Bayesian Approach to Estimation:
Natural Condition of Control and DSGE Models[*]

Martin Fukač[†]   Vladimír Havlena[‡]

October 2010

Abstract

This paper is written by authors from technical and economic fields, motivated to find a common language and views on the problem of the optimal use of information in model estimation. The center of our interest is the natural condition of control − a common assumption in the Bayesian estimation in technical sciences. We discuss its validity in the estimation of dynamic stochastic general equilibrium (DSGE) models in economics. We argue that the models contain implicit control variables that have observable/measurable counterparts, and they carry information that affects estimation efficiency and model adaptability. We illustrate our points on a basic RBC model.

Key words: natural condition of control, Bayesian estimation, DSGE model estimation

JEL: C11, C18

1   Introduction

Since the seminal paper by Peterka (1981), it has been well understood in the technical sciences that on the way from the Bayesian formula to the standard recursive least square method for an ARX model estimation, or the Kalman filter estimation, several assumptions about information contained in observed input and output variables have to be adopted. While such assumptions are well justified and easy to interpret in technical applications like LQG observer-based state feedback or adaptive control, the technical assumptions may be violated in some other areas like economics.

Our attention is focused on the *natural condition of control* (henceforth NCC or condition). In this way, we would like to stimulate the discussion on the proper use of the information available to econometricians and on the adaptation of theoretical model concepts to a particular estimation algorithm. We review the development of the model estimation from a conceptual Bayesian solution – resulting in a generic functional recursion on conditioned probability density functions (c.p.d.f.) – to famous Kalman filter equations, and demonstrate the loss of optimality in the case when the assumptions used for the development of the standard Kalman filter are not satisfied.

The natural condition of control is an assumption made in the control system literature. This mathematical assumption significantly simplifies the algorithm for the optimal estimation of unknown variables like parameters or state (latent) variables. The condition says that if an external observer (econometrician/statistician) is at the same time also the one who controls the system, then his control decisions, if optimal, do not provide any additional information about the state of the system, and vice versa.

In contrast to economic applications, the NCC is a credible assumption in the technical sciences. The observer and controller is one person, the system under his control is well identified and he uses algorithms that lead to optimal estimation and decisions. On the other hand in economic applications, it is always almost difficult to argue that the condition holds, because the observer (econometrician) is almost always different from the controller. The econometrician observes the real-time decisions (of workers, employers, central bankers, treasurers, etc.) with a substantial time delay. Moreover, economic models that he uses are often poorly identified.

The NCC is a mathematical assumption, and its violation is difficult to detect in the data. The problem is different from the one of model misspecification that manifests itself in residuals, shock estimates, or model implied expectations. But, if we (economic agents or econometricians) objectively know that the condition does not hold, i.e. there are observed control variables that are not explicitly included in our models (the NCC is violated), the NCC entitles us to use that knowledge in our favor. There are two effects we can expect. The first, we improve the efficiency of our estimates. The second, because the observed control variable is a result of optimal de-

cision, we can use that variable to infer the encoded underlying information to improve our own knowledge about the modelled system.

We are going to review the condition's validity for the estimation of dynamic stochastic general equilibrium (DSGE) models. We choose them because they have become the norm for an optimal policy and decision analysis in policy institutions such as central banks. At the same time, they are exactly the class of economic models for which the NCC is the most relevant. In the experiments we are going to conduct with the models, we will assume that 1.) econometrician differs from the controller, but 2.) the econometrician knows the data generating process, 3.) he also knows the technology for recovering hidden (unobserved) variables, but, and this an important twist, 4.) out of a data-rich database, the econometrician decides to use only a limited subset of observed time series, which he *assumes*, carry all relevant information. 1.)-4.) is a scenario very close to the reality.

In fact, the choice of observable variables matters. There are many decision variables in DSGE models that are implicit in the DSGE's reduced forms but the variables have direct observable counterparts. For instance, such variables may be labor income, capital income, or all kinds of fees or tax revenues. For model dynamics they may be viewed as being of a second-order importance because they do not carry any extra information for the dynamics. Output, prices, and interest rates carry the whole piece of information. But from the estimation point of view, the variables, if observed, carry an important piece of information about the optimal (control) decisions from which we can infer the believes of the other agents, and use them to improve our beliefs. By definition of the DSGE models, every decision is optimal. And by the construction of the Kalman filter, every decision variable must be present in the estimation. Otherwise the filter does not provide optimal estimates.

Our paper is closely related to the literature on DSGE models in data rich environment. Boivin and Giannoni (2006) *propose* a framework for exploiting information from a large datasets to improve the estimation of DSGE models. Our arguments go a similar direction. In comparison to the data-rich literature we provide justification why the use of all available information in estimation is a must. It is the dictate of the natural condition of control. In contrast to Boivin and Giannoni, who work with empirical relationships, we use only the information that can be linked directly to a decision process captured by the model (like the variables listed above).[1] In that respect we are also using the information springing from cross-equation restrictions.

