Bayesian Estimation of a Dynamic Partial-Equilibrium Model for Investment*

Matthias Kredler†

August 26, 2005

Abstract

This paper revisits the question if the user cost of capital plays an important role for investment decisions using Bayesian estimation techniques. These methods offer advantages over classical econometric tools in this area: The most important are that prior distributions offer a convincing way to confine the support of model parameters and that confidence intervals are more reliable when model parameters approach the bounds of their support. I use aggregate investment data from six industrial sectors in the UK to estimate a parsimonious partial-equilibrium model. The Kalman Filter is used to evaluate the likelihood and MCMC methods are employed to draw from the posterior distribution. The main finding is that the real interest rate accounts for less than 10 percent of the variance in investment under the 99-percent confidence level; this result is robust across sectors.

*This paper is handed in as a term paper for the course "Advanced Macroeconomics", taught by T. Sargent and F. Schorfheide at *New York University* in the Fall 2004 semester.

†Department of Economics, New York University, mk1168@nyu.edu

1 Introduction

Surprisingly, it is still an open question in macroeconomics if the user cost of capital matters for investment. On a theoretical level, all standard models predict that the real interest rate should have a major impact on investment.[1] On an empirical level, however, this effect is hard to find. As Caballero (2000) notes, "the empirical investment literature has been nearly merciless in evaluating [these] theories". He summarizes the empirical literature by saying that even studies that *do* find an impact of the user cost of capital on investment leave a large fraction of the variance in investment unexplained.

In the last two decades, yet, evidence has accumulated that the impact of capital costs on investment is larger than previously thought. Caballero (2000) cites studies that focus on long-run capital formation and episodes where investment incentives underwent changes due to tax reform; in all of these studies, the coefficients on the cost of capital are larger than in previous work. Bernanke (1983) found that high real interest rates were largely responsible for the low pace of capital expenditure in the late 1970s and early 1980s in the U.S.

Recently, new methods have been added to the econometrician's toolkit that warrant a new look at this topic: The vast increase in computing power and the development of Markov-Chain Monte-Carlo (MCMC) methods have made the estimation of Bayesian models feasible on standard personal computers. These techniques offer some significant advantages over frequentist methods.

First, they give the researcher the possibility to force certain model parameters into the economically sensible range by specifying their prior distribution. Moreover, as Johannes & Polson (2005) note, they offer a coherent treatment of both model parameters and hidden state variables in the model — they treat both as random variables which have a joint distribution with the observables. In many cases, also, Bayesian methods are more reliable and robust than classical maximization-based methods; this is especially the case when filtering methods are employed to evaluate the likelihood of an observed time series. Ill-behaved likelihood surfaces and the lack of analytical derivatives can make maximization in a high-dimensional parameter space a formidable task.

Lastly, and maybe most importantly, the Bayesian approach yields exact

[1] See for example the models by Abel & Eberly (1994) and Hayashi (1982).

confidence intervals for the parameters as well as for functions of these; in our case, for example, the fraction of variance in investment explained by movements in the interest rate. If this fraction is close to zero, the Bayesian approach yields confidence intervals that are much more credible than the ones obtained from first-order approximations in a classical framework. The latter are bound to be symmetric and will extend below zero in many cases although they refer to a fraction, which by definition has to be positive.

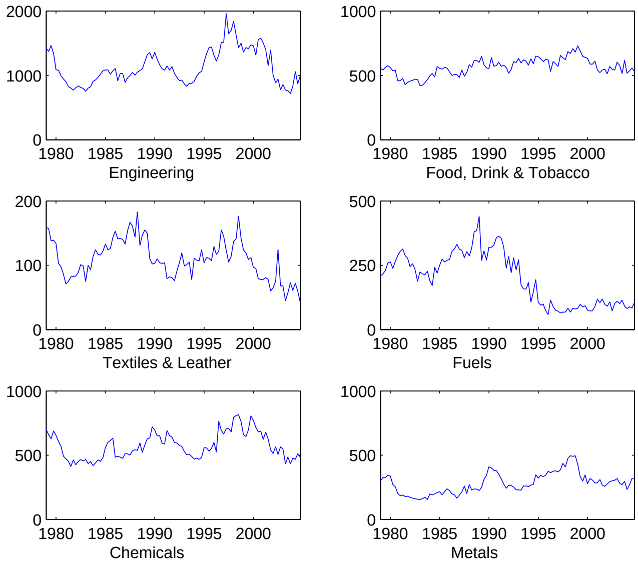
In this paper, I estimate a parsimonious partial-equilibrium model for aggregate investment in six industrial sectors in the UK with Bayesian methods. I use the Kalman Filter to evaluate the likelihood and draw from the posterior distribution of the parameters and the hidden state of the model using MCMC methods. The main finding is that the real interest rate accounts for less than 10 percent of the variance in investment in all of these sectors; this statement holds under a 95-percent confidence level. However, as in probably any model, there are indications of misspecification which suggest to compare this parsimonious framework to different, possibly more elaborate, models in future research.

2 Data

The data set for investment is taken from the UK's *Office for National Statistics*'s web site. It consists of six quarterly time series for aggregate business investment by industry at 2001 prices in millions of British Pounds. The six industries are: Chemicals; Engineering; Food, Drink & Tobacco; Fuels; Metals; Textiles & Leather. The data set ranges from the first quarter of 1979 to the fourth quarter of 2004, yielding 104 observations. The time series are seasonally adjusted and measure investment by total capital expenditure in the sector. The time series are plotted in figure 1.

