ISSN 1936-5330

THE FEDERAL RESERVE BANK of KANSAS CITY
ECONOMIC RESEARCH DEPARTMENT

Note on the Role of Natural Condition of Control in the Estimation of DSGE Models

Martin Fukac and Vladimir Havlena

August 2011
RWP 11-03

RESEARCH WORKING PAPERS

Note on the Role of Natural Condition of Control in the Estimation of DSGE Models*

Martin Fukač[†] Vladimír Havlena[‡]

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, which may be violated in economic applications. In estimating dynamic stochastic general equilibrium (DSGE) models, typically only a subset of endogenous variables are treated as measured even if additional data sets are available. The natural condition of control dictates the exploitation of all available information, which improves model adaptability and estimates efficiency. We illustrate our points on a basic RBC model.

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

JEL: C11, C18

*This version: July 2011. Comments are welcome. This paper was prepared for the Conference in honor of Osvald Vašiček, held at Masaryk University in Brno, Czech Republic, 22-23 October 2010. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

[†]Federal Reserve Bank of Kansas City, 1 Memorial Drive, Kansas City, Missouri 64198; e-mail: *first.lastname* kc.frb.org.

[‡]ČVUT and Honeywell; email: *first.lastname*honeywell.com

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 must 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). 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 particular estimation algorithms. We review the development of 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. We 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 that simplifies the algorithm for the optimal estimation of unknown variables like parameters or state (latent) variables using the Kalman Filter. The condition says that if an external observer (econometrician/statistician) simultaneously observes and controls the system, then his control decisions, if optimal, do not provide any additional information about the state of the system, and vice versa.

The violation of NCC is difficult to detect in the data. It may be more of an argument than a directly testable hypothesis. The problem is different from that of model misspecification which manifests itself in residuals, shock estimates, or inconsistent, model implied expectations. But if econometricians and economic agents with a significant market power objectively know that the condition does not hold (i.e. there are observed control variables that are not explicitly included in their models while they should be and thus the NCC is violated), the NCC entitles them to use that knowledge in their favor.

In contrast to many economic applications, the NCC is a credible assumption in the technical sciences. The observer and controller are 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 almost always 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 (about tax revenues, production, consumption or prices) with a substantial time delay.

To avoid this problem, we assume that the models here describe economic agents with significant market power. Natural candidates for the controller-observer in economics are policy institutions like monetary or fiscal authorities which observe the markets' behavior and, most importantly, have effective tools to influence them. The NCC does not apply to economic agents in perfectly competitive markets because their size prevent their behavior from impacting the aggregate markets. The condition may apply to the abstract concept of a representative agent (which we do not explore here), or agents with a significant market power such as policy institutions.

We 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. At the same time, they are exactly the class of economic models for which the NCC is the most relevant because they capture optimal decisions.

There are two direct implications of the NCC on DSGE models. First, we can improve the efficiency of our estimates. Second, because the observed control variable is a result of optimal decisions in these models, we can use that variable to infer the encoded underlying information to improve our own knowledge about the modeled system. The performance of our own model can benefit if the underlying model of the particular optimal decision is of a better quality than ours.

We show that the choice of observable variables matters. There are many decision variables that are implicit in the DSGE models, but the variables have direct observable counterparts such as labor income, capital income, or all kinds of fees or tax revenues. For model dynamics, they are of second-order importance because they do not carry any extra information for the aggregate dynamics, because output, prices, and interest rates carry the entire set of information. But from the estimation point of view, the variables, if observed, carry an important piece of information from which we can infer the beliefs of the other (representative) agents with higher precision and use them to improve our own (policy authority) beliefs. Every decision is a DSGE model is, by definition, optimal. And by the construction NCC, every decision variable must be present in the estimation. Otherwise the Kalman filter does not provide optimal estimates.