For the convention, the paper is structured as follows. We begin by

---

[1] On the ground of NCC, the methodology proposed in Schorfheide, Sill and Kryshko (2010) is an example of an inconsistent and inefficient use of available information. The off-model variables, if relevant at all, should be used to update the information of the model states (variables, factors or whatever you wish to call it) and not be treated exogenously.

reviewing the derivation of the basic Kalman filter equations from the engineering point of view. It will help us to understand the motivation and consequences of the natural condition of control. In the third section, we show how the engineering world maps in to the world of dynamic stochastic general equilibrium models. In the fourth section, we illustrate our points on a neoclassical growth model.

2   State Estimation and Output Prediction

In engineering typically, the motivation of parameter and/or state estimation is the optimal control problem. The definition of the model is then implied by this task. Consider a discrete-time dynamic system depicted in Figure 1 with the observable/measurable input sequence u_t and output sequence y_t and some hidden variables that can be interpreted as the system parameters θ or system state x_t. The input sequence enters in a closed loop, in which the control decision is based on the system states estimates.

Figure 1: Dynamic system (DSGE model)
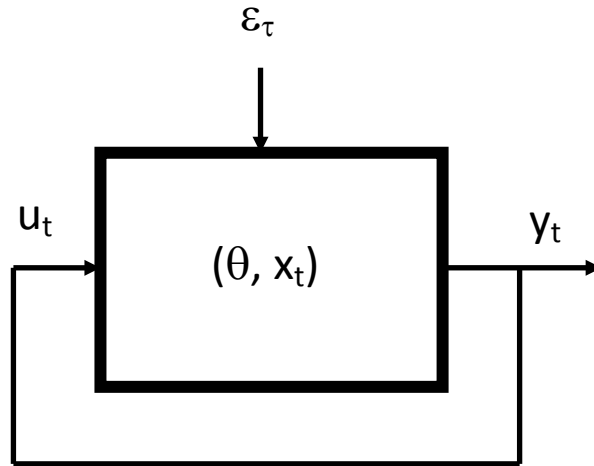
2.1 Optimal control problem

Let the sequence of input and output data observed at time interval from t_1 up to time t_2 be $D_{t_1}^{t_2} = \{u_{t_1}, y_{t_1}, ..., u_{t_2}, y_{t_2}\}$. If the initial time $t_1 = 1$, it can be omitted, i.e. $D_1^t = D^t$.

Suppose, we have observed the input-output sequence up to the time t and are looking for optimal control on T step prediction horizon with optimality criterion $\min E\left\{ J\left(D_{t+1}^{t+T}\right) \middle| D^t \right\}$. This optimization problem requires the joint probability density function $p(D_{t+1}^{t+T}|D^t)$. Using the chain rule, this c.p.d.f. can be written as

$$p(D_{t+1}^{t+T}|D^t) = p(y_{t+T}|D_1^{t+T-1}, u_{t+T})p(u_{t+T}|D_1^{t+T-1}) \times ... \times p(y_{t+1}|D_1^t, u_{t+1})p(u_{t+1}|D^t)$$

The set of c.p.d.f.s $p(y_\tau|D^{\tau-1}, u_\tau)$ for $\tau = t + 1, ..., t + T$ defines the dependence of system output y_τ on system history up to the time $\tau - 1$, and the system input at time τ. These c.p.d.f.s define the model of the system.

The set of c.p.d.f.s $p(u_\tau|D_1^{\tau-1})$ for $\tau = t = 1, ..., t + T$ is a general description of the law by which the input u_τ is generated. We will call this set of c.p.d.f.s as the *control law*. Note the information delay in the control law; while the input u_τ is applied to the system to generate its output in the τ-th period, the output y_τ is not available to calculate the control law u_τ. [2]

2.2 State Estimation

If there exists a hidden (latent) variable x_t of fixed dimension such that

$$p(x_{t+1}, y_t|D^{t-1}, x_t, u_t) = p(x_{t+1}, y_t|x_t, u_t)$$

it is called the *state of the system*. x_t constrains all the information about the system history that is relevant to predict the values $\{x_{t+1}, y_t\}$. Using the state definition above, the output model can be obtained as a marginal distribution

$$p(y_t|x_t, u_t) = \int p(x_{t+1}, y_t|x_t, u_t)dx_{t+1}$$

and the state transition model as a conditioned distribution

$$p(x_{t+1}|x_t, u_t, y_t) = \frac{p(x_{t+1}, y_t|x_t, u_t)}{p(y_t|x_t, u_t)}.$$

This reflects the fact that for the prediction of state x_{t+1}, the information about the output in the t-th period is available and should be incorporated in the optimal prediction (see the sampling scheme in Figure 2). To calculate

---

[2] In engineering applications it is typically assumed that a continuous process is observed at regular intervals $\tau = tT_s$ with sampling period T_s and the input is constant during the sampling period, i.e. $u(\tau) = u_t$ for $tT_s \leq \tau < (t+1)T_s$.

the output prediction

$$p(y_t|D^{t-1}, u_t) = \int p(y_t|u_t, x_t)p(x_t|D^{t-1}, u_t)dx_t$$

information about the state given by the c.p.d.f. $p(x_t|D^{t-1}, u_t)$ is required at each step of the recursion. That is the point when the NCC comes in to play.

