Short–Run Restrictions: An Identification Device?

Fabrice Collard
University of Toulouse (CNRS–GREMAQ and IDEI)

Patrick Fève*
University of Toulouse (GREMAQ and IDEI)
and Banque de France (Research Division)

Julien Matheron
Banque de France (Research Division)

March, 2006

Abstract

This paper examines the extent to which short–run restrictions can be used as a way to identify the deep parameters of a DSGE model. We propose a simple statistical framework designed to quantitatively assess this issue. We then consider a fully–fledged RBC model with habit formation and investment adjustment costs that we take to the data using Minimum Distance Estimation (MDE). We show that as long as the model possesses weak propagation mechanisms, MDE effectively treats structural parameters as deep parameters. On the contrary as soon as propagation mechanisms are rich enough, the ability of MDE to identify deep parameters is questioned because timing restrictions propagate over time.

Key Words: Minimum Distance Estimation, DSGE models, SVARs, Identification.

JEL Class.: E32, C52.

*Address: GREMAQ–University of Toulouse, manufacture des Tabacs, bât. F, 21 allée de Brienne, 31000 Toulouse. email: patrick.feve@univ-tlse1.fr. We would like to thank F. Canova, M. Dupaigne, A. Guay, L.P. Hansen, F. Portier, H. Uhlig and the participants to the first AMEN meeting (Barcelona, November 2005)for helpful discussions. The traditional disclaimer applies. The views expressed herein are those of the authors and not necessarily those of the Banque de France.

Introduction

The econometrics of Dynamic Stochastic General Equilibrium (DSGE) models has witnessed substantial advances over the recent years. It is nowadays more and more common to bring DSGE models to the data using a variety of formal statistical techniques, including Maximum Likelihood estimation (Altug [1989], Ireland [2004]), Generalized Method of Moments (Christiano and Eichenbaum [1992], Burnside, Eichenbaum and Rebelo [1993]), Bayesian techniques (Schorfheide [2000], Smets and Wouters [2003]), and Minimum Distance Estimation (Rotemberg and Woodford [1997], Christiano, Eichenbaum and Evans [2005a]). The present paper is concerned with the principles underlying the latter econometric technique.

The Minimum Distance Estimation (MDE) technique consists in estimating the structural parameters of DSGE models so as to minimize a weighted distance between theoretical impulse responses of key macroeconomic variables to structural shocks and those derived from a Structural Vector Autoregression (SVAR). The attractive feature of this method is that it allows researchers to bring structural and empirical approaches into closer conformity. Moreover, such a limited information approach does not impose fully specifying the whole stochastic structure of the DSGE model as attention is focused only on those shocks that are relevant for the question under study.

This method requires that an auxiliary SVAR model be estimated prior to estimating the DSGE parameters. In doing so, a researcher has access to at least two broad types of identifying restrictions.[1] Blanchard and Quah [1989] and Galí [1999] have proposed to identify shocks based on long-run restrictions. For example, Galí [1999] identifies technology shocks as the only shocks that have an effect on the long–run level of labor productivity. The attractive feature of this approach is that this type of restrictions holds in a broad class of competing DSGE models. However, Erceg, Guerrieri and Gust [2005], and Chari, Kehoe and McGrattan [2005] have recently questioned the ability of long–run restrictions in SVAR to properly recover the shocks. An alternative approach consists in imposing short–run restrictions through which some variables are forbidden to react to some shocks on impact (Sims [1980], Christiano, Eichenbaum and Evans [1999], Christiano et al. [2005a]). Christiano, Eichenbaum and Vigfusson [2005b] show that this latter technique performs remarkably well compared to long–run restrictions. In particular, such restrictions are able to correctly and precisely pin down the shocks. This suggests that SVAR with short-run restrictions can be a useful assessment device when constructing and evaluating a DSGE model, even though a broad class of DSGE models do not a priori meet

[1] Another class of identifying assumptions relies on sign restrictions, which are generally robust. However, using such restrictions requires that the modeler has a prior knowledge of the expected sign of the responses under study. This prior knowledge generally stems from a DSGE model. Notice that when some deep parameters, unknown to the econometrician, vary, the sign of the response of interest might change as well. Our model will exemplify this case.

these identifying assumptions (see Canova and Pina [2005]).