The NCC provides a theoretical explanation and support, for example, to the literature on the choice of observable variables (Guerron-Quintana, 2010), or DSGE models in a data rich environment (Boivin and Giannoni, 2006). Guerron-Quintana (2010) addresses the question of "why one should be concerned with the choice of observable?" He experiments with different sets of observables, and on a standard New Keynesian model he shows the effects that their choice have on the parameter estimates and overall model dynamical behavior. At first sight his approach may appear as data mining, because it is a very data intensive analysis, but in the light of our argument, Guerron-Quintana exploration and findings may be justified by natural-condition-of-control arguments.

Our arguments also go a similar direction like in Boivin and Giannoni (2006) who propose a framework for exploiting information from a large datasets to improve the estimation of DSGE models. 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.[1] In that respect we are also using

[1] From the logic of NCC, the methodology proposed in Schorfheide, Sill and Kryshko (2010) may be viewed as an inefficient use of available information. The off-model variables, if relevant at all, should be used to update the information about the model states (variables, endogenous factors), instead of being treated exogenously.

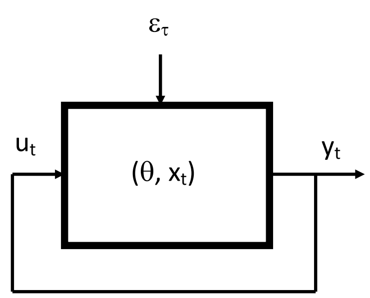
the information springing from cross-equation restrictions.

The rest of the paper is structured as follows. In the next section, we review the derivation of the basic Kalman filter equations from an engineering perspective. 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, the typical 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 (e.g., 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 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*. 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

[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 $tT_s \le \tau < (t+1)T_s$.

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 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 at which 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 than 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.

3 General Equilibrium Models

Now we turn our attention to the dynamic stochastic general equilibrium (DSGE) models. Their (log)linear form is

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

where x_t is a $(n \times 1)$ vector of endogenous variables (log-deviations from their steady state), and ε_t is a $(k \times 1)$ vector of unobservable exogenous *i.i.d.* shocks. For a simple 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 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

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

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

$$y_t = Cx_t, \tag{3}$$

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 into y_t.

Equations (2) and (3) establish together the state-space representation of the original model (1).[3] When estimating (2) and (3), it is standard to assume that (i) model (1) is a reasonable representation of the world and the decisions taken in it, and (ii) y_t is the only information that the outside observer has available to estimate and evaluate x_t.

If an external observer does not use all available information, the NCC is violated and the Kalman filtering may not be optimal, which sacrifices the estimation efficiency of parameters and unobservable variables. We consider two instances in which the NCC is violated.

3.1 Learning from others

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 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

[3] It is useful to note at this point that if $n = m$ the state-space model can be written as a finite order VAR. If $n > m$, the model can be written as an infinite order VAR. If $n < m$, the state space model is stochastically deficient.

evaluate their optimal control strategy, their behavior is not optimal from the Bayesian inference/information accumulation point of view.

As an example, assume a statistician 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 statistician 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

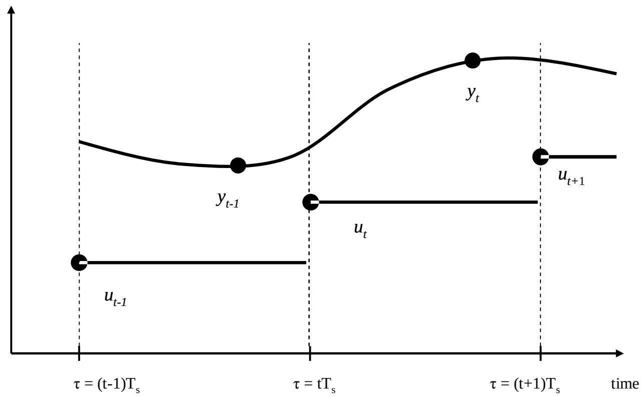
$$p(x_t|D^t) = p(x_t|D^{t-1}, u_t, y_t) \propto p(y_t|x_t, u_t)p(x_t|D^{t-1}, u_t).$$

If the statistician 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 are not satisfied [4] 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 $p(u_t|x_t, K)$.