Suppose the information about the state $p(x_t|D^{t-1})$ based on the data up to time $t-1$ is available. This information can be updated after a new input-output observation $\{u_t, y_t\}$ has been obtained using the Bayes formula

$$p(x_t|D^t) = \frac{p(y_t|D^{t-1}, x_t, u_t)p(x_t|D^{t-1}, u_t)}{p(y_t|D^{t-1}, u_t)} = \frac{p(y_t|x_t, u_t)}{p(y_t|D^{t-1}, u_t)}p(x_t|D^{t-1}),$$

where the properties of the state and the *natural condition of control* for the state estimation (Peterka, 1981) $p(x_t|D^{t-1}, u_t) = p(x_t|D^{t-1})$ are used to get the second term.

The NCC assumption cannot be deduced from the properties of the dynamic system itself but rather from the process of information accumulation. In the technical context, its interpretation is twofold:

1. The condition $p(x_t|D^{t-1}, u_t) = p(x_t|D^{t-1})$ says that the control variable u_t does not provide any additional information about the state of the system x_t. This assumption is valid e.g. in the framework of observer-based LQG control – incomplete information feedback – with the control variable based on state estimate $u_t = f\left(E[x_t|D^{t-1}]\right)$. In this case the control variable u_t does not provide any additional knowledge then the information contained in the data set D^{t-1}.

2. Using the equality

$$p(x_t|D^{t-1}, u_t)p(u_t|D^{t-1}) = p(u_t|x_t, D^{t-1})p(x_t|D^{t-1}),$$

the condition $p(x_t|D^{t-1}, u_t) = p(x_t|D^{t-1})$ implies that also $p(u_t|D^{t-1}, x_t) = p(u_t|D^{t-1})$. If the state-estimation and control is performed by the same subject, the system input is based only on the available data and is not modified by the state estimate, which does not provide any "new" information for the calculation of the control law.

If any additional information about the system state is available to calculate the control law, the standard Kalman filter is not optimal from the Bayesian inference/information accumulation point of view. That is why some applications in the economic literature may not fully comply with the NCC assumption: typically in multi-agent environment where individual agents operate based on different information content, the control action of

6

one agent may provide additional information to the remaining agents, i.e. $p(x_t|D^{t-1}) \neq p(x_t|D^{t-1}, u_t)$. If this additional information is not used to evaluate their optimal control strategy, their behavior is not optimal from the Bayesian inference/information accumulation point of view.

As an example, assume a statistitian observing a linear system controlled by (complete information) state feedback. Then his (noisy) observation of controlled variable $u_t = -Kx_t + e_t^u$ provides significant information about the state.

If the statistitian knows the control law K, interpreting the control variable u_t as an additional observation defined by c.p.d.f. $p(u_t|x_t) = p_{e^u}(u_t + Kx_t)$ in parallel to the observed outputs $y_t = Cx_t + Du_t + e_t^y$ defined by $p(y_t|x_t, u_t) = p_{e^y}(y_t - Cx_t - Du_t)$, the optimal *data update* step of state estimation process (Kalman filter) should cover *input update* step

$$p(x_t|D^{t-1}, u_t) \propto p(u_t|x_t)p(x_t|D^{t-1})$$

and *output update* step

$$p(x_t|D^t) = p(x_t|D^{t-1}, u_t, y_t) \propto p(y_t|x_t, u_t)p(x_t|D^{t-1}, u_t).$$

If the statistitian does not know the control law K, he is not able to incorporate this information into the state estimation process. However, if he knows that NCC is *not* satisfied[3] and he is sure that the observed control variable u_t provides additional information about the state x_t, he may try to recover this information. One of his options is adaptation of his behavior based on estimation of the control law K as an unknown parameter of the observation model $p(u_t|x_t, K)$.

In the DSGE models the rational expectations operate over the information set that is pooled by the agents. But the idea of information heterogeneity has becoming popular in the adaptive learning literature.