It is thus important to stress that applying the MDE approach requires that DSGE and SVAR share the same restrictions. With short-run restrictions, this implies that some form of recursiveness be imposed in the DSGE model to make it compatible with the SVAR. This requires that information sets in the model be manipulated, i.e. some decisions are taken prior to observing the current period shocks. To some, imposing this type of constraints in a DSGE model might seem restrictive.

Yet, restricting the information sets has proved useful in many empirical studies. First, this has proved a fruitful research strategy, as exemplified by factor hoarding models (Burnside et al. [1993], Burnside and Eichenbaum [1996]), limited participation models (Christiano [1991], Christiano, Eichenbaum and Evans [1997]), or preset prices models (Obstfeld and Rogoff [1995], Galí [1999]). Second, this method has sparked important developments in applied monetary economics, e.g. Rotemberg and Woodford [1997], Christiano et al. [2005a], Boivin and Giannoni [2005], Giannoni and Woodford [2004]. This is exemplified in Woodford [2003] who discusses how monetary DSGE models can be brought into closer conformity with monetary SVARs by suitably restricting the information set used to base agents decisions. Short–run restrictions are now used as an identification and estimation device.

In this paper, we investigate the ability of such restrictions to properly identify the deep parameters of DSGE models. More precisely we ask the question of whether the MDE method treats structural parameters as deep parameters — *i.e.* parameters that remain invariant to the short-run identification scheme used to identify shocks. For example, Christiano et al. [2005a] show that manipulating the information set has negligible effects on the dynamic properties of their model, in response to a monetary policy shock identified with short-run restrictions. Similarly, Christiano et al. [2005b] show that in a standard Real Business Cycle (RBC) model, the response of hours to a permanent technology shock is not dependent upon the timing of decisions regarding labor supply – except on impact. We argue that this invariance property stems from the lack of strong propagation mechanisms with regards to the shock under study. In contrast, in models with richer propagation properties, the timing of decisions may deeply affect both the sign, magnitude, and persistence of responses to shocks. This suggests that MDE with short–run restrictions might not identify truly deep parameters.

We first develop a simple statistical framework designed to quantitatively assess the role of timing restrictions. The general idea of our procedure is to ask whether it is possible to find a value of deep parameters in an unrestricted version of the DSGE model under consideration that allows us to match the responses obtained in its restricted version. If so, both versions are observationally equivalent in terms of the auxiliary model (SVAR). It is then possible to test whether the estimated structural parameters in both versions of the DSGE model are equal. When the null of equivalence is not rejected, MDE is found to truly estimate deep parameters, *i.e.* parameters are invariant to the timing of decisions.

We then consider a fully–fledged RBC model with habit formation and investment adjustment costs in the line of Christiano et al. [2005a]. A version of this model with predetermined hours is estimated on US data via MDE using the impulse response functions of output and hours to a technology shock as implied by a SVAR with short–run restrictions.[2] The model is found to fit satisfactorily the data and estimated structural parameters point to a significant degree of real frictions (habit, investment adjustment costs).