The NCC adds on an additional dimension to adaptive learning. The basic Kalman filter algorithm already utilizes the information from one's own past prediction errors. In contrast to learning from one's own errors, the violation of NCC calls for learning from the decisions and errors of others.

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

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



3.2 Unobserved-observed variables

The state equation (2) can be viewed as a model of control in closed loop (or full-state control). The latent endogenous variables x_t can be split in to two parts:

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

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

If there is u_t that is observed, we have to extend the observation equation (3) to inform the estimates of θ and x_t, similarly like in the previous subsection. Therefore, we augment measurement equation (3) to the form of

$$\begin{aligned} x_t &= A(\theta)x_{t-1} + B(\theta)\varepsilon_t && (4) \\ \begin{bmatrix} u_t \\ y_t \end{bmatrix} &= \begin{bmatrix} \Gamma_0(\theta_2) & 0 \\ 0 & CA(\theta) \end{bmatrix} x_{t-1} + \begin{bmatrix} I & 0 \\ 0 & CB(\theta) \end{bmatrix} + \begin{bmatrix} \varepsilon_t^u \\ \varepsilon_t \end{bmatrix} && (5) \end{aligned}$$

An example of a variable u_t in DSGE models can serve income tax, consumption tax, or capital (property) tax revenues. They almost never explicitly appear in DSGE models. These variables are determined by a passive fiscal policy. Tax rates affect dynamics indirectly via resource allocation, but the tax revenues per se never explicitly appear in the minimum-state representation because they do not bring any additional information about the aggregate dynamics. From the estimation perspective, including observations on tax revenues may be important. It is very often the case that some of the variables in the minimum-state model do not have an observable counterpart (e.g. capital stock or output gap). Then to minimize the uncertainty around their estimates, any information on capital tax revenues is very useful because it is structurally linked to the unobserved capital stock and thus helps to effectively infer its level.[5]

[5] Similarly, the demand for money may add additional power to output gap forecasts. It does not carry any additional information about the inflation rate and output gap, because money demand on those factors. But exactly the very same reason dictates to include money demand among observable variables if output gap is unobservable.

4 Illustration

This section illustrates our point that using the whole disposable information may improve models' adaptability and estimation efficiency. We use a simple real business cycle (RBC) model to generate artificial data of private consumption, hours worked, investment, consumption tax receipts, and of disposable income, which form our set of disposable information. First, we assume that an econometrician (observer) uses only two series out of the complete information set. Next, we gradually expand the set that the econometrician utilizes. Then a similar exercise is repeated on actual data.

The RBC model is comprised of two sets of agents – household and firm. Households maximize their expected lifetime 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 $w_t L_t + (1 - r_t - \delta)K_{t-1} + T_t = (1-\tau_c)C_t + K_t$. The parameter $\sigma > 0$ is the measure of household's risk aversion, and the parameter $\beta \in (0,1)$ is the time discount factor. The household's welfare derives from consumption C_t and leisure $1 - L_t$. The level of consumption is fueled by the habit H_t, which depends on the past consumption and an *i.i.d.* 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 spent by work L_t causes disutility but it is compensated by the hourly real wage w_t. The consumption is taxed by the government at the rate of $\tau_c \in (0,1)$. 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 in return, but 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, which operates on a balanced budget.

Firms maximize their profits $\Pi_t = Y_t - r_t K_{t-1} - w_t L_t$ by optimally hiring labor and capital to produce the consumption good Y_t using 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: $\log A_t = \rho \log A_{t-1} + \varepsilon_t^A$. The exogenous shock $\varepsilon_t^A \sim N(0, \sigma_A^2)$ and we interpret it as the productivity shocks. $\alpha \in (0,1)$ is the share of capital in production.

In equilibrium, all (labor, capital, and consumption goods) markets clear.

