Gertler, M. and Karadi, P. (2009). A Model of Unconventional Policy. Mimeo, NYU.

Geweke, J. (1999). Using Simulation Methods for Bayesian Econometric Models:Inference, Development, and Communication. *Econometric Reviews*, **18**, 1–127.

Jones, C. I. (2003). Growth, capital shares, and a new perspective on production functions. mimeo, Stanford University.

Jones, C. I. (2005). The shape of production functions and the direction of technical change. *Quarterly Journal of Economics*, **120**(2), 517–549.

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

Kennedy, C. (1964). Induced bias in innovation and the theory of distribution. *Economic Journal*, **74**, 541–547.

Klump, R. and Preissler, H. (2000). CES production functions and economic growth. *Scandinavian Journal of Economics*, **102**, 41–56.

Klump, R., McAdam, P., and Willman, A. (2007). Factor Substitution and Factor Augmenting Technical Progress in the US. *Review of Economics and Statistics*, **89**(1), 183–92.

La Grandville, O. de. (1989). In Quest of the Slutzky Diamond. *American Economic Review*, **79**, 468–481.

La Grandville, O. de. (2009). *Economic Growth: A Unified Approach*. Cambridge University Press.

León-Ledesma, M. and Satchi, M. (2010). A Note on Balanced Growth with a Less than Unitary Elasticity of Substitution. University of Kent, School of Economics DP 1007.

León-Ledesma, M. A., McAdam, P., and Willman, A. (2010). Identifying the elasticity of substitution with biased technical change. *American Economic Review*, **100**(4), 1330–57.

Mankiw, N. G. (1995). The growth of nations. *Brookings Papers on Economic Activity*, **1**, 275–310.

Marquetti, A. A. (2003). Analyzing historical and regional patterns of technical change from a classical-marxian perspective. *Journal of Economic Behavior and Organization*, **52**, 191–200.

McAdam, P. and Willman, A. (2008). Medium Run Redux. Working Paper No. 915, European Central Bank.

Rowthorn, R. (1999). Unemployment, wage bargaining and capital-labour substitution. *Cambridge Journal of Economics*, **23**, 413–425.

Santeulália-Llopis, R. and Ríos Rull, J. (2009). Redistributive shocks and productivity shock. Mimeo.

Smets, F. and Wouters, R. (2003). An estimated Stochastic Dynamic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association*, **1(5)**, 1123–1175.

Smets, F. and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review*, **97**(3), 586–606.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, **70**, 65–94.

7 Appendix

A Expressing Summations as Difference Equations

In the first order conditions for Calvo contracts and expressions for value functions we are confronted with expected discounted sums of the general form

$$\Omega_t = E_t \left[\sum_{k=0}^{\infty} \beta^k X_{t,t+k} Y_{t+k} \right] \tag{A.1}$$

where $X_{t,t+k}$ has the property $X_{t,t+k} = X_{t,t+1}X_{t+1,t+k}$ (for example an inflation, interest or discount rate over the interval $[t, t+k]$).

Lemma

Ω_t can be expressed as

$$\Omega_t = X_{t,t}Y_t + \beta E_t \left[X_{t,t+1}\Omega_{t+1}\right] \tag{A.2}$$

Proof

$$\begin{aligned}
\Omega_t &= X_{t,t}Y_t + E_t\left[\sum_{k=1}^{\infty}\beta^k X_{t,t+k}Y_{t+k}\right] \\
&= X_{t,t}Y_t + E_t\left[\sum_{k'=0}^{\infty}\beta^{k'+1} X_{t,t+k'+1}Y_{t+k'+1}\right] \\
&= X_{t,t}Y_t + \beta E_t\left[\sum_{k'=0}^{\infty}\beta^{k'} X_{t,t+1}X_{t+1,t+k'+1}Y_{t+k'+1}\right] \\
&= X_{t,t}Y_t + \beta E_t\left[X_{t,t+1}\Omega_{t+1}\right] \qquad \square
\end{aligned}$$

B Proof of Price and Wage Dispersion Results

For prices and without indexation, in the next period, ξ_p of these firms will keep their old prices, and $(1-\xi_p)$ will change their prices to P^O_{t+1}. By the law of large numbers, we assume that the distribution of prices among those firms that do not change their prices is the same as the overall distribution in period t. It follows that we may write

$$\begin{aligned}
\Delta_{p,t+1} &= \xi_p \int_{m\ no\ change}\left(\frac{P_t(m)}{P_{t+1}}\right)^{-\zeta} + (1-\xi_p)\left(\frac{P^0_{t+1}}{P_{t+1}}\right)^{-\zeta} \\
&= \xi_p\left(\frac{P_t}{P_{t+1}}\right)^{-\zeta}\int_{m\ no\ change}\left(\frac{P_t(m)}{P_t}\right)^{-\zeta} dm + (1-\xi_p)\left(\frac{P^0_{t+1}}{P_{t+1}}\right)^{-\zeta} \\
&= \xi_p\left(\frac{P_t}{P_{t+1}}\right)^{-\zeta}\int_{m}\left(\frac{P_t(m)}{P_t}\right)^{-\zeta} dm + (1-\xi_p)\left(\frac{P^0_{t+1}}{P_{t+1}}\right)^{-\zeta} \\
&= \xi_p\Pi^{\zeta}_{t+1}\Delta_{p,t} + (1-\xi_p)\left(\frac{P^0_{t+1}}{P_{t+1}}\right)^{-\zeta}
\end{aligned} \tag{B.1}$$

The generalization to indexation is straightforward.