Finally, we apply our proposed statistical procedure and assess the role of timing restrictions. We first show that the estimated DSGE models with restricted and unrestricted hours differ sharply. When hours cannot respond to technology shocks on impact, they increase smoothly in the subsequent periods, whereas they dramatically drop when they can freely react to the shock. As a complementary exercise, we investigate a stripped down version of the model where propagation mechanisms are drastically weakened (habit and adjustment costs parameters set to zero). In this case, we find that restricting the information sets has no substantial consequences on the dynamics of hours and other aggregate variables. Note however that this version of the model is rejected by the data. Second, we investigate how information lags propagate over the entire economy. We find that output, productivity, consumption and investment are all affected by the timing of hours decisions. Our testing approach formally confirms the quantitative importance of restricting hours. In particular, we show that the structural parameters estimated so that the unconstrained model matches its constrained counterpart statistically differ from the original estimates. This means that these cannot be treated as deep parameters. Finally, we show that the model with restricted hours faces important identification problems, except when investment is taken as the informative variable.

The paper is organized as follows. In section 1, we start by expounding the basic principles of DSGE models evaluation from SVARs with short-run restrictions. We lay out a number of statistical tools designed to assess the role of timing decisions. Section 2 considers a fully-fledged DSGE model with habit formation and investment adjustment costs, which we bring to the data. In section 3, we use the estimated model and the method outlined in section 1 to illustrate how information restrictions are not quantitatively innocuous. The last section briefly concludes.

[2] Note that in the present case, short–run restrictions deliver responses very similar to what would stem from a SVAR with long–run restrictions. This suggests at first glance that short–run restrictions do not a priori distort the IRFs as estimated by the SVARs.

1 Short-Run Restrictions as an Identification Device

This section presents a brief summary of the Minimum Distance Estimation method and our evaluation procedure.

1.1 Minimum Distance Estimation with Short-Run Restrictions

Let us consider a sequence of data $\{x_t\}_{t=1}^T$, where x_t is a vector of dimension n_x. We assume that x_t can be represented by a canonical VAR of the form

$$A(L)x_t = \varepsilon_t,$$

where $A(L) = (I - A_1 L - \ldots - A_\ell L^\ell)$ and ℓ denotes the number of lags. ε_t are the canonical innovations with zero mean and covariance matrix $\mathscr{E}\{\varepsilon_t \varepsilon_t'\} = \Sigma$.

Economists are mostly interested in uncovering the structural shocks that hit the economy. These are intimately related to the canonical innovations, as the latter can be viewed as linear combinations of structural shocks η_t,

$$\varepsilon_t = S\eta_t,$$

which hit the economy in each and every period. S is a non singular matrix that specifies the mapping between canonical innovations and structural shocks. The data only provide information on the canonical innovations, and some additional assumptions have to be placed to identify the structural shocks. We therefore impose an orthogonality assumption on the structural shocks and a normalization condition ($\mathscr{E}\{\eta_t \eta_t'\} = I_{n_x}$). Doing so, we set out $n_x(n_x+1)/2$ constraints out of the n_x^2 needed to completely identify S. Imposing the remaining $n_x(n_x-1)/2$ identifying constraints usually requires restrictions borrowed from economic theory.

Many types of of identifying restrictions can be imposed on a system. Blanchard and Quah [1989] and Galí [1999] have proposed to identify shocks based on long–run restrictions. Uhlig [2005] and Canova and de Nicolò [2002] propose to identify shocks based on sign restrictions. In this paper, we follow Sims [1980] and consider short-run restrictions. In such a setting, S can then be recovered by solving $SS' = \Sigma$. A structural $MA(\infty)$ representation of the dynamics of x_t is then given by

$$x_t = C(L)\eta_t$$

where $C(L) = \sum_{i=0}^{\infty} C_i L^i = A(L)^{-1} S$.