The dynamic equilibrium is characterized by the Euler equation for consumption, labor demand, resource constraint, and the exogenous supply of technology.

$$\begin{aligned}
\left(\frac{C_t + \phi C_{t-1} e^{\varepsilon_t}}{E_t\{C_{t+1}\} + \phi C_t}\right)^{-\sigma} &= \beta(1 - \delta - \alpha A_t K_{t-1}^{\alpha-1} L_t^{\alpha}) && (6)\\
\xi \frac{1 - L_t}{C_t^{\sigma}} &= \left(\frac{1-\tau_c}{1-\alpha}\right) \frac{L_t^{\alpha}}{A_t K_{t-1}^{\alpha}} && (7)\\
A_t K_{t-1}^{\alpha} L_t^{1-\alpha} &= C_t + K_t - (1-\delta) K_{t-1} && (8)\\
\log A_t &= \rho \log A_{t-1} + \varepsilon_t^A && (9)
\end{aligned}$$

We use the following values to parameterize 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 Rebelo (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 (1). Finally, the log-linear model is solved for the rational expectations $E_t(.)$. The result is the state equation (2), in which there are four endogenous variables $\begin{bmatrix} C_t & K_t & L_t & A_t \end{bmatrix}$, three of which are truly state variables,[6] and there are two structural shocks $\begin{bmatrix} \varepsilon_t & \varepsilon_t^A \end{bmatrix}$.

Having the model, we 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 $\bar{DI}_t$. We assume that the measures of these variables are published by a statistical office. We do not measure any direct counterparts of the physical capital stock and the total factor productivity. Those variables remain latent states in the exercise.

The first two observed 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's endogenous variables. The gross private investment is defined as $\bar{I}_t = K_t - (1-\delta)K_{t-1}$, consumption tax receipts

[6] $[C_{t-1}, K_{t-1}, A_{t-1}]$ are the truly state variables. They form the minimum-state-variable solution to the DSGE model. The equilibrium level of L_t follows from the marginal rate of substitution between work and consumption, which is an intratemporal/static relationship.

are $\bar{T}_t = \tau_c C_t$, and the disposable income is equal to equilibrium production $\bar{DI}_t = A_t K_{t-1}^{\alpha} L_t^{1-\alpha}$. All the variables are measured with an error that is *i.i.d.*

4.1 Monte Carlo Experiment

We assume that an econometrician (observer), who wants to estimate the model (6)-(9) decides to use the information contained only 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 estimate the latent states K_t and A_t.

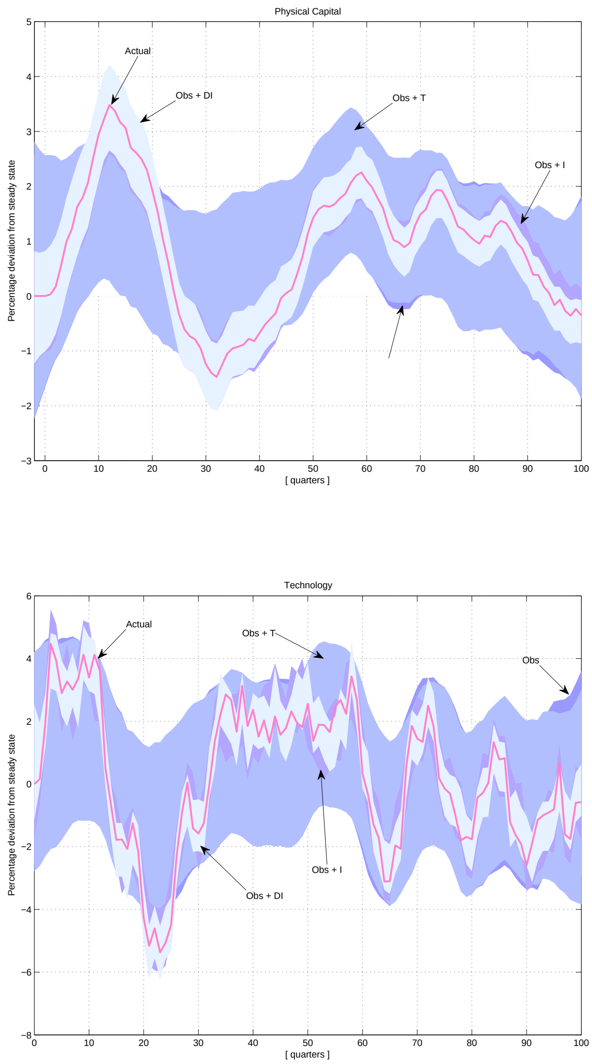
The assumptions of this experiment resemble a set up common to economic applications. The econometrician knows the disposable data, but he decides to use the measures that are naturally the closest to the model variables. Because there are no direct counterparts of K_t nor A_t in his database, he treats them as latent and estimate them. The structural parameters are known to him and thus the econometrician only seeks K_t and A_t. The Kalman filter can deliver their optimal estimates.