C Posteriors

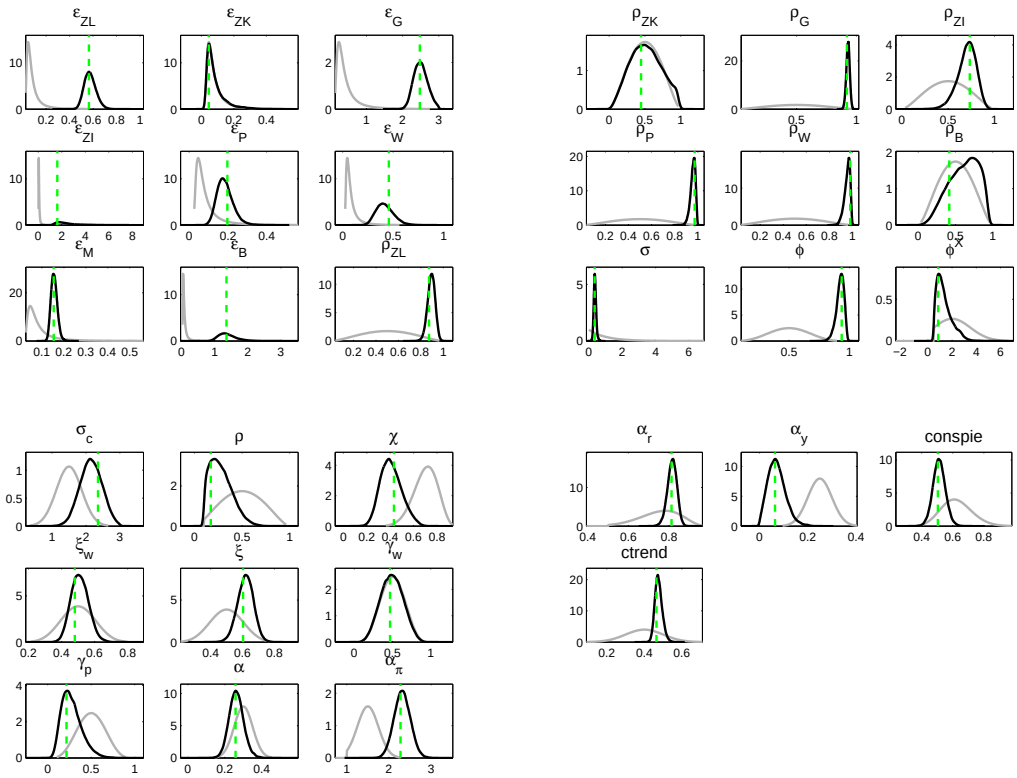


Figure 5: Priors and Posteriors distributions

D Figures

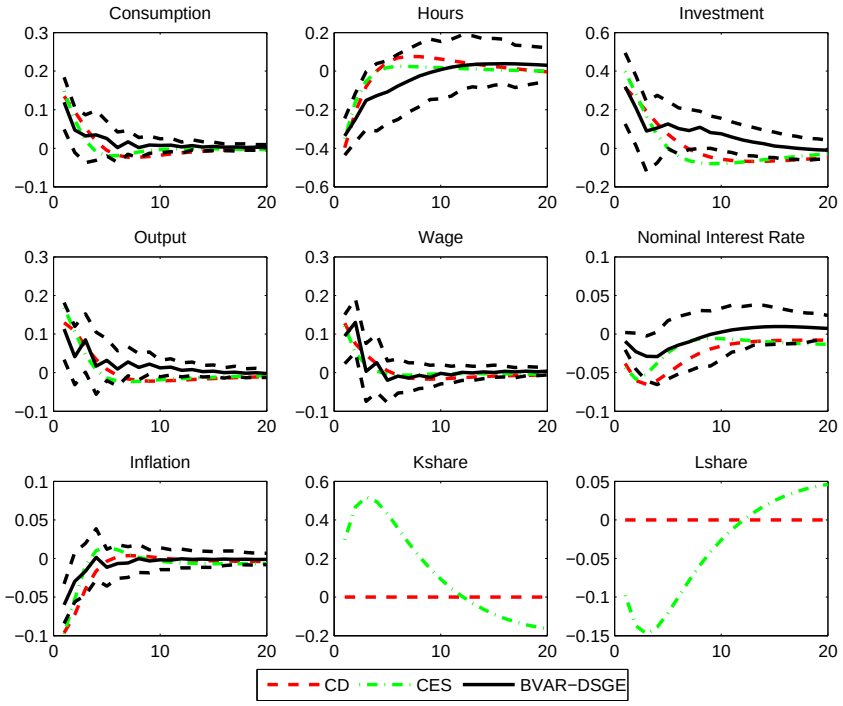


Figure 6: Bayesian IRFs - Labour-augmenting shock$^{\diamondsuit}$

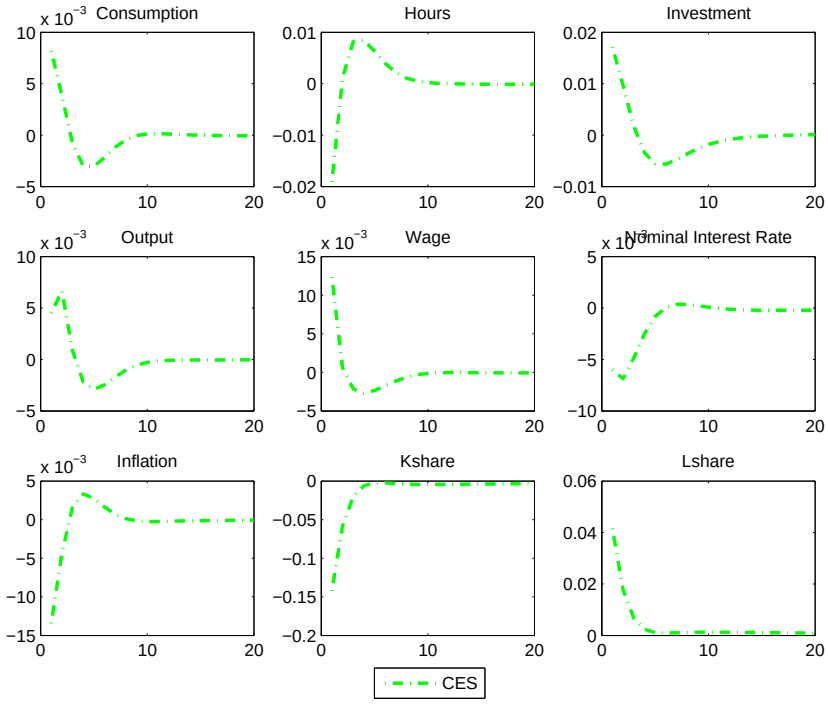


Figure 7: Bayesian IRFs - Capital-augmenting shock

$\diamondsuit$ BVAR-DSGE($\hat{\lambda} = 1.1$): the dashed lines are the first and ninth posterior deciles of the VAR's IRFs. The bold black curve is the posterior mean of the VAR's IRfs.