Imposing short-run identifying assumptions amounts to restrict the impact response of certain variables to structural shocks, i.e. set to zero some elements of $C(0)$. Notice that since $A(0) = I_{n_x}$, it must be the case that $C(0) = S$. The selection of the variables that are allowed to contemporaneously react to a shock is undertaken by *(i)* imposing a particular ordering in x_t and *(ii)* setting matrix S equal to the Cholesky decomposition of Σ (see Sims [1980]). As an example, let us consider a VAR modeling the joint dynamics of hours and productivity. One may then recover a technology shock imposing that this shock has no impact effect on hours, while productivity may react. This assumption can be imposed by assuming *(i)* that hours come first in the vector x_t, and *(ii)* that $C(0)$ has the following structure

$$C(0) = \begin{pmatrix} \times & 0 \\ \times & \times \end{pmatrix}$$

Let us define $C_h^{(i,j)}$ the (i,j) element of matrix C_h, the h–th coefficient of the matrix polynomial $C(L)$. Thus, the response of variable $x_{i,t}$ to shock j at horizon h is given by

$$I^{ij}(h) \equiv \frac{\partial x_{i,t+h}}{\partial \eta_{j,t}} = C_h^{(i,j)}.$$

In the MDE approach, these responses constitute the objects to be matched by the DSGE model. Let $\widehat{I}_T(\mathcal{H})$ denote the vector collecting the responses of interest from the SVAR for the horizons $\mathcal{H} = \{1, \ldots, h\}$, obtained from the data $\{x_t\}_{t=1}^T$. In the case of short-run restrictions and for practical reasons, it is important to exclude the impact response of constrained variables, since the latter is a degenerate random variable.

These empirical impulse responses are then used to identify the structural parameters of a DSGE model. A generic linear(ized) DSGE model admits the following representation

$$\mathscr{E}_t^{\star}\left\{\sum_{i=-\tau_{\rm b}}^{\tau_{\rm f}} H_i(\theta_1)\, y_{t+i} + \sum_{i=-r_{\rm b}}^{r_{\rm f}} R_i(\theta_1)\, s_{t+i}\right\} = 0$$

where the τ's and the r's are integer, the $H_i(\theta_1)$'s and the $R_i(\theta_1)$'s are matrices whose elements are possibly complicated functions of the deep parameters θ_1. y_t is a vector of n_y endogenous variables, and s_t is a vector of n_s exogenous variables, including current and past values of the structural shocks. Notice that for the MDE to be well defined, the variables in x_t must be a subset of the variables about which the DSGE model has something to say, i.e. x_t is composed of elements included in y_t. Finally, $\mathscr{E}_t^{\star}\{\cdot\}$ is a conditional expectation operator satisfying

$$\mathscr{E}_t^{\star}(F(y_i, s_j)\} = \begin{bmatrix} \mathscr{E}\{F_1(y_i, s_j)|\mathscr{I}_{1,t}\} \\ \mathscr{E}\{F_2(y_i, s_j)|\mathscr{I}_{2,t}\} \\ \vdots \\ \mathscr{E}\{F_{n_y}(y_i, s_j)|\mathscr{I}_{n,t}\} \end{bmatrix} \text{ with } i = -\tau_b \ldots, \tau_f, j = -r_b \ldots r_f$$

where F denotes the set of equations of the model. This makes it explicit that each endogenous variable is possibly determined on the basis of a specific information set, and can therefore be prevented from reacting to a particular shock in the current period, as long as the latter is excluded from $\mathscr{I}_{i,t}$. It is important to stress that the model should share the same information structure as that implied by the recursive identification scheme imposed in the SVAR. In other words, the information sets $\mathscr{I}_{i,t}$, $i = 1, ..., n_y$, in the DSGE should incorporate the exact same short–run exclusion restrictions as in the VAR. As for illustrative purposes, let us go back to our simple example. Since the technology shock is identified, in the VAR model, assuming that hours do not react to this shock, we would assume that agents take their labor supply decisions prior to observing the shocks in the DSGE model.

Note that these information restrictions have critical implications for the interpretations of the structural shocks and the definition of the relevant innovation. Indeed, without any information constraint, the shock taking place in the current period constitutes the innovation of any of the variables of the economy. In contrast, when some variables are decided prior to observing the structural shocks, the relevant innovations for such variables are those pertaining to the latest period. As an example, let us consider again the case where hours in period t are decided prior to observing the technology shock of period t. The relevant innovation to be considered for hours is the technology shock that occurred in period $t-1$.