The estimation results of this experiment 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.

The econometrician's choice of observable variables results in the loss of efficiency and consistency. We see that very often the estimate of capital or technology is indistinguishable from zero. The confidence intervals are wide but at the same time may not include the actual series.

In the next step, the econometrician exploits the disposable information a bit more. He realizes that the statistical office provides more data that is structurally linked to his model and they can help to inform his estimates. The light intervals in Figure 3 show the gain in efficiency and consistency when the disposable income is introduced in the set of observable variables, data on investment are similarly informative. The confidence intervals for the capital stock and technology estimates shrink. The estimates become statistically significant and more closely match the actual underlying trajectory of the latent variables. The tax revenues contribute only marginally to the estimates accuracy.

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 parameterized 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.

This Monte Carlo experiment also illustrates why the violations of the NCC differ from the problem of model misspecification and therefore is hard to detect. Unlike model misspecification, the violation of the NCC does not have a clear manifestation in the estimation outcome, e.g. shock estimates are not *i.i.d.* The Kalman filter optimally process data information. It always provides optimal estimates of all latent variables. For example, all confidence intervals plotted in Figure 3 are based on a well specified model and well behaved estimates so more of them can be dismissed as inefficient or inconsistent. And still some of the data sets well outperform others in estimation.

In practice, it is difficult to make *ad hoc* claims which of available data add the most efficient information, but the model structure may help with the inference. Prior to any estimation we can evaluate the Fisher information matrix. We may infer how much new information we can expect to obtain when asking a particular set of 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 available computation power. In contrast to the selection criteria proposed in Guerron-Quintana (2010), the analysis of the Fisher information matrix appears as a cleaner way to prioritize among observable variables, because it does not require any prior data information.

4.2 Estimated Model

Now we repeat the above experiments with actual 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 Federal Reserve Economic Data managed by the Federal Reserve Bank of St. Louis.[7] We use the annual series of the real personal consumption expenditures (mnemonic: 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 and 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 growth model and estimate it on an annual frequency. Because of the non-stationary nature of the actual data, we modify the technological process to include a stochastic trend. 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.[8] The other parameters are kept fixed at their parameterized values mentioned above. We will not report their estimates and instead we again focus on the estimates of the capital stock K_t and labor augmenting technology A_t.

Figure 4 summarizes the basic results. In the top two panels, we compare the confidence intervals for the *smoothed* estimates of K_t and A_t. The panels show the relative efficiency of the capital stock estimate (top left) and labor

[7] Web page `http://research.stlouisfed.org/fred2/`.

[8] The parameters are estimated to allow for a better data match. It does not pose a fundamental change to the setup of the experiment.

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 then computed as $100(std(X_{t,model1})/std(X_{t,model2}) - 1)$. Positive values mean that the model two – model with more information, outperforms the model one – model with less information.