3   General Equilibrium Models

Now we turn our attention to the dynamic stochastic general equilibrium (DSGE) model. We work with the following (log-linear) DSGE model:

$$\Gamma_0(\theta)x_t \;=\; \Gamma_1(\theta)E_t x_{t+1} + \Gamma_2(\theta)x_{t-1} + \Gamma_3(\theta)\varepsilon_t, \tag{1}$$

where x_t is $(n \times 1)$ vector of endogenous variables (log-deviations from their steady state), and ε_t is $(k \times 1)$ vector of unobservable exogenous *iid* shocks. For simplicity in notation, we assume that $n = k$. This assumption will be relaxed in the later discussion. $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, $\Gamma_2(\theta)$ and $\Gamma_3(\theta)$ are (time invariant) matrices of structural parameters. Their elements are functions

---

[3]detection of NCC violation may be a separate topic of interest

Figure 2: Sampling from a continuous process - logic for the Kalman filter timing
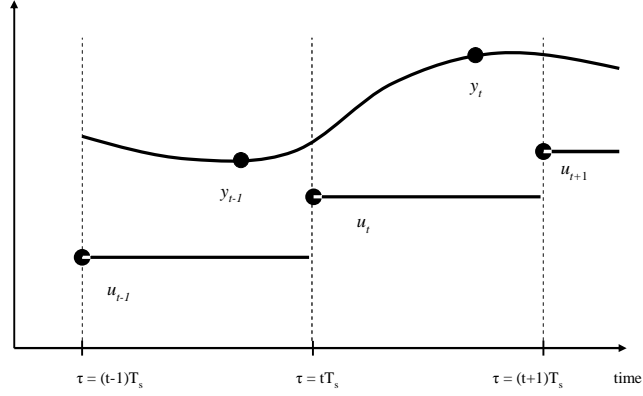
Note:

of deep structural parameters, θ. $E_t(.)$ is the rational expectation operator conditional on the model M and information available to the economic agents at time t – the information matrix is $\Omega_t \in (x_t, x_{t-1}, ..., x_0, \varepsilon_t, M)$. The structural matrices $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, $\Gamma_2(\theta)$ and $\Gamma_3(\theta)$ are such that the model has unique and stable equilibrium.

Solving for the rational expectations $E_t(.)$, model (1) has a minimum state representation

$$x_t = A(\theta)x_{t-1} + B(\theta)\varepsilon_t. \qquad (2)$$

Equation (2) characterizes the dynamic equilibrium in the reduced form. $A(\theta)$ and $B(\theta)$ are functions of Γs and through them they are functions of the deep structural parameters θ.

The states of the model x_t are linked to their observed counterparts via the measurement equation

$$y_t = Cx_t, \qquad (3)$$

where y_t is $(m \times 1)$ vector of observable variables, and C is the $(m \times n)$ (usually identity) matrix that maps the model variables in to y_t.

(2) and (3) establish together the state-space representation of (1).

Estimating (2) and (3), it is standard to assume that (i) model (1) is the accurate representation of the world and the decisions taken in it, and (ii) y_t of the dimension $m \times 1$ is the only information that the outside observer has available to estimate and evaluate x_t. Of course, in reality this is not always true, and for policy oriented models such assumptions are unjustified. Here is one example for all.

DSGE models include implicit definitions of variables that have observable counterparts, e.g. income tax or capital (property) tax revenues. They almost never explicitly appear in DSGE models. Tax rates affect dynamics indirectly via resource allocation, but the tax revenues per se never appear in the minimum-state representation as they do not bring any additional information about the dynamics. Unless we are interested in estimating the model, carrying around the capital tax revenues as an extra variable is not necessary. But if we estimate the model, we must include it.[4]

In the language of the previous section, x_t are model states and (2) can be view as a model of control in closed loop (or full-state control). The hidden (implicit) control variables in DSGE models can be contemporaneously expressed as follows

$$\Gamma_0(\theta)x_t = \Gamma_0(\theta_1)x_t + u_t.$$

$u_t = \Gamma_0(\theta_2)x_t$ is the cumulative effect of the structural parameters $\Gamma_0(\theta_2)$.

If there is u_t that it is observed, we have to extend the observation equation (3) to inform the estimates of θ and x_t. To facilitate that fact, we augment measurement equation (3) to

$$x_t = A(\theta)x_{t-1} + B(\theta)\varepsilon_t \tag{4}$$

$$\begin{bmatrix} u_t \\ y_t \end{bmatrix} = \begin{bmatrix} \Gamma_0(\theta_2) & 0 \\ 0 & CA(\theta) \end{bmatrix} x_{t-1} + \begin{bmatrix} I & 0 \\ 0 & CB(\theta) \end{bmatrix} + \begin{bmatrix} \varepsilon_t^u \\ \varepsilon_t \end{bmatrix} \tag{5}$$

Let us illustrate the benefit of such a state-space specification on a simple example.

4   Illustration

We run a Monte Carlo experiment to demonstrate our point that optimally using the whole disposable information improves estimates of latent (state) variables. This is done first under the assumption of information pooling but its partial use, and then under the assumption that one of the control

---

[4]For similar reasons, the demand for money is excluded from the model, because it does not carry any additional information about the inflation rate and output gap. But when the output gap is not observable, then money has additional forecasting power.

variables is generated based on a more precise information about the state variables.