The vector of exogenous variables s_t is assumed to evolve according to the law of motion

$$s_t = P\left(\theta_2\right) s_{t-1} + Q\left(\theta_2\right) \varepsilon_t,$$

where θ_2 is a vector of parameters governing the dynamics of the forcing variables. A rational expectations solution to the model then admits the following representation

$$y_t = D_1\left(\theta_1\right) y_{t-1} + \cdots + D_{\tau_b}\left(\theta_1\right) y_{t-\tau_b} + M\left(\theta_1, \theta_2\right) s_t.$$

Using this solution, it is easy to uncover the mapping $I\left(\theta, \mathcal{H}\right)$ from the structural parameters to the model counterpart of $\widehat{I}_T\left(\mathcal{H}\right)$, where $\theta \equiv (\theta_1, \theta_2)$. Then, the MDE estimator of θ is defined as follows

$$\widehat{\theta}_T = \underset{\theta \in \Theta}{\operatorname{argmin}} \; (I(\theta, \mathcal{H}) - \widehat{I}_T\left(\mathcal{H}\right))' W_T (I(\theta, \mathcal{H}) - \widehat{I}_T\left(\mathcal{H}\right)), \tag{1}$$

where Θ is the set of admissible values for θ and W_T is a positive semi-definite weighting matrix. W_T denotes the inverse of a consistent estimate of the covariance matrix of $\widehat{I}_T(\mathcal{H})$. It is common practice in the literature to use instead a diagonal weighting matrix that involves the inverse of each impulse response's variance on the main diagonal (see Christiano et al. [2005a]). With this choice, the structural parameters θ are selected so that $I(\theta, \mathcal{H})$ lies as much as possible inside the confidence interval of $\widehat{I}_T\left(\mathcal{H}\right)$. Efficient MDE would suggest to use the complete variance-covariance matrix of impulse response functions as a weighting matrix. In most empirical applications, this may be impossible. Indeed, the nh impulses ($n \leqslant n_x$ and h is the horizon of each response) that we want to match depend on the vector of VAR parameters, whose dimension is $k = n_x^2 \ell + n_x(n_x+1)/2$. In practice, however, it is not uncommon that

$nh > k$. For example, Altig, Christiano, Eichenbaum and Linde [2005] study a SVAR with four lags and 10 variables, implying $k = 455$. They seek to match the impulse response functions of these 10 variables to three shocks, for $h = 20$, implying $nh = 592 > k$, once short-run exclusion restrictions have been properly taken into account.

The structural parameters θ are pinned down so as to minimize the discrepancy between SVAR–based impulse response functions and their DSGE counterparts. The value of the objective function at convergence can then be used to test the ability of the model to match the impulse response functions implied by the SVAR. This is undertaken using an overidentification test à la Hansen [1982] (Hereafter J–stat). The statistics is distributed as a chi–square with $nh - \dim(\theta)$ degrees of freedom. $\widehat{\theta}_T$ is normally distributed with covariance matrix

$$V(\widehat{\theta}_T) = \left(\left.\frac{\partial I(\theta,\mathcal{H})'}{\partial\theta}\right|_{\theta=\widehat{\theta}_T} W_T \left.\frac{\partial I(\theta,\mathcal{H})}{\partial\theta'}\right|_{\theta=\widehat{\theta}_T} \right)^{-1}.$$

In the sequel, all the quantitative assessments of the role of timing assumptions rely on $\hat{\theta}_T$ and $V(\widehat{\theta}_T)$.

1.2 Assessing the Role of Timing Restrictions

The deep parameters as estimated by the method described in the previous section are obtained imposing a set of short–run restrictions on the behavior of variables both in the data and in the model. This section is concerned with the exact quantitative role played by these restrictions. This assessment is therefore in line with the robustness check conducted by Christiano et al. [2005a]. They show that, as far as their model is concerned, the timing of decisions does not have a substantial impact on the dynamic properties of their model economy. Their assessment essentially stems from comparing impulse response functions where informations lags are kept and dropped. In this section, we propose to extend the analysis and provide additional tools to gauge the role of timing restrictions.