As for the cost of capital, it would be optimal to use some direct measure of the real interest rate in the estimation. Inflation-indexed bonds traded on the British capital markets come very close to such a direct measurement. However, these instruments are only traded at very long maturities. Also, as Barr & Campbell (1996) note, these bonds are only imperfect measures of real interest rates — even in the long term — since they leave the buyer unprotected against inflation occurring in the last months before maturity. Instead, I choose to use the ex-post real interest rate and regard it as a noisy signal of the expected real interest rate in the market. This procedure uses

Figure 1: Investment Data



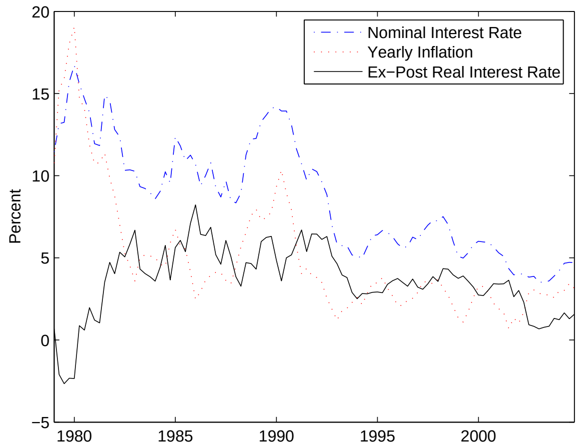
the Kalman Filter and is due to Fama & Gibbons (1982). The ex-post real interest rate is defined as the difference between the nominal interest rate and inflation. The data are taken from the *Bank of England*'s web site. Figure 2 shows the respective time series over the sample period.

For inflation, I take the difference between the logarithm of the price level over one year.[2] Hamilton (1994*b*) suggests this as a simple procedure to remove seasonality. Technically, the resulting figure is a 4-quarter moving average over annualized quarterly inflation (seasonally adjusted) and *not* annualized quarterly inflation. Yet, this smoothing of the inflation time series essentially removes noise from the data and should not have a significant effect on the results — the relevant variable in the estimation procedure is the underlying *expected inflation* (in seasonally adjusted terms), which has certainly less variance in the high frequencies than realized inflation.

Visual inspection suggests that the series can be assumed to be stationary over the sample period. Hence, no effort will be made to include a time trend in the theoretical model.

[2]Technically: $\pi_t = \log CPI_{t+4} - \log CPI_t$, where π_t is annualized inflation in quarter t and CPI_t is the consumer price index in quarter t.

Figure 2: Interest-Rate Data



3 The Model

The model describes a dynamic partial equilibrium in a competitive industrial sector, where the interest rate and demand are exogenous. The representative firm in the sector produces a single good with a constant-returns-to-scale technology. The only factor employed is capital:

$$y_t = Ak_t,$$

where y_t is the quantity of the good produced by the firm in period t, k_t is the capital stock in t, and $A > 0$ is a productivity coefficient. I use lowercase letters to denote quantities on the firm level and capitals to denote their aggregate analogon. Capital accumulation is frictionless:

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

where i_t is investment by the firm in period t and $0 \leq \delta \leq 1$ is the depreciation rate. The investment good i_t can be obtained at the constant price of one unit of the output good. Demand for the good produced in the sector is given by

$$P_t = C_d Y_t^{-\gamma} \nu_t = C_d A^{-\gamma} K_t^{-\gamma} \nu_t, \tag{2}$$

where P_t is the price of the good, $C_d > 0$ is a sector-specific constant, and Y_t is aggregate supply in the sector. The elasticity of demand $\gamma \geq 0$ is constant and time-invariant. The demand shifter ν_t induces time-varying investment incentives in the sector; its logarithm $\tilde{\nu}_t \equiv \log \nu_t$ is assumed to follow an AR(1) process with a Gaussian innovation:

$$\tilde{\nu}_{t+1} = \rho_\nu \tilde{\nu}_t + \sigma_\nu \tilde{\varepsilon}_{\nu,t+1}, \tag{3}$$

where $\tilde{\varepsilon}^\nu_t$ is a serially uncorrelated normal shock with variance 1. To simplify notation, introduce $C \equiv C_d A^{-\gamma} > 0$.

Notice that technology shocks can easily be accommodated in this framework. Suppose we introduce an adequately scaled process ξ_t for the productivity of capital: $y_t = Ak_t\xi_t$. Then, equation (2) becomes: $P_t = CK_t^{-\gamma}\nu_t\xi_t^{-\gamma}$. In this alternative framework, $\hat{\nu}_t \equiv \nu_t\xi_t^{-\gamma}$ would be the demand shifter. Unfortunately, I did not find data series that allow me to disentangle the effects of demand and technology. This would be an interesting extension of this project.

The firm is risk-neutral and has access to an incomplete capital market. There is an exogenously given stochastic sequence for the real interest rate R_t. Specifically, I assume that its deviation from the logarithmic mean, $\tilde{r}_t \equiv \log R_t - E[\log R_t]$, follows an AR(1) process with a standard normal serially uncorrelated shock:

$$\tilde{r}_{t+1} = \rho_r \tilde{r}_t + \sigma_r \tilde{\varepsilon}_{r,t+1} \tag{4}$$

The following notation is adopted to facilitate the ensuing discussion:

$$R_{0,t}^{-1} = \prod_{i=0}^{t-1} R_i^{-1},$$

The firm ranks stochastic sequences of profits by the criterion

$$Q\big(\{k_{t+1}\}_{t=0}^{\infty}, \cdot\big) = E_0 \sum_{t=0}^{\infty} R_{0,t}^{-1}\Big[CK_t^{-\gamma}\nu_t k_t - k_{t+1} + (1-\delta)k_t\Big]$$
$$\text{given} \qquad k_0, \{K_{t+1}\}_{t=0}^{\infty}, \{R_t\}_{t=0}^{\infty},$$

where the sequences $\{K_{t+1}\}_{t=0}^{\infty}, \{R_t\}_{t=0}^{\infty}$ are exogenous and the sequence $\{k_{t+1}\}_{t=0}^{\infty}$ is under control of the firm. The term in brackets is the firm's profit in period t, which consists of the revenue from sales and the cost of investing in the capital stock.