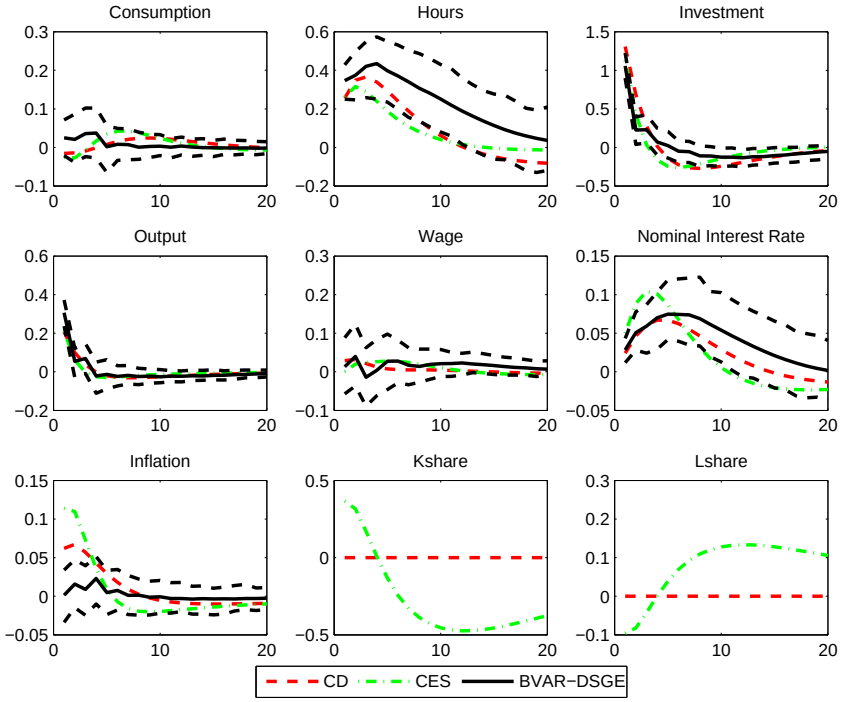


Figure 8: Bayesian IRFs - Investment-specific shock

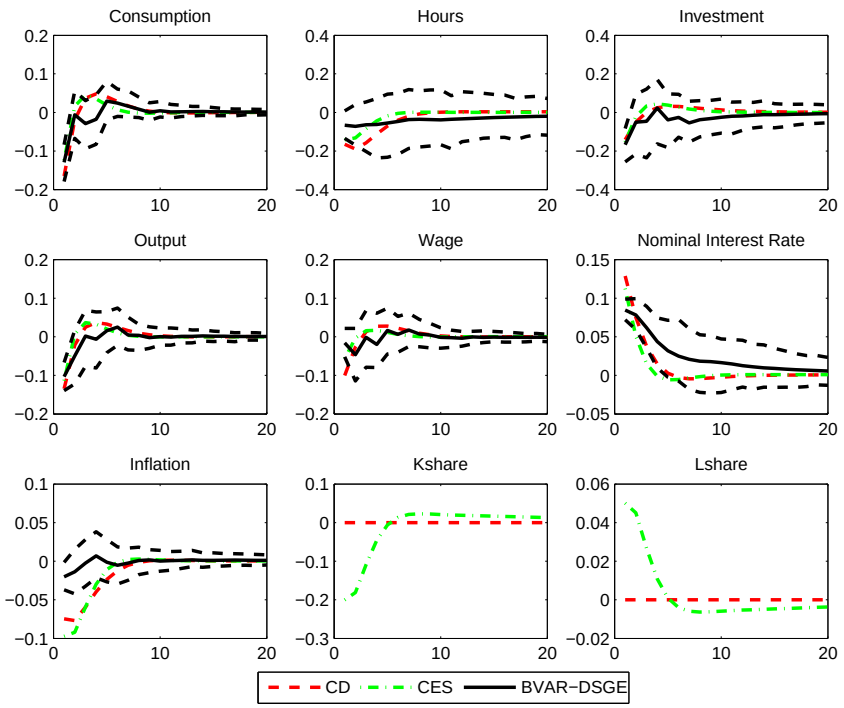


Figure 9: Bayesian IRFs - Monetary policy shock

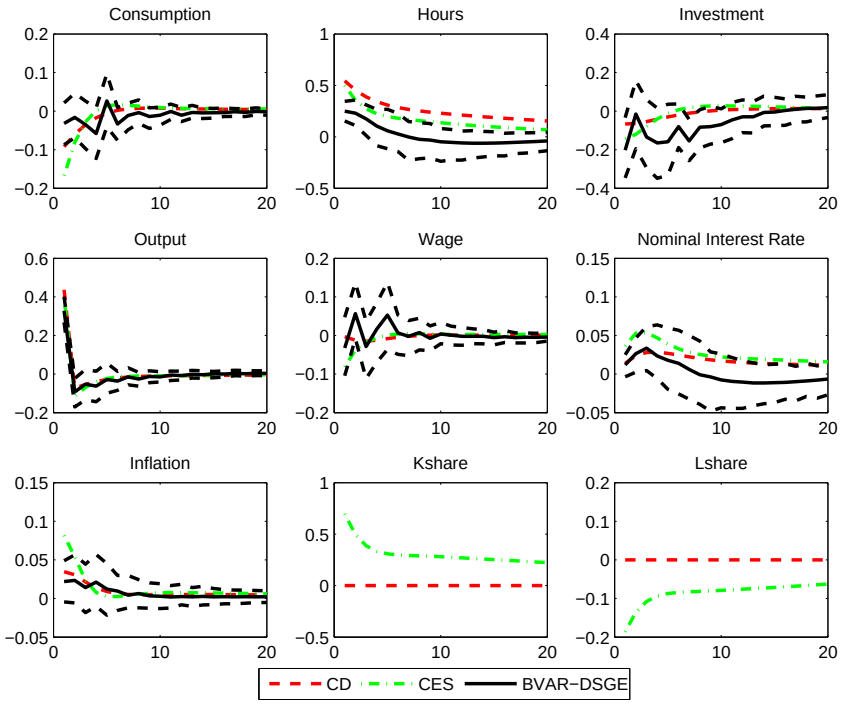


Figure 10: Bayesian IRFs - Government spending shock

Forecast horizon	Observable variables	Productivity (K)	Productivity (L)	Government spending	Mark-up (price)	Investment	Mark-up (wage)	Monetary policy	Preference
		Shocks of the Estimated DSGE Models							
t=1	Output	0.0 (0.00)✠	5.1 (7.1)	32.8 (31.9)	21.4 (19.4)	14.2 (12.6)	20.3 (23.3)	3.9 (2.9)	2.4 (2.8)
	Consumption	0.0 (0.01)	6.8 (5.0)	4.2 (10.0)	12.2 (7.2)	1.8 (0.9)	32.5 (31.6)	9.1 (3.9)	33.4 (41.4)
	Investment	0.0 (0.01)	5.4 (9.7)	0.2 (2.0)	24.7 (26.4)	51.3 (38.6)	14.8 (19.3)	1.2 (2.1)	2.3 (2.0)
	Inflation	0.0 (0.09)	15.6 (12.3)	2.4 (12.9)	24.7 (13.5)	15.0 (30.7)	16.6 (3.5)	13.5 (17.4)	12.3 (9.7)
	Real wage	0.0 (0.02)	5.7 (3.4)	0.01 (2.3)	49.5 (33.0)	0.02 (0.1)	29.6 (52.8)	4.0 (0.9)	11.2 (7.6)
	Interest rate	0.0 (0.05)	6.4 (7.7)	1.0 (8.0)	9.6 (6.5)	6.1 (18.6)	6.4 (1.3)	65.3 (51.2)	5.1 (6.6)