We use a simple real business cycle model to generate artificial data of private consumption, hours worked, investment, consumption tax receipts, and of disposable income. Those series form our set of disposable information. In the first experiment, we assume that an econometrician (observer) uses only two series out of the set. In the following experiments we gradually expand the information set that the econometrician utilizes and observe the efficiency and consistency improvement. But let us define the real business cycle model first.

In the model, there are two sectors – household and firm.

Households maximize their expected life-time welfare $E_0[\sum_{t=0}^{\infty} \beta^t \frac{(C_t+H_t)^{1-\sigma}}{1-\sigma} + \xi \log(1-L_t)]$ subject to a budget constraint $w_t L_t + (1 - r_{t-1} - \delta)K_{t-1} + T_t = (1 - \tau_c)C_t + K_t$. The household's welfare derives from consumption C_t and leasure $1 - L_t$. The level of consumption is fuelled by the habit H_t, which depends on the past consumption and an *iid* habit shock: $H_t = \phi C_{t-1}e^{\varepsilon_t}$, with $\phi \in (0,1)$ and $\varepsilon_t \sim N(0, \sigma_c^2)$. Time spend by work L_t brings disutility but it is compensated by hourly wage w_t. The consumption is taxed by the government at the rate of $\tau_c \in (0,1)$. Next, the household is the only owner of physical capital K_t in the economy, which is, together with labor, a factor of production. Firms rent the capital and pay the households the interest r_{t-1} in return. The physical capital depreciates over time by the rate $\delta \in (0,1)$. The household further receives the lump-sum transfers T_t from the government, who operates on the balanced budget. The household discounts the future by the discount factor $\beta \in (0,1)$. $\sigma > 0$ is the measure of household's risk aversion.

Firms maximize their profits $\Pi_t = Y_t - r_t K_{t-1} - w_t L_t$ by optimally hiring labour and capital to produce the consumption goods Y_t. They use a Cobb-Douglas technology: $Y_t = A_t K_{t-1}^{\alpha} L_t^{1-\alpha}$. A_t is the total factor productivity and follows a log-linear AR(1) process: $\log A_t = \rho \log A_{t-1} + \varepsilon_t^A$. $\varepsilon_t^A \sim N(0, \sigma_A^2)$ and we interpret is as the productivity shocks. $\alpha \in (0,1)$ is the share of capital in production.

In equilibrium, all (labour, capital, and consumption goods) markets clear in this economy. The dynamic equilibrium of the model economy is then characterised by four equations: (i) the Euler equation for consumption, (ii) labour demand, (iii) resource constraint, and (iv) exogenous supply of

technology.

$$\left(\frac{C_t + \phi C_{t-1}e^{\varepsilon_t}}{E_t\{C_{t+1}\} + \phi C_t}\right)^{-\sigma} = \beta(1 - \delta - \alpha A_t K_{t-1}^{\alpha-1} L_t^{\alpha}) \qquad (6)$$

$$\xi\frac{1 - L_t}{C_t^{\sigma}} = \left(\frac{1 - \tau_c}{1 - \alpha}\right)\frac{L_t^{\alpha}}{A_t K_{t-1}^{\alpha}} \qquad (7)$$

$$A_t K_{t-1}^{\alpha} L_t^{1-\alpha} = C_t + K_t - (1 - \delta)K_{t-1} \qquad (8)$$

$$\log A_t = \rho \log A_{t-1} + \varepsilon_t^A \qquad (9)$$

We use the following values to parametrize the equilibrium: $\alpha = 0.60$, $\beta = 0.97$, $\xi = 1$, $\sigma = 3$, $\tau_c = 0.2$, $\delta = 0.01$, $\phi = 0.5$, $\rho = 0.9$, $\sigma_c = \sigma_A = 0.01$.

The model is solved using the methodology proposed by King, Plosser and Rebello (1988). First, the model steady state is computed. Second, (6) - (9) are log-linearized around the steady state, and we obtain the model in the form of (similar to 1). Third, the log-linear model is solved for the rational expectations $E_t(.)$. The result is the state equation (2). There are four state variables $\begin{bmatrix} C_t & K_t & L_t & A_t \end{bmatrix}$ and two structural shocks $\begin{bmatrix} \varepsilon_t & \varepsilon_t^A \end{bmatrix}$.

The model is used to simulate the set of disposable information. The set consists of the measures of private consumption $\bar{C}_t$, hours worked $\bar{L}_t$, gross private investment $\bar{I}_t$, sales tax receipts $\bar{T}_t$, and disposable income $\bar{D}I_t$. The first two variables are direct counterparts of the state variables C_t and L_t. The other three are definitions implicitly included in (6)-(9). They are functions of the model variables (states). The gross private investment is defined as $\bar{I}_t = K_t - (1-\delta)K_{t-1}$, consumption tax receipts $\bar{T}_t = \tau_c C_t$, and the disposable income $\bar{D}I_t = A_t K_{t-1}^{\alpha} L_t^{1-\alpha}$. All the variables are measured with an error that is *iid*. We cannot measure any direct counterparts of the physical capital stock and the total factor productivity. Those variables remain latent states.