Taking first-order conditions with respect to k_{t+1} and re-arranging yields the following familiar expression:

$$R_t = CK_{t+1}^{-\gamma} E_t[\nu_{t+1}] + (1-\delta) \tag{5}$$

This equation says that the expected profits from investing a marginal unit in a company in the sector must be equal to the interest rate on the capital market. Notice that no variable under control of the firm enters in (5). The equation only gives a restriction on aggregate quantities which makes a risk-neutral investor indifferent between investing a marginal unit in the sector or holding a bond in the wider capital market. In order to obtain a tractable linear expression, I take a first-order Taylor approximation of (5) in logarithms and re-arrange to obtain[3]

$$\tilde{k}_{t+1} = \frac{\rho_\nu}{\gamma}\tilde{\nu}_t - \frac{\bar{R}}{\gamma[\bar{R}-(1-\delta)]}\tilde{r}_t, \tag{6}$$

where $\tilde{k}_t \equiv \log K_t - E[\log K_t]$ and $\bar{R} \equiv E[\log K_t]$. Using equations (3), (4) and the expectation of (6), the impulse response for capital with respect to the driving processes is

$$\tilde{k}_{t+1} = \frac{\rho_\nu \sigma_\nu}{\gamma}\sum_{j=0}^{\infty} \rho_\nu^j \tilde{\varepsilon}_{\nu,t-j} - \frac{\bar{R}\sigma_r}{\gamma[\bar{R}-(1-\delta)]}\sum_{j=0}^{\infty}\rho_r^j \tilde{\varepsilon}_{r,t-j}. \tag{7}$$

To obtain an expression for investment, use the log-linearized law of motion of capital, $\tilde{k}_{t+1} = (1-\delta)\tilde{k}_t + \delta\tilde{i}_t$:

$$\tilde{i}_t = \frac{\rho_\nu \sigma_\nu}{\gamma}\left(\frac{1}{\delta}\tilde{\varepsilon}_{\nu,t} + \sum_{j=1}^{\infty}\rho_\nu^j \tilde{\varepsilon}_{\nu,t-j}\right) - \frac{\bar{R}\sigma_r}{\gamma[\bar{R}-(1-\delta)]}\left(\frac{1}{\delta}\tilde{\varepsilon}_{r,t} + \sum_{j=1}^{\infty}\rho_r^j \tilde{\varepsilon}_{r,t-j}\right) \tag{8}$$

The variance of investment can be decomposed as

$$Var(\tilde{i}_t) = \left(\frac{\rho_\nu \sigma_\nu}{\gamma}\right)^2 \left(\frac{1}{\delta} + \frac{\rho_\nu^2}{1-\rho_\nu^2}\right) - \left(\frac{\bar{R}\sigma_r}{\gamma[\bar{R}-(1-\delta)]}\right)^2 \left(\frac{1}{\delta} + \frac{\rho_r^2}{1-\rho_r^2}\right).$$

The second observed variable is the ex-post real interest rate $r_{p,t}$, which is defined as $r_{n,t} - \pi_t$, where $r_{n,t}$ is the nominal interest rate and π_t is inflation.

[3]Note that when taking expectations of this first-order approximation (or equivalently, in the so-called "deterministic steady state"), the following identity follows: $\bar{R} - (1-\delta) = C\bar{K}^{-\gamma}$.

The ex-ante real interest rate r_t is defined as $r_{n,t} - \pi_{e,t}$, where $\pi_{e,t}$ is the inflation rate expected for period t by the public when entering this period. I follow Fama & Gibbons (1982) and decompose realized inflation into $\pi_t = \pi_{e,t} + \eta_t$, where η_t is a rational-expectations error. Regarding r_t as a hidden state and combining the expressions before, we have

$$r_{p,t} = r_{i,t} - \pi_t = r_t - \eta_t. \tag{9}$$

Since η_t is a rational-expectations error, it is uncorrelated over time and also uncorrelated to other variables known to the agents in the market at time t, as for example the contemporaneous shocks $\varepsilon_{\nu,t+1}$ and $\varepsilon_{r,t+1}$. I furthermore assume that η_t has constant variance σ_η.

Now, everything is in place to write down a state-space system combining (3), (4), (8) and (9):

$$\begin{pmatrix} \tilde{\nu}_{t+1} \\ \tilde{r}_{t+1} \\ \tilde{k}_0 \end{pmatrix} = \begin{pmatrix} \rho_\nu & 0 & 0 \\ 0 & \rho_r & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\nu}_t \\ \tilde{r}_t \\ \tilde{k}_0 \end{pmatrix} + \begin{pmatrix} \sigma_\nu & 0 \\ 0 & \sigma_r \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\varepsilon}_{\nu,t} \\ \tilde{\varepsilon}_{r,t} \end{pmatrix} \tag{10}$$

$$\begin{pmatrix} \tilde{i}_t \\ \tilde{r}_{p,t} \end{pmatrix} = \begin{pmatrix} -\sum_{j=1}^{t}(1-\delta)^j L^j \\ 0 \end{pmatrix} \tilde{i}_t + \begin{pmatrix} \frac{\rho_\nu}{\delta\gamma} & \frac{-\bar{R}}{\delta\gamma[\bar{R}-(1-\delta)]} & -\frac{(1-\delta)^t}{\delta} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\nu}_t \\ \tilde{r}_t \\ \tilde{k}_0 \end{pmatrix} + \begin{pmatrix} 0 \\ \sigma_\eta \end{pmatrix} \eta_t,$$

where $(\tilde{\varepsilon}'_{\nu,t+1}, \tilde{\varepsilon}'_{r,t+1}, \eta'_t)'$ is Gaussian white noise with covariance matrix $\mathbf{I}_3$ and L is the lag operator: $Lx_t = x_{t-1}$. In the language of the Kalman Filter, $(\tilde{\nu}_t, \tilde{r}_t, \tilde{k}_0)'$ is the *hidden state* and $(\tilde{i}_t, \tilde{r}_{p,t})'$ is the *observed state*. The lagged values of $\tilde{i}_t$ are pre-determined at t and can hence be treated as fixed. The timing convention is as follows: The first observation is made at $t = 1$, the last at $t = T$.