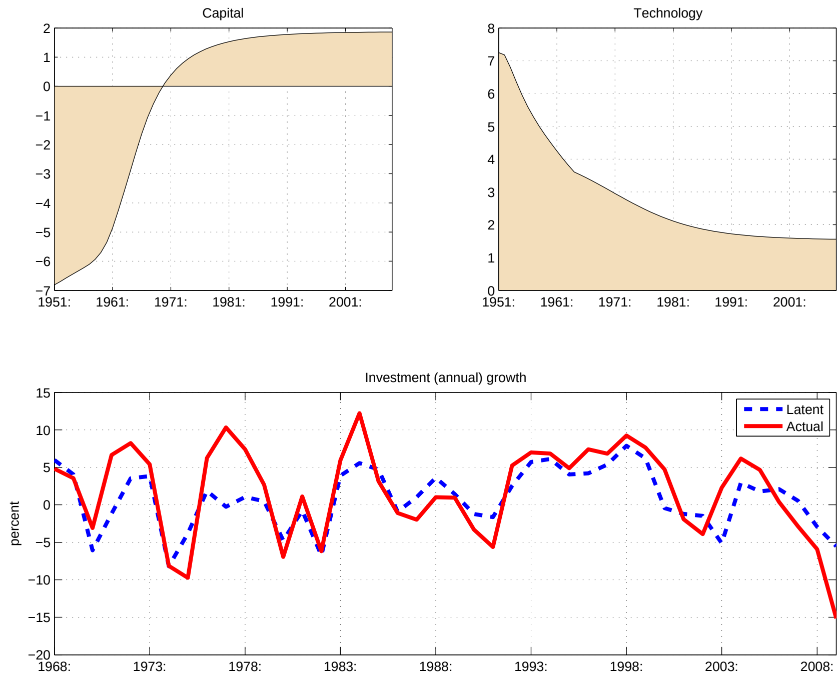
The model with more information (Figure 4, model two) helps to improve the estimate of the labor augmenting technology (right panel). Early in the sample the relative efficiency of model two is 7 times higher than of the model with less information. The relative advantage of model two gradually diminishes as model one adapts and as its efficiency improves over time. In 2010, the relative efficiency of model two is already only 2 times better than of model one.

We recall that the estimates are smoothed estimates; that is, the early estimates of A_t are based on the complete information set available at time T. That is why the model with more information performs better over the whole sample even if the extra information in the form of investment growth comes in after 1967.

In the instance of capital stock (Figure 4, left panel), model two starts to outperform model one shortly before the year 1971, when the new information from hours worked and investment begins to feed in. Early in the sample, the model with less information performs almost 7 times better than the model two. This is the price of the improved estimate for the technological process in this time period. After 1971, the model with more information again clearly outperforms model one, delivering estimates of the capital stock that are twice as efficient as in model one.

One may wonder why the observations on investment do not provide even higher gains in efficiency. The model's good (in-sample) predictive power for investment provides an explanation. The graph in the bottom panel of Figure 4 compares the model implied investment (when treated as latent in model one) to the actually realized (observed) data. Clearly, the data on private consumption expenditures and hours worked by themselves contain enough information about investment and thus the capital stock and also technology.

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 when that information is extended with (ii) the investment growth (model 2). The shaded areas are computed as $100(std(X_{t,model1})/std(X_{t,model2}) - 1)$.

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 in what this condition implied 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 illustrated that this leads to a significantly improved estimate efficiency.

The NCC can provide an alternative structural perspective for 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 rather with the flow of information in it.

In future work we would like to look at the possible avenues for formal testing of the NCC, which can be used for an empirical assessment of endogenous decision rules. DSGE models consist of optimal decision (control) rules, so each equation can be subject to testing. Another possible avenue for research is to relax the assumption of information pooling, and look at the case of an agent with significant market power and private information. If the NCC should hold, the remaining market players can try to infer the private information encoded in the decisions of the dominant player and adapt to it.

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*

Schorfheide F., K. Sill, and M. Kryshko (2010), "DSGE Model-Based Forecasting of Non-Modeled Variables," *International Journal of Forecasting*, vol 26(2), pages 348-373.

Guerron-Quintana P. A. (2010), "What You Match Does Matter: The Effects of Data on DSGE Estimation," *Journal of Applied Econometrics*, vol. 25(5), pages 774-804.