4.1 Monte Carlo Experiment

In the first experiment, we assume that an econometrician (observer), who wants to estimate the model (6)-(9), decides to use the information contained in the measures of consumption $\bar{C}_t$ and hours worked $\bar{L}_t$. He knows the structural model and its parametrization, and he wants to learn the latent states K_t and A_t.

The assumptions of this experiment resemble a set up common to economic applications. (6)-(9) contains four states $\{C_t, L_t, K_t, A_t\}$. The econometrician knows the disposable data, but he decides to use the measures that are naturally the closest to the model variables. Because there are no direct counterparts of K_t nor A_t in his database, he estimates them as latent. The results are plotted in Figure 3. The solid (pink) lines are the actual (simulated and known to us) series of K_t (top panel) and A_t (lower panel). The

widest (blue) interval corresponds in both panels to the uncertainty of the estimates. We see that very often the estimate of capital or technology is indistinguishable from zero.

The econometrician's choice of observable variables results in the loss of efficiency and consistency. The confidence intervals are wide but at the same time they do not often cover the actual series.

Next, the econometrician exploits the disposable information. The econometrician thinks carefully and realizes that the whole disposable information in fact is structurally linked to his model and can help to inform his estimates. The light intervals in Figure 3 show the gain in efficiency and consistency. While the observation on disposable income and investment are very informative about the level of capital and productivity, the tax revenues are found to contribute only marginally.

Prior to any estimation we can evaluate the Fisher information matrix in order to see, in principle, how much new information we can expect to obtain when asking the data. It is a coherent way to summarize and analyze the information content for example presented in Figure 3. The Fisher Information matrix can help us to prioritize among variables we consider to select from the set of available information, which may be particularly helpful if we happen to have a constraint on the computational power available.

4.2   Estimated Model

Now we repeat the above experiments with the US data. In contrast to the prior analysis, we will see that empirically consumption and hours worked sufficiently inform the estimates of capital and technology. Adding the observations on the fixed private investment does not add to the efficiency very much.

The data used in this section are taken from the Fred (the Federal Reserve Economic Data managed by the Federal Reserve Bank of St. Louis[5]). We use the annual series of the real personal consumption expenditures (mnenomic: PCECCA96), annual series of real private fixed investment (mnemonic: FPICA), and annualized series of total hours worked, which is the product of monthly seasonally adjusted series of average weekly hours worked in private industries (mnemonic: AWHNONAG) and of the total non-farm payrolls (mnemonic: PAYEMS). Per capita terms are taken with respect to the total civilian labor force (mnemonic: CLF16OV). There are two time spans we consider. The first one is relevant for consumption that spans from 1949 to 2009. The second time span is for hours worked and investment that we observe from 1965 and 1967, respectively, to 2009. The model is estimated on the relevant samples between 1950 and 2009.

Because model (6)-(8) is without nominal rigidities, we treat it as a

---

[5]Web page *http://research.stlouisfed.org/fred2/*.

Figure 3: Estimates of capital stock and labor productivity (MC experiment)
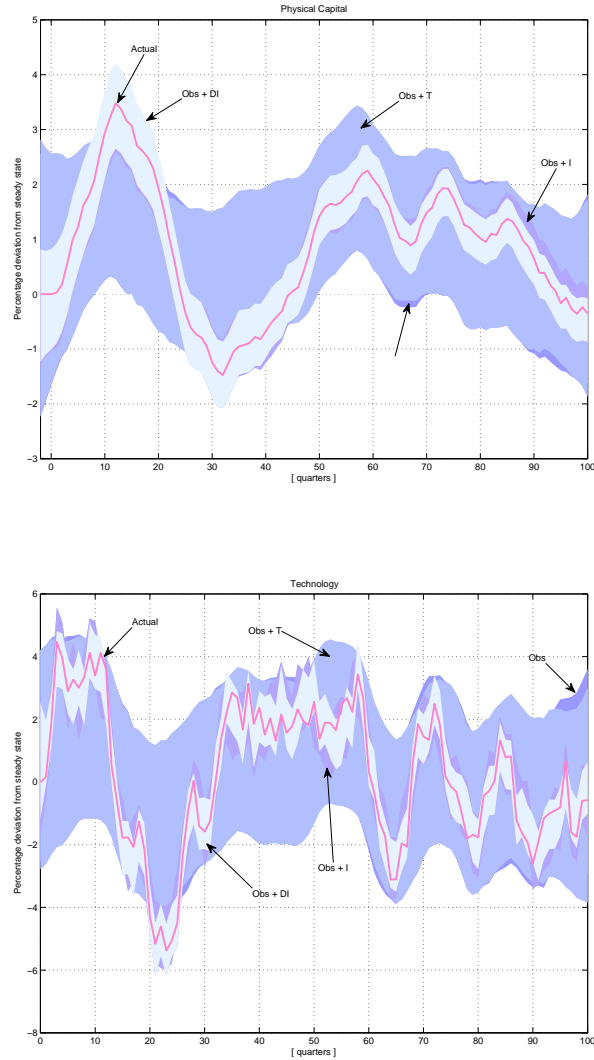
Note: The graph presents the results of a Monte Carlo experiment with model (6)-(9) as parametrized in the text. The shocks are drawn from *iid*s. In both panels, the solid line is the actual series (capital stock - top panel; technology - lower panel). The shaded bands around the actual series are the estimated 2std (smoothed) confidence intervals conditioned on a set of observables. The baseline information set (Obs)– consumption and hours work observed – and the baseline set extended for the consumption tax receipts (Obs+T) yield the two widest confidence intervals. The baseline set extended for disposable income (Obs+DI) or investment (Obs+I) provide the most narrow confidence intervals and pin down the level of the actual states very precisely.