The distribution of the hidden state at $t = 0$, which is needed to start the recursions of the Kalman Filter, is the unconditional variance of the variables. From (8), the unconditional variance of $\tilde{k}_0$ can be seen to be

$$Var(\tilde{k}_{t+1}) = \left(\frac{\rho_\nu}{\gamma}\right)^2 Var(\tilde{\nu}_t) + \left(\frac{\bar{R}}{\gamma[\bar{R} - (1-\delta)]}\right)^2 Var(\tilde{r}_t),$$

where $Var(\tilde{\nu}_t) = \frac{\sigma_\nu^2}{1-\rho_\nu^2}$ and $Var(\tilde{r}_t) = \frac{\sigma_r^2}{1-\rho_r^2}$. The covariances of $\tilde{k}_{t+1}$ with $\tilde{\nu}_t$

and $\tilde{r}_t$ are

$$Cov(\tilde{k}_{t+1}, \tilde{\nu}_{t+1}) = Cov(\tilde{k}_{t+1}, \rho_\nu \tilde{\nu}_t) + Cov(\tilde{k}_{t+1}, \sigma_\nu \tilde{\varepsilon}_{\nu,t+1}) = \frac{\rho_\nu^2}{\gamma} Var(\tilde{\nu}_t)$$

$$Cov(\tilde{k}_{t+1}, \tilde{r}_{t+1}) = Cov(\tilde{k}_{t+1}, \rho_r \tilde{r}_t) + Cov(\tilde{k}_{t+1}, \sigma_r \tilde{\varepsilon}_{r,t+1}) = \frac{\rho_r \bar{R}}{\gamma[\bar{R} - (1-\delta)]} Var(\tilde{r}_t)$$

4 Estimation and Computational Issues

I adopt a two-stage strategy to estimate the model described in (10). In the first stage, the parameters ρ_ν, σ_ν and σ_η are determined solely from the interest-rate data by maximum-likelihood estimation (MLE). In the second stage, these parameters are fixed at the estimated values and the remaining four parameters in (10) are estimated from the investment data in a particular sector and the interest-rate data employing Bayesian methods.

This two-stage strategy is not only computationally easier to implement than joint estimation of all seven parameters. It is also in another way a natural way to proceed. Since the six investment time series from the different sectors all contain a noisy signal about the real interest rate that adds to the information contained in the ex-post real rate, it would be theoretically appealing to estimate a model with all sectors jointly. However, this is obviously hard to implement. The alternative strategy of estimating the seven parameters in (10) jointly in *each* of the six sectors, on the other hand, is somewhat awkward as well, since all estimation procedures would yield slightly different posterior distributions for the time path of the real rate. Therefore, the two-stage strategy seems a reasonable way to proceed. At any rate, potential inefficiencies of this method should be small since the investment time series can be expected to contain very little additional information on the real interest rate.

The first step evaluates the likelihood of observing the data for $\{\tilde{r}_{p,t}\}_{t=1}^{104}$ for a fixed triplet of parameters $(\rho_\nu, \sigma_\nu, \sigma_\eta)'$ by applying the Kalman Filter to the simplified system[4]

$$\begin{aligned} \tilde{r}_{t+1} &= \rho_\nu \tilde{r}_t + \sigma_\nu \varepsilon_{\nu,t} \\ \tilde{r}_{p,t} &= \tilde{r}_t + \sigma_\eta \eta_t. \end{aligned}$$

[4]See Hamilton (1994*a*), for example, for the exact procedure of obtaining the likelihood in a state-space system.

Table 1: Prior Distribution

Parameter	Distribution Family	Shape Parameter	Scale Parameter	Support from	to
ρ_ν	Beta	2	1.2	0	1
σ_ν	Inverse Gamma	2	4.5	0	∞
δ	Beta	1.2	3	0	1
γ	Inverse Gamma	2	3	0	∞

The series $\{\tilde{r}_{p,t}\}_{t=1}^{104}$ is obtained by de-meaning the logarithm of the ex-post real interest rate, as described in section 2. Notice that this procedure yields a method-of-moments estimate for the parameter $\bar{R}$, which will be used in the second step as well.

As in all subsequent estimation procedures, I calculate the matrices arising in the application of the Kalman Filter dynamically until they converge to the stationary steady state. These steady-state matrices are computed using the doubling algorithm[5]. This proved to be important since the matrices in the dynamic algorithm diverged after a large number of periods, probably due to small calculation inaccuracies accumulating over time.

In the second stage, the Metropolis-Hastings algorithm[6] is used to draw from the posterior distribution of the remaining four parameters ρ_ν, σ_ν, δ and γ. The prior distribution of the parameters is given in table 4. The four parameters are independently distributed under the prior.

The likelihood of observing the data $\{\tilde{r}_{p,t}, \tilde{i}_t\}_{t=1}^{104}$ given a vector of parameters is evaluated applying the Kalman Filter to the state-space system described in (10). To obtain the series $\{\tilde{i}_t\}_{t=1}^{104}$, again I de-mean the logarithm of the investment time series in the respective sector, which is tantamount to estimating the model parameter C by a moment condition.

Since (10) contains a dynamic component, it is necessary to work with a time-varying system for a certain number of periods. For small values of δ, the matrices calculated for the Kalman Filter do not converge for a long time and sometimes the time-varying system has to be used until the end of the data series. In the other cases, the matrices from the doubling algorithm of the limiting system are utilized after a number of periods applying a convergence

[5]Specifically, I use a program that Ljungqvist & Sargent (2004) provide.

[6]See Johannes & Polson (2005), for example, for a description of how to MCMC techniques.

criterion. For this stable system, the component $\tilde{k}_0$ is dropped from the hidden state.

As suggested by Johannes & Polson (2005), a fat-tailed distribution is used for the jump proposals in the Metropolis-Hastings algorithm. Specifically, I use a t-distribution with 5 degrees of freedom. In the estimations, it proved very important to adjust the scaling of the jump-proposal density to the covariance of the parameters under the posterior. As figure 8 in the appendix suggests, especially the covariance between the parameters γ and σ_ν is very high; in all sectors, their correlation coefficient under the posterior distribution exceeds 0.9. Implementing a well-scaled proposal density allowed me to increase the efficiency of the algorithm by using a relatively large variance for the jump proposals. The jump size was tuned such that the acceptance rate lay in the optimal range between 0.25 and 0.4.[7]

The computations were carried out in *Matlab* and were greatly simplified by the use of object-oriented programming to calculate the likelihood under different parameter vectors and for different dimensions of the hidden state vector.[8]

For each sector, I carried out 20,000 draws from the posterior distribution. For every tenth draw, an additional draw was taken from the process for the hidden state in (10). This strategy is due to the high computational demands of carrying out these draws — note that drawing from this process requires both more memory usage and additional computations to those of the Kalman Filter.