As a first check, we propose to follow Christiano et al. [2005a] and consider two alternative models corresponding to two short–run identification schemes. In the first model, labeled $\mathcal{M}_1(\theta)$, we assume that all variables can instantaneously and freely react to any unexpected shock on a given exogenous variable. The second model, labeled $\mathcal{M}_2(\theta)$, shares the same information structure as the SVAR. It is therefore constrained. Accordingly, it is the model used to estimate the deep parameters, $\widehat{\theta}_T$. Then, for a given $\widehat{\theta}_T$, it is possible to assess the role of the timing restrictions by comparing the impulse responses generated by model $\mathcal{M}_1(\widehat{\theta}_T)$ and $\mathcal{M}_2(\widehat{\theta}_T)$. This comparison should obviously be conducted for the set of variables used to estimate the deep parameters as well as the other variables the model has something to say about. This reveals the way the dynamic properties of the models do propagate information lags through the economy.

The strict comparison of the impulse response functions (sign, amplitude, persistence) may be supplemented by the computation of the ratio of the two impulse responses for the relevant innovation. Should this ratio be constant over time, the propagation of this innovation would be the same in the two models. An alternative way of assessing the problem is to have a look at the autocorrelation function as computed by restricting the space of shocks to the shock of interest.

When the impulse responses are the same in both models, short–run restrictions have no impact on the dynamic properties of the model. A direct consequence of this result is that the estimation method would lead to the same estimates for θ no matter the timing imposed on decisions. In such a setting the estimation method deals with the structural parameters as deep parameters — *i.e.* invariant to the timing. When, on the contrary, the impulse responses differ, we have $\mathcal{M}_1(\widehat{\theta}_T) \neq \mathcal{M}_2(\widehat{\theta}_T)$, meaning that short–run restrictions have a long–lasting effect on the dynamics of the model. On the one hand, these are good news as this means that information lags carry a lot of information usable by the estimation method.[3] This may actually be used as a model selection device. On the other hand, this has a direct consequence on the estimation of the parameters that ought to differ depending on the timing of decisions. Therefore, this raises the question of the way this econometric approach deals with structural parameters: Are they deep parameters? Indeed, when $\mathcal{M}_1(\widehat{\theta}_T) \neq \mathcal{M}_2(\widehat{\theta}_T)$, it is legitimate to ask whether there exists a set of parameters $\widetilde{\theta}_T$ such that $\mathcal{M}_1(\widetilde{\theta}_T) = \mathcal{M}_2(\widehat{\theta}_T)$ — i.e. model $\mathcal{M}_1$, when parameterized with $\widetilde{\theta}_T$ is able to mimic the model with timing restrictions. This suggests a second check of the estimation method.

This second way of evaluating the method amounts to conducting a test in the line of encompassing principles. The idea is first to find a vector of parameters $\widetilde{\theta}_T$ that minimizes the distance between the impulse response function of unconstrained variables as computed in model $\mathcal{M}_1(\theta)$ — the model without timing restrictions — and the impulse response function from model $\mathcal{M}_2(\widehat{\theta}_T)$ — i.e. for the parameter set such that the impulse response of $\mathcal{M}_2$ match the data

$$\widetilde{\theta}_T = \underset{\theta\in\Theta}{\text{argmin}} \left(I_1(\theta,\mathcal{H}) - I_2(\widehat{\theta}_T,\mathcal{H})\right) \Omega_2(\widehat{\theta}_T) \left(I_1(\theta,\mathcal{H}) - I_2(\widehat{\theta}_T,\mathcal{H})\right)' \quad (2)$$

where $I_j(\theta,\mathcal{H})$ collects all the