growth model and estimate it on an annual frequency. Because of the non-stationary nature of the actual data, we have to have a stochastic trend

within the model. Therefore, instead of (9) we now assume that the technology A_t is labor augmenting and follows the first-difference stationary process with drift:

$$\Delta \log A_t = (1 - \rho)\Delta \bar{A} + \rho \Delta \log A_{t-1} + \varepsilon_t^A. \tag{10}$$

$\Delta \bar{A} > 0$ is the drift term, which sets the economy on an exogenous but balanced growth path. Both capital and consumption grow at that rate in the long run.

The transitory parameters $\{\phi, \rho\}$ and the variances $\{\varepsilon_t^c, \varepsilon_t^a\}$ are estimated using the maximum likelihood. The other parameters are kept fixed at their parametrized values mentioned above. We will not report the results here but instead we show the estimates of the capital stock K_t and labor augmenting technology A_t.

In the top two panels of Figure 4, we compare the confidence intervals for K_t and A_t. Explicitly appreciating that a capital accumulation happens through investment, we see that adding the investment among our observable variables improves our confidence by about 2 percent. The standard errors become lower.

The model's good (in-sample) predictive power for investment explains the marginal gain in efficiency. The graph in the third panel of Figure (4 compares the model implied investment (when treated as latent) to the actually realized (observed) data. Clearly, the private consumption expenditures and hours worked data carry enough information about investment at thus the capital stock and technology.

The same conclusion can be made from the (out-of-sample) forecasting exercise. The investment as an observable has a weak information content about the transitory parameters, variances and the mean value of K_t and A_t. In Figure 5 we report the results on pseudo real-time forecasts. The 5-year rolling forecasts coincide under the two setups.

Figure 4: Estimates of capital stock and labor productivity (US data)
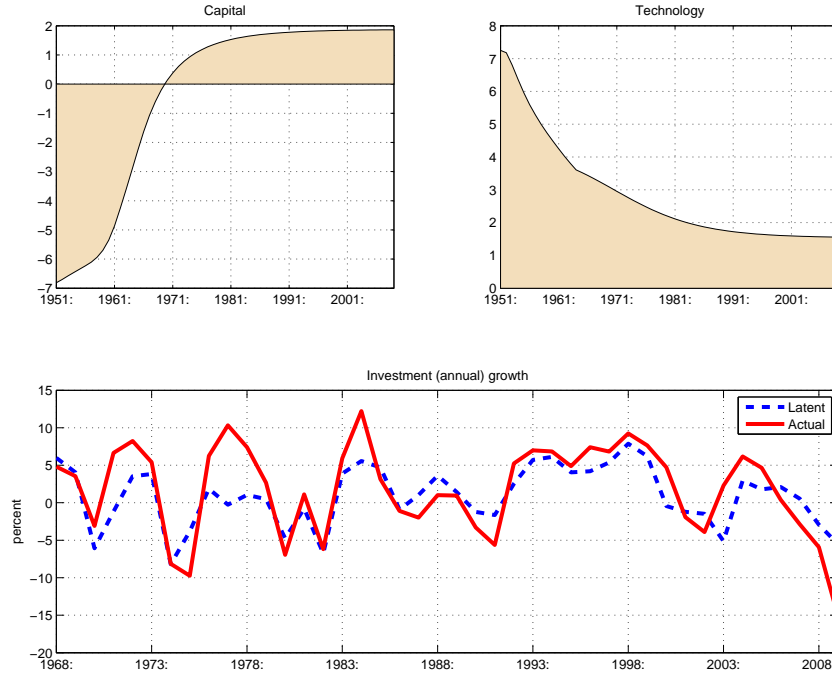
Note: The top two panels show the relative efficiency of the capital stock estimate (top left) and labor augmenting technology estimated (top right) when (i) the information on the growth of consumption and hours worked is used (model 1), and (ii) when that information is extended with the investment growth (model 2). The shaded areas are computed as $100(std(X_{t,model1})/std(X_{t,model2}) - 1)$. Positive values mean that the model with more information (model 2) outperforms the model with less information (model 1). This is the case of the estimate of labor augmenting technology over the whole sample period 1951 - 2009. In case of the capital stock, model 2 starts to outperform after the year 1971 when we accumulate enough new information from hours worked and investment that we start observing in 1965 and 1967, respectively.