The technique I use to draw from the hidden state is inspired by the smoothing algorithm as described, for example, in Hamilton (1994*a*).[9] To simplify notation, denote a generic state-space system (as the one in (10)) as

[7] See Johannes & Polson (2005) for an overview of theoretical results on the optimal acceptance rate of MCMC algorithms.

[8] Recall that the initial capital stock is dropped from the system once convergence is reached and so the number of hidden states decreases from three to two.

[9] "Smoothing" describes a recursive procedure for obtaining the conditional expectation (or projection, in the non-Gaussian case) $E[x_t|y^T]$ with its associated conditional variance (or mean squared error). Note that the Kalman Filter only gives a recursive formula for finding $E[x_t|y^t]$ and $E[x_{t+1}|y^t]$ — it does not take into account future observations of y_t, which potentially contain useful information about the distribution of x_t.

follows:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{C}\boldsymbol{\varepsilon}_t$$
$$\mathbf{y}_t = \mathbf{F}\mathbf{x}_t + \mathbf{G}\boldsymbol{\eta}_t,$$

where $\mathbf{x}_t$ is a vector of hidden state variables, $\mathbf{y}_t$ is a vector of observed variables, $(\boldsymbol{\varepsilon}_t', \boldsymbol{\eta}_t')'$ is vector white noise and the matrices $\mathbf{A}$, $\mathbf{C}$, $\mathbf{F}$ and $\mathbf{G}$ are known at t and fulfill the obvious conformability conditions with the state and shock vectors. Note that the following derivation is valid for time-varying matrices $\mathbf{A}_t$, $\mathbf{C}_t$, $\mathbf{F}_t$ and $\mathbf{G}_t$ as long as they are known at t — the subscript is dropped for notational convenience only.

The following identity is at the heart of the algorithm:

$$E[\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y}^T] = E[\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y}^t],$$

where $\mathbf{y}^t \equiv (\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_t)$. This can easily be seen by decomposing $\mathbf{y}_{t+j}$ into

$$\mathbf{y}_{t+j} = \mathbf{F}\left[\mathbf{A}^{j-1}\mathbf{x}_{t+1} + \sum_{k=0}^{j-1}\mathbf{A}^k\mathbf{C}\boldsymbol{\varepsilon}_{t+j-k}\right] + \mathbf{G}\boldsymbol{\eta}_{t+j}$$

and noting that the errors $\boldsymbol{\varepsilon}_{t+k}$ and $\boldsymbol{\eta}_{t+k}$ are uncorrelated with $\mathbf{x}_t$ for $k > 0$ by assumption. To facilitate notation, introduce $\hat{\mathbf{x}}_{t|t} \equiv E[\mathbf{x}_t|\mathbf{y}^t]$, $\hat{\mathbf{x}}_{t+1|t} \equiv E[\mathbf{x}_{t+1}|\mathbf{y}^t]$ and $\hat{\mathbf{x}}_t^d \equiv E[\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y}^t]$, where the d in the superscript indicates that this value will be used for drawing from the posterior. Now, update the projection of $\mathbf{x}_t$ on $\mathbf{y}^t$, which is a by-product of the Kalman Filter, with the information about $\mathbf{x}_{t+1}$. Make the "news" from $\mathbf{x}_{t+1}$ orthogonal on $\mathbf{y}^t$ by introducing the innovation $\mathbf{a}_{t+1} \equiv \mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t}$ and update using the formula for updating a linear projection (as given in Hamilton (1994*b*), for example):

$$\hat{\mathbf{x}}_t^d = E[\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y}^t] = \hat{\mathbf{x}}_{t|t} + \underbrace{E\Big[(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})\mathbf{a}_{t+1}'\Big]}_{\equiv \boldsymbol{\Omega}_{xa}} \Big\{\underbrace{E[\mathbf{a}_{t+1}\mathbf{a}_{t+1}']}_{\equiv \boldsymbol{\Omega}_{t+1|t}}\Big\}^{-1} \mathbf{a}_{t+1} \quad (11)$$

Notice that $\boldsymbol{\Omega}_{t+1|t}$ is also an ingredient of the Kalman Filter and hence does not require additional computations. As for $\boldsymbol{\Omega}_{xa}$, it is given by

$$\boldsymbol{\Omega}_{xa} = E\Big[(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})\mathbf{a}_{t+1}'\Big] = E\Big[(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})\left(\mathbf{A}(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}) + \mathbf{C}\boldsymbol{\varepsilon}_{t+1}\right)'\Big] = \boldsymbol{\Omega}_{t|t}\mathbf{A}',$$

where $\boldsymbol{\Omega}_{t|t} \equiv E[(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t})']$ is again an ingredient of the Kalman Filter.

The variance of the true state $\mathbf{x}_t$ around $E[\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y}^t]$ is given by the usual formula for the mean squared error of an updated linear projection:

$$\mathbf{\Omega}_t^d \equiv E[(\mathbf{x}_t - \hat{\mathbf{x}}_t^d)(\mathbf{x}_t - \hat{\mathbf{x}}_t^d)'] = \mathbf{\Omega}_{t|t} - \mathbf{\Omega}_{xa}\mathbf{\Omega}_{t+1|t}^{-1}\mathbf{\Omega}_{xa} \tag{12}$$

Since the variables $\{\mathbf{y}_t, \mathbf{x}_t\}_{t=1}^T$ are by assumption jointly normally distributed, the conditional variance $E[(\mathbf{x}_t - \hat{\mathbf{x}}_t^d)(\mathbf{x}_t - \hat{\mathbf{x}}_t^d)'|\mathbf{x}_{t+1}, \mathbf{y}^t]$ is equal to the variance $\mathbf{\Omega}_t^d$ which is unconditional on $(\mathbf{x}_{t+1}, \mathbf{y}^t)$.