	Hours worked	0.0 (0.01)	12.3 (12.2)	30.5 (30.2)	19.0 (17.7)	13.2 (11.9)	19.5 (22.8)	3.5 (2.7)	2.1 (2.5)
t=4	Output	0.0 (0.00)	6.8 (7.7)	23.8 (27.9)	25.7 (20.5)	12.5 (10.4)	25.1 (26.9)	3.0 (2.6)	3.1 (4.0)
	Consumption	0.0 (0.01)	7.6 (4.9)	4.5 (10.0)	13.0 (6.9)	2.1 (1.3)	35.3 (32.1)	8.4 (3.8)	29.2 (41.0)
	Investment	0.0 (0.01)	6.6 (9.9)	0.3 (2.5)	30.2 (28.6)	39.5 (30.7)	20.0 (24.6)	0.8 (1.8)	2.7 (1.9)
	Inflation	0.0 (0.06)	11.4 (8.7)	2.3 (10.2)	15.5 (9.9)	25.1 (38.6)	17.5 (3.1)	16.8 (21.8)	11.4 (7.5)
	Real wage	0.0 (0.01)	6.7 (3.4)	0.2 (2.4)	53.1 (33.7)	0.4 (0.4)	25.4 (50.8)	3.7 (0.9)	10.5 (8.5)
	Interest rate	0.0 (0.03)	10.7 (7.0)	2.5 (12.0)	10.1 (3.4)	26.5 (53.5)	16.3 (1.2)	21.9 (15.9)	11.9 (7.0)
	Hours worked	0.0 (0.00)	2.0 (2.4)	10.7 (9.5)	35.2 (31.8)	13.2 (7.9)	36.9 (46.7)	1.5 (1.1)	0.5 (0.6)
t=10	Output	0.0 (0.01)	6.8 (8.3)	23.2 (27.4)	25.2 (20.22)	13.7 (11.0)	24.8 (26.5)	3.1 (2.7)	3.3 (4.0)
	Consumption	0.0 (0.01)	7.6 (4.8)	4.3 (9.8)	12.8 (6.9)	3.3 (3.7)	34.3 (31.1)	8.4 (3.7)	29.4 (39.9)
	Investment	0.0 (0.01)	6.3 (10.5)	0.3 (2.5)	28.4 (27.0)	42.2 (32.8)	19.1 (23.1)	0.9 (1.8)	2.9 (2.3)