The lower panel plots the estimate of the latent investment growth implied by the model 1 (dashed line). The solid line is the actual observed series used in the estimation of model 2. From the graph we see that the model structure (6)-(8) and (10) has a strong predictive power about investment.

Figure 5: Pseudo real-time forecast (US data)
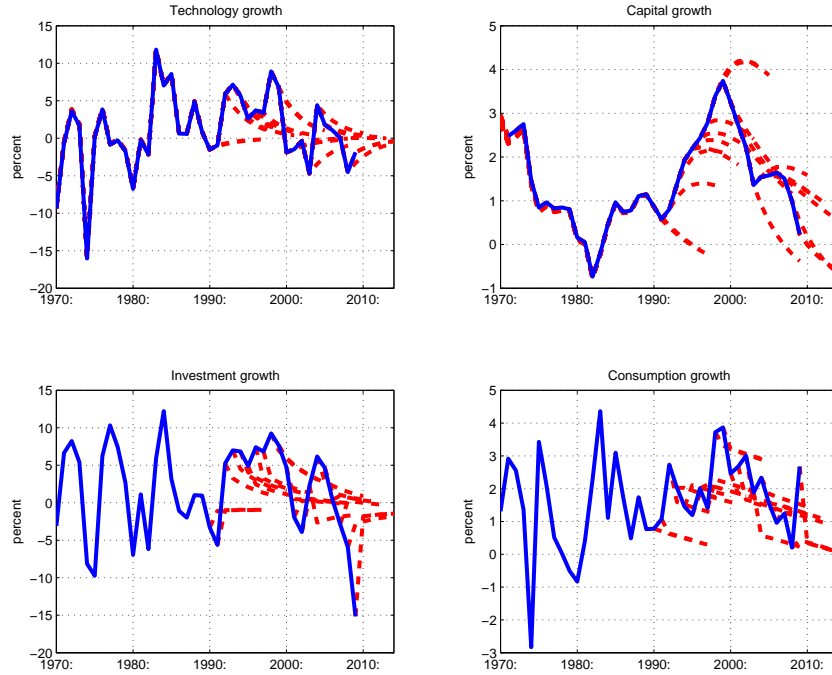
Note: The pseudo real-time forecast exercise starts in 1990, and we roll 5-year forecast window through the end of our sample in 2009. The solid line is the actual or smoothed estimate available in 2009, and the dash lines are the rolling 5-year forecasts. The forecasts based on both model 1 (consumption and hours worked observed) and model 2 (including investment growth) exhibit the same central tendencies. That is why we only report the results based on model 1. The two forecasts differ in their efficiency.

5 Final Remarks

We reviewed the basic derivation of Kalman filter equations with the focus on the role of the natural condition of control. We were interested what this condition implies for the estimation of DSGE models used in economics. We provided a theoretically consistent justification for the use of all available (observable) information that can be structurally linked to the model. Under the assumption of information pooling, we showed that that leads to a significantly improved efficiency.

The NCC can provide an alternative structural perspective for the DSGE model developers. The model may be well specified but the NCC still can be

violated. It is because the condition does not deal with the model structure per se but with the flow of information in it.

We would like to look at the possible avenues for formal testing of the NCC. We would like to use them for the empirical assessment of decision rules. DSGE models consist of optimal decision (control) rules, thus, in principle, each equation can be subject to the testing.

In the future work we also relax the assumption of information pooling, and look at the case of an agent with a significant market power and private information. If the NCC should hold the remaining market players should try to infer the private information encoded in the decisions of the dominant player and adapt to it. The NCC then leads to the adaptive learning arguments.

References

Boivin J., and M. Giannoni (2006), "DSGE Models in a Data-Rich Environment," *NBER Working Paper* No.12772.

King R. G., Plosser Ch., and Rebelo S. T. (1988), "Production, growth and business cycles : I. The basic neoclassical model," *Journal of Monetary Economics*, vol. 21(2-3), pages 195-232.

Peterka, V. (1981), "Bayesian Approach to System Identification " in P. Eykhoff (ed) *Trends and Progress in System Identification* (Pergamon Press), Oxford. Also available at *http://moodle.utia.cas.cz/moodledata/4/peterka.pdf*