The algorithm to draw a sequence $\{\mathbf{x}_t\}_{t=1}^T$ from its distribution conditional on a sequence $\{\mathbf{y}_t\}_{t=1}^T$ is as follows:

- Initialize by setting $t := T$, $\hat{\mathbf{x}}_t^d = \hat{\mathbf{x}}_{T|T}$ and $\mathbf{\Omega}_t^d = \mathbf{\Omega}_{T|T}$, which can both be obtained from the Kalman Filter.[10]
- Draw $\mathbf{x}_t = \hat{\mathbf{x}}_t^d + \left[\mathbf{\Omega}_t^d\right]^{1/2} \boldsymbol{\chi}_t$, where $\boldsymbol{\chi}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- Set $t := t - 1$. Stop if $t = 0$ is reached. If not, update $\hat{\mathbf{x}}_t^d$ according to (11) and $\mathbf{\Omega}_t^d$ according to to (12).

5 Results

The results of this first step are summarized in table 2. Figure 3 shows the filtered series (i.e. $E[r_t|r_{p,t}, r_{p,t-1}, \ldots]$) with confidence intervals, derived from the conditional variance of the true state r_t around its conditional mean $E[r_t|r_{p,t}, \ldots]$.

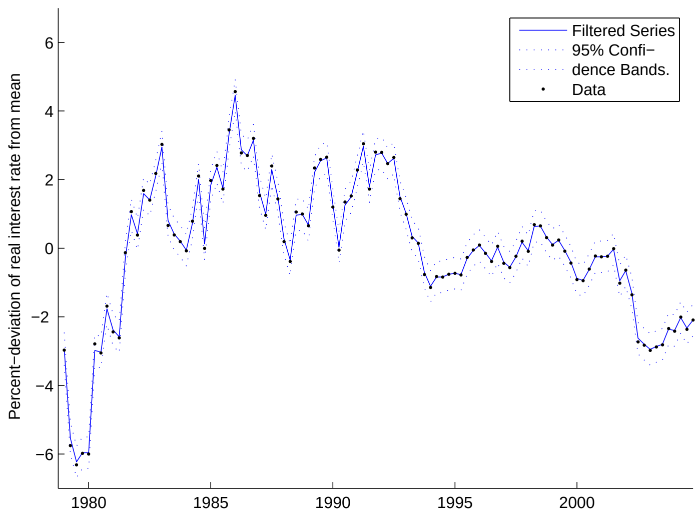
Table 2: Results of 1st-stage Estimation (MLE)

Parameter	ρ_ν	σ_ν	σ_η
Estimate	0.95	1.2	0.35

In the second step, the MCMC estimation yielded very similar results across the six sectors. Table 3 shows posterior mean and variance for the

[10]There were some computational difficulties involved in this step since the updated matrices lost the property of positive definiteness for some parameter draws. These irregularities occurred very rarely and were obviously due to computational inaccuracies when the vectors in Ω_t^d were close to collinear. The problem was fixed by adding very small numbers to the elements of the matrix to make it positive definite whenever this property was lost.

Figure 3: Results of 1st-stage Estimation: Filtered Series for r_t



model parameters and $\tilde{k}_0$. Figure 4 shows the marginal distribution of the parameters and the proportion of variance in capital and investment caused by interest-rate movements under the posterior for the sector *Engineering*. The black curves describe the distribution of the respective parameter under the prior[11] — it is easy to see that the priors are not very restrictive and have little influence on the posterior.

There are two features of the results that are unexpected: First, the rather high estimates for the depreciation rate δ are surprising. They suggest that more than 80 percent of the capital stock are obliterated every quarter in the *Chemicals* sector, for example, if the underlying story of the model is taken seriously. This result hints at a misspecification issue here and suggests to test other specifications for the investment process, e.g. specifying some form

[11]For the two bottom panels, which describe complicated functions of the deep model parameters, the calculation of the prior density is a non-trivial task — using the theorem on the transformation of probability densities, it would be possible to obtain the densities of these random variables in closed form. Since this would involve integration over several complicated functions, I opted for a simulation method. A very large number of draws (1,000,000) was taken from the independent distributions of the deep model parameters and the respective function values were calculated . Then, a simple histogram was obtained.

Figure 4: Posterior Distribution for the Sector *Engineering*

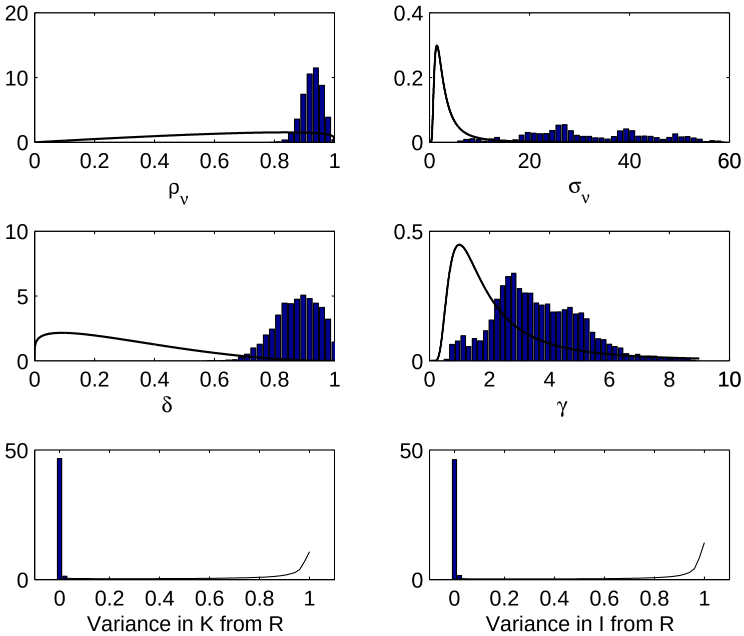


Table 3: Posterior Mean and Standard Deviation

Sector	ρ_ν	σ_ν	δ	γ	$\tilde{k}_0$
Engineering	0.926	31.845	0.876	3.619	0.253
	0.033	12.01	0.072	1.479	9.8
Food, Drink & Tobacco	0.916	46.891	0.667	10.245	0.047
	0.035	21.159	0.076	4.702	6.129
Textiles & Leather	0.927	13.91	0.766	1.1	0.247
	0.039	8.39	0.111	0.716	26.399
Fuels	0.974	24.275	0.608	2.28	-0.09
	0.014	7.456	0.086	0.805	18.095
Chemicals	0.903	32.671	0.841	4.128	0.249
	0.04	14.65	0.085	1.951	8.988
Metals	0.928	16.758	0.895	1.647	0.558
	0.03	13.047	0.067	1.309	24.983

(Means are given above in normal script size, standard deviations below in footnote size.)

of adjustment cost or lumpiness in investment.