	Inflation	0.0 (0.06)	10.9 (8.7)	2.3 (9.9)	16.7 (10.3)	26.1 (38.5)	16.5 (4.0)	16.2 (20.8)	11.3 (7.8)
	Real wage	0.0 (0.01)	6.7 (3.3)	0.2 (2.4)	51.6 (33.0)	0.4 (1.0)	26.1 (51.1)	3.9 (0.9)	11.1 (8.4)
	Interest rate	0.0 (0.03)	8.6 (5.3)	3.2 (12.7)	7.9 (8.7)	41.0 (54.6)	15.5 (1.7)	14.0 (11.9)	9.8 (5.2)
	Hours worked	0.0 (0.00)	1.3 (1.1)	6.0 (4.6)	41.1 (32.3)	7.1 (3.4)	43.5 (57.8)	0.6 (0.5)	0.5 (0.3)
t=100	Output	0.0 (0.01)	7.1 (8.4)	22.2 (27.1)	25.4 (20.4)	13.5 (10.9)	25.7 (26.7)	3.0 (2.6)	3.2 (3.9)
	Consumption	0.0 (0.01)	7.6 (4.8)	4.3 (9.8)	12.9 (7.7)	3.9 (3.8)	34.3 (31.5)	8.2 (3.6)	28.8 (38.8)
	Investment	0.0 (0.01)	6.6 (10.6)	0.3 (2.6)	28.8 (27.3)	40.4 (31.9)	20.2 (23.6)	0.8 (1.7)	2.9 (2.3)
	Inflation	0.0 (0.03)	10.3 (5.6)	2.6 (6.4)	18.1 (16.5)	25.2 (23.8)	18.7 (30.1)	14.8 (12.4)	10.3 (4.6)
	Real wage	0.0 (0.01)	6.9 (3.3)	0.2 (2.3)	51.8 (32.5)	0.5 (1.1)	26.0 (51.8)	3.9 (0.8)	10.8 (8.1)
	Interest rate	0.0 (0.01)	7.9 (3.7)	4.7 (8.0)	13.7 (18.2)	34.8 (25.2)	20.6 (37.8)	10.8 (4.9)	7.6 (2.2)
	Hours worked	0.0 (0.00)	1.0 (0.4)	6.3 (2.1)	45.2 (27.5)	5.2 (1.4)	41.5 (68.3)	0.4 (0.2)	0.4 (0.1)