Second, the high variability of the model parameters σ_ν and γ paired with their large covariance could be an indicator for problems with identification in this model. As described in the appendix B, however, this problem does not arise with simulated data where interest rates have a *significant* impact upon investment — note that this is not the case for either of the data series in the estimation. In fact, the following results on the importance of the interest rate for investment are not affected by the "misbehavior" of the deep model parameters.

The most striking result of the estimations is the very low fraction of variance accounted for by the interest rate — see table 4 for some percentiles of this statistic across the different sectors. Its mean is below 2.5 percent in all sectors, and only in one sector does the 99th percentile exceed 10 percent.

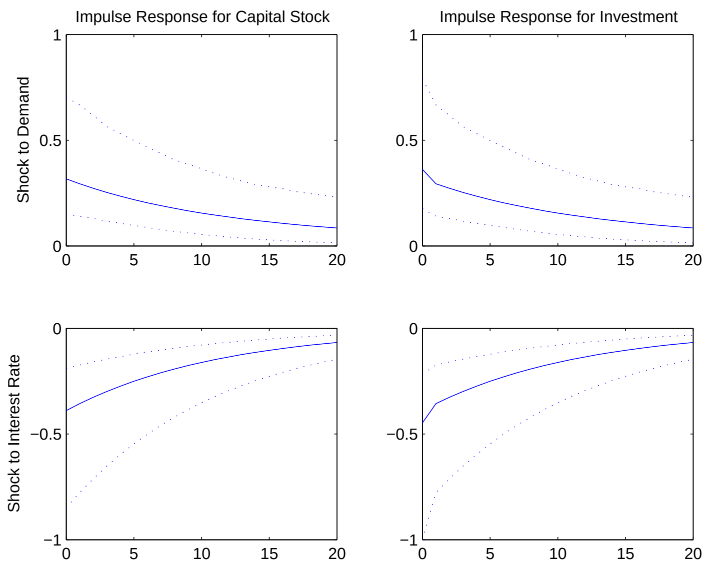
The impulse responses of investment to a one-standard-deviation shock to the demand shifter and the real interest rate are given in figure 5. The solid line depicts the mean of the analytical impulse response as given in (7) and (8) under the posterior; the dotted lines are 95-percent confidence intervals.

The draws from the posterior distribution of the hidden process $\{\tilde{\nu}_t\}_{t=1}^{104}$

Table 4: Proportion of Variance in Investment due to Interest Rate

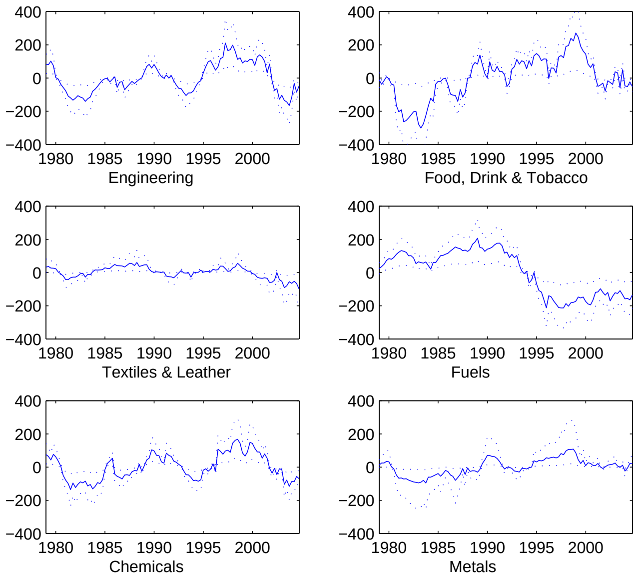
Sector	Mean	95th Percentile	99th Percentile
Engineering	0.0024	0.0081	0.0218
Food, Drink & Tobacco	0.0061	0.0371	0.0739
Textiles & Leather	0.0211	0.0714	0.1068
Fuels	0.0025	0.0096	0.0190
Chemicals	0.0041	0.0162	0.0318
Metals	0.0161	0.0627	0.0987

Figure 5: Impulse Responses for the Sector *Engineering*



are depicted in figure 6. Note that these distributions mirror both the uncertainty about the deep model parameters remaining after the estimation and the uncertainty about the hidden state that remains after any filtering exercise. For this reason, the size of the confidence intervals is time-varying, which may be striking at first for the reader accustomed to the constant conditional variances that characterize the Kalman filter under parameter certainty. It is important to bear in mind the additional dimension of parameter uncertainty that arises in a Bayesian estimation framework.

Figure 6: The Demand Shifter ν_t



Visual inspection suggests that the investment incentives faced by the firms in the different sectors are lowly correlated among each other. It seems that macroeconomic factors which affect all sectors and are not captured in interest rates (e.g. multi-purpose technologies, common components in demand, government policies) play a minor role in determining investment in the different sectors. Further econometric analysis of the joint properties of these processes or a model encompassing more than one sector would be an interesting extension of this exercise.

6 Conclusions

The estimations in this paper show that movements in the real interest rate accounted for very little variance in investment in six industrial sectors in the UK over the last 26 years. This statement can be made on a statistically sound basis in a Bayesian estimation framework — I argue that confidence intervals derived from first-order approximations employed in frequentist statistics are inferior to the Bayesian concept of the posterior distribution when this fraction is very close to zero. Moreover, the Bayesian approach offers the possibility to force model parameters into the economically sensible range and is very robust even when the dimensionality of the parameter space increases. All these are reasons to use econometric tool in the empirical evaluation of dynamic investment models.

Furthermore, the hidden investment incentives in the different sectors seem to have very little in common. This is surprising, since many macroeconomic conditions other than the interest rate should have a similar impact on all sectors. An extension of this model to investigate this issue systematically would be an interesting exercise.