Table 4: Variance Decomposition - Comparison of CD and CES Specifications (in Percent)◇

◇ All the variance decomposition is computed from the model solutions (order of approximation = 1). The results are based on the models' posterior distribution.
✠ We report the results obtained from the most general model with the CES production function in parentheses.

parameter	prior mean	post. mean CD (SW07)	post. mean CES	5% CES	95% CES	Prior	pstdev CES
ρ_{ZL}	0.5	0.9004 (0.95)	0.8919	0.8386	0.9476	beta	0.2
ρ_{ZK}	0.5	N/A (N/A)	**0.5120**	0.1729	0.8790	beta	0.2
ρ_G	0.5	0.9561 (0.97*)	0.9297	0.9096	0.9533	beta	0.2
ρ_{ZI}	0.5	0.7162 (0.71)	0.6992	0.5397	0.8617	beta	0.2
ρ_P	0.5	0.9239 (0.89*)	0.9566	0.9235	0.9908	beta	0.2
ρ_W	0.5	0.9046 (0.96*)	0.9528	0.9142	0.9912	beta	0.2
ρ_B	0.5	0.6324 (N/A)	0.6057	0.3109	0.9183	beta	0.2
ε_{ZL}	0.1	0.5735 (0.45)	0.5743	0.4899	0.6532	invg	2.0
ε_{ZK}	0.1	N/A (N/A)	**0.0853**	0.0238	0.1612	invg	2.0
ε_G	0.5	2.4634 (0.53*)	2.4958	2.1909	2.8166	invg	2.0
ε_{ZI}	0.1	3.6743 (**0.45**)	**2.4553**	1.2221	3.9710	invg	2.0
ε_P	0.1	0.1136 (0.14*)	**0.1846**	0.1169	0.2486	invg	2.0
ε_W	0.1	0.3835 (0.24*)	0.4187	0.2769	0.5577	invg	2.0
ε_M	0.1	0.1585 (0.24)	0.1556	0.1318	0.1778	invg	2.0
ε_B	0.1	1.4830 (N/A)	**1.3850**	0.9620	1.7984	invg	2.0

Table 5: Posterior results for the exogenous shocks

parameter	prior mean	post. mean CD (SW07)	post. mean CES	5% CES	95% CES	Prior	pstdev CES
σ	1	1 (1)	**0.3622**	0.2391	0.4755	gamma	1
ϕ	0.5	0.9002 (0.54)	0.9215	0.8725	0.9738	beta	0.15
ϕ^X	2	2.5530 (2.87)	**1.3515**	0.5	2.2683	norm	1.5
σ_c	1.5	2.0158 (1.38*)	**2.1431**	1.6078	2.6755	norm	0.3750
ϱ	0.5	0.5061 (1.83*)	**0.2785**	0.1	0.4398	beta	0.20
χ	0.7	0.5092 (0.71*)	**0.4065**	0.2597	0.5540	beta	0.1
ξ_w	0.5	0.5632 (0.7)	0.5093	0.4233	0.5950	beta	0.1
ξ_p	0.5	0.7173 (0.66)	**0.6171**	0.5367	0.6994	beta	0.1
γ_w	0.5	0.4948 (0.58)	0.4903	0.2459	0.7259	beta	0.15
γ_p	0.5	0.2215 (0.24)	0.2749	0.0929	0.4478	beta	0.15
α	0.3	0.2239 (0.19)	0.2615	0.1985	0.3253	norm	0.05
α_π	1.5	2.1699 (2.04)	**2.3108**	1.9897	2.6187	norm	0.25
α_r	0.75	0.8221 (0.81)	0.8146	0.7772	0.8520	beta	0.1
α_y	0.25	0.0407 (0.08)	0.0710	0.0110	0.1254	norm	0.05
conspie	0.625	0.5201 (0.78)	0.5116	0.4461	0.5774	gamma	0.1
ctrend	0.4	0.4730 (0.43)	0.4756	0.4432	0.5072	norm	0.1

Table 6: Posteriors results for model parameters

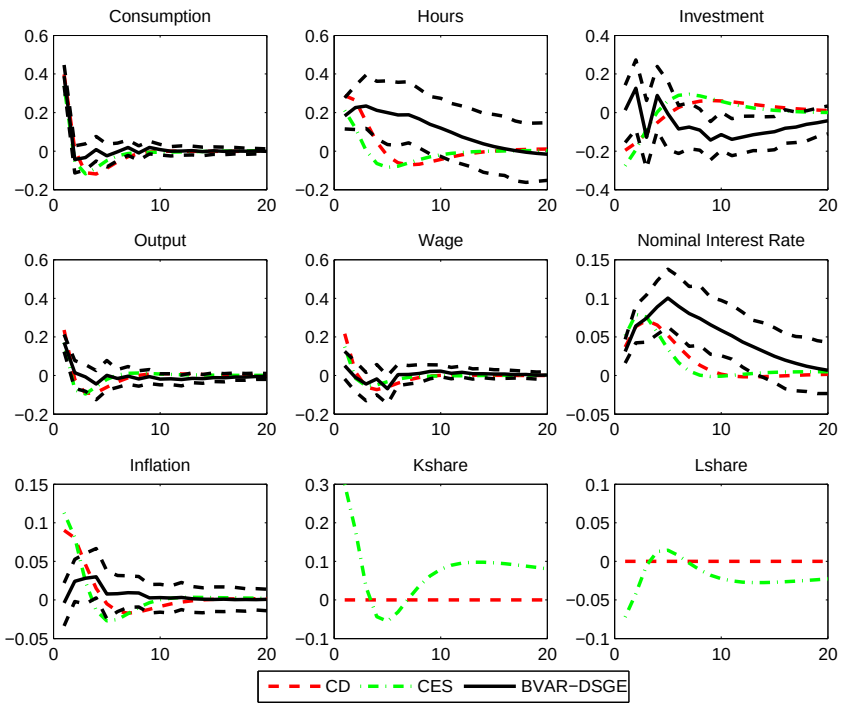


Figure 11: Bayesian IRFs - Preference shock

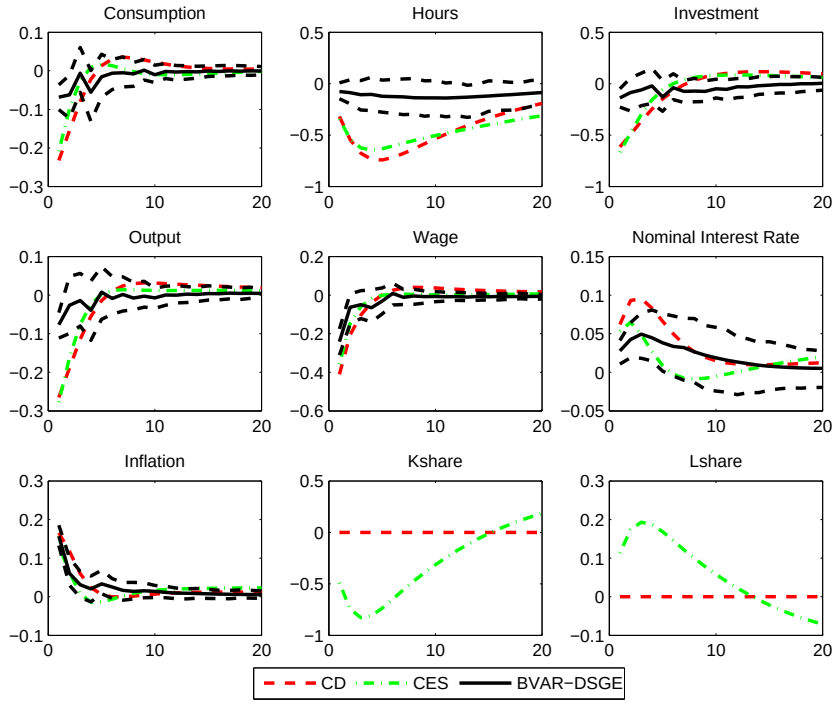


Figure 12: Bayesian IRFs - Price mark-up shock

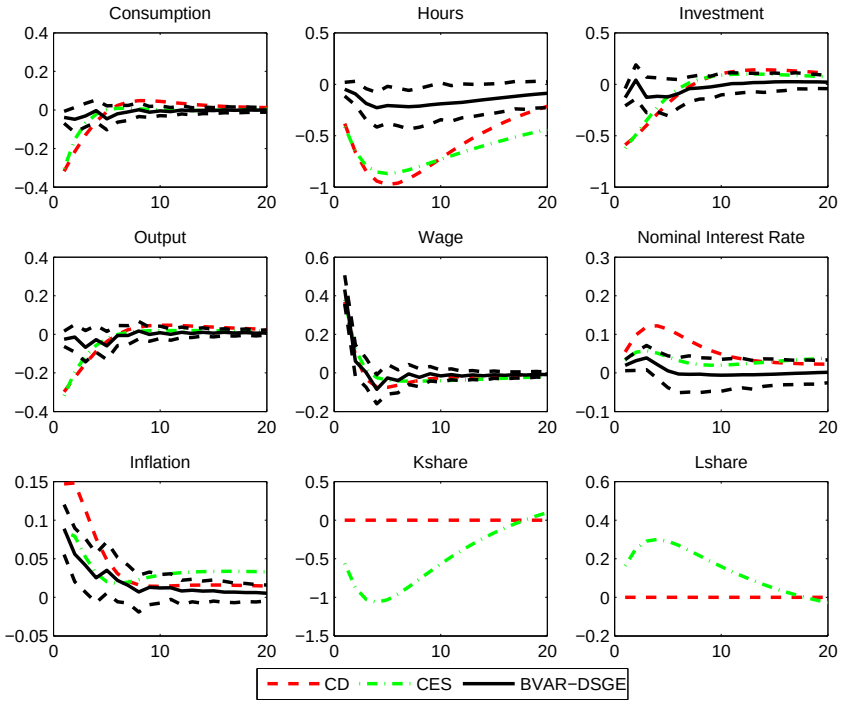
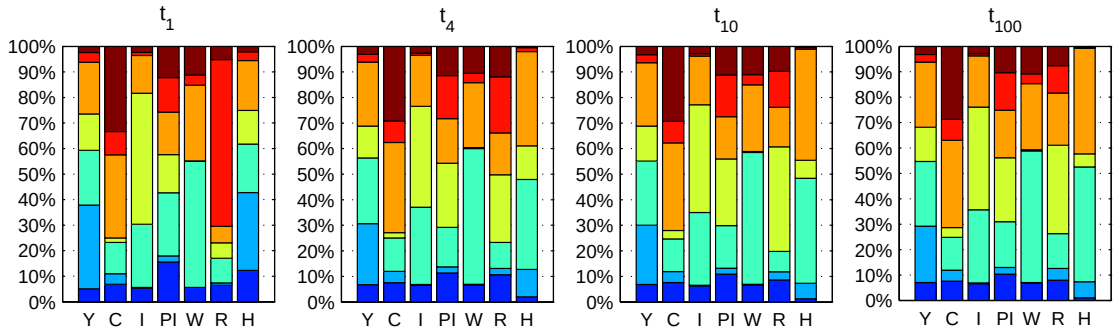
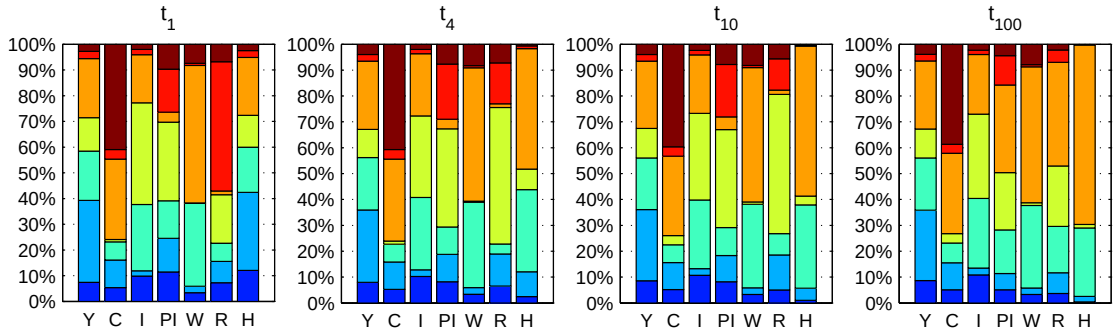


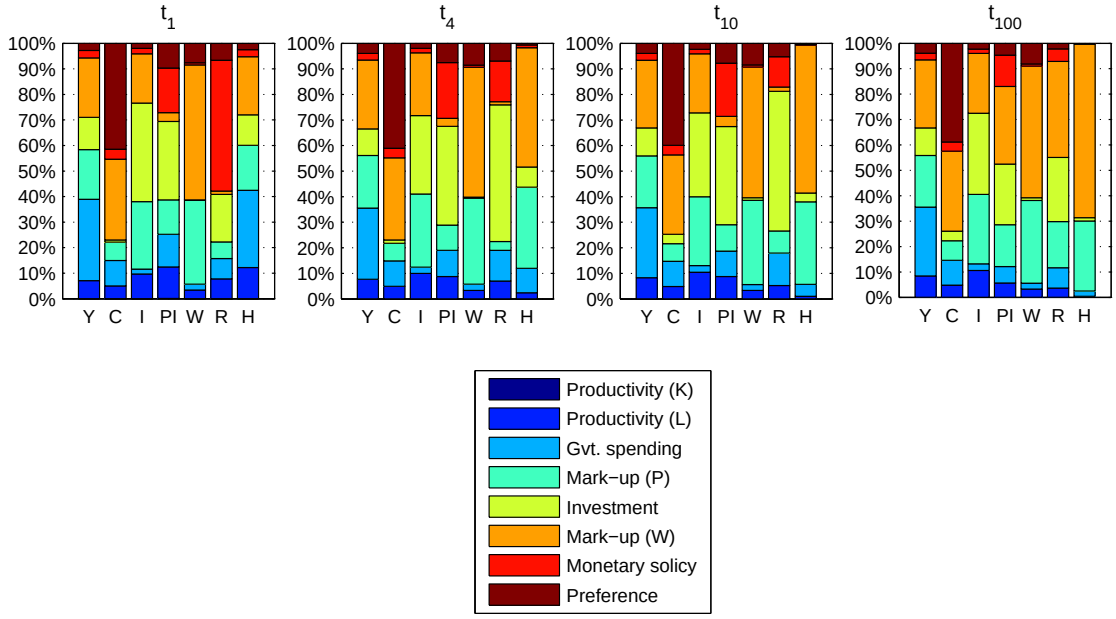
Figure 13: Bayesian IRFs - Wage mark-up shock



(a) C-D



(b) CES (one shock)



(c) CES (two shocks)

Figure 14: Variance Decomposition