Finally, one caveat is in place: The very high estimates for quarterly depreciation rates — in most sectors the posterior mean of this parameter is above 60 percent — suggest that the model employed in this paper is potentially misspecified. Future research could apply Bayesian methods to models with adjustment costs and irreversibilities in the investment process (as those of Abel & Eberly (1994) and Hayashi (1982)) or models that acknowledge the lumpiness of investment, a point that is emphasized by the empirical investment literature[12]. Another fascinating possibility is the the use of particle filters or other non-linear methods, as described by Johannes & Polson (2005) and already used by Fernández-Villaverde & Rubio-Ramírez (2004), to explore the implications of theoretical models that going beyond first-order effects.

[12]Caballero (2000) suggests this avenue of research in his survey article.

References

Abel, A. B. & Eberly, J. C. (1994), 'A unified model of investment under uncertainty', *American Economic Review* **84**(5), 1369–1384.

Barr, D. G. & Campbell, J. Y. (1996), 'Inflation, real interest rates, and the bond market: A study of UK nominal and index-linked government bond prices', *NBER Working Paper 5821* .

Bernanke, B. S. (1983), 'The determinants of investment: Another look', *American Economic Review* **73**(2), 71–75. Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association.

Caballero, R. J. (2000), 'Aggregate investment: Lessons from the previous millenium', *Working Paper* . AEA Session. In Memoriam: Robert Eisner.

Fama, E. F. & Gibbons, M. R. (1982), 'Inflation, real returns, and capital investment', *Journal of Monetary Economics* **9**(3), 297–323.

Fernández-Villaverde, J. & Rubio-Ramírez, J. F. (2004), 'Estimating dynamic equilibrium economies: Linear versus nonlinear likelihood', *Journal of Applied Econometrics* . Forthcoming.

Hamilton, J. D. (1994*a*), State-space models, *in* R. Engle & D. McFadden, eds, 'Handbook of Econometrics', Vol. 4, Elsevier Science, Amsterdam: North Holland, chapter 50.

Hamilton, J. D. (1994*b*), *Time Series Analysis*, Princeton University Press.

Hayashi, F. (1982), 'Tobin's marginal q and average q: A neoclassical interpretation', *Econometrica* **50**(1), 213–224.

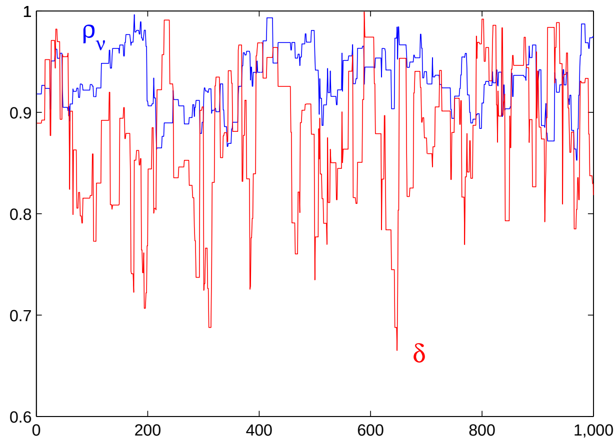
Johannes, M. & Polson, N. (2005), MCMC methods for continous-time financial econometrics, *in* Y. Ait-Sahalia & L. P. Hansen, eds, 'Handbook of Financial Econometrics', Elsevier Science, Amsterdam: North Holland.

Ljungqvist, L. & Sargent, T. J. (2004), *Recursive Macroeconomic Theory*, 2 edn, MIT Press, Cambridge and London.

A Parameter Paths in the MCMC Algorithm

Figures 7 and 8 show the path of the Markov chain for the four model parameters in the estimation for the *Engineering* sector. For the parameters ρ_ν and δ, the sensible range of parameter values is thoroughly explored — note that only 1,000 draws are shown in figure 7, whereas all 20,000 draws are shown in figure 8. For σ_ν and γ, however, the picture is different — it would be quite confident to assert that the chain has converged to a stationary distribution for *these* two parameters.

Figure 7: MCMC Path of ρ_ν and δ (Sector *Engineering*)

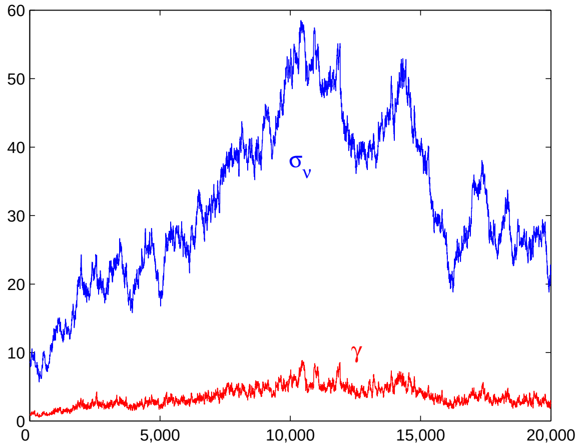


Due to their high correlation, the ratio between σ_ν and γ stays in a relatively stable range, so that other statistics of interest (as impulse responses and variance decompositions) do not vary much along the chain. The system described in (10) seems to come close to non-identifiability when the influence of interest rates on investment goes close to zero; yet, the main conclusion of the exercise — interest rates plays a minor role in determining investment — is very robust.

B Estimation with Simulated Data

The data in this simulation were generated by drawing 104 independent standard normal shock vectors $(\tilde{\varepsilon}_{\nu,t}, \tilde{\varepsilon}_{r,t}, \eta_t)'$ and feeding them into the dynamic

Figure 8: MCMC Path of σ_ν and γ (Sector *Engineering*)



linearized system described in (10). The initial state was drawn from the stationary distribution of the system. The parameters ρ_r, σ_r and σ_η were fixed at the values obtained in the first estimation stage (described in sections 4 and 5). For the remaining parameters, the following values were chosen: $\rho_\nu = 0.9$, $\sigma_\nu = 2$, $\delta = 0.1$ and $\gamma = 1.5$. These true values are marked with a black diamond in figure 9, which gives the results of an estimation that was carried out in exactly the same fashion as the estimation on the real data (described in section 4). The estimation procedure yields reasonable results. The posterior distribution does not depart far from the true values and the variance of the parameters under the posterior is modest. This indicates strongly that the model does not suffer from an identification problem.

Figure 9: Posterior Distribution in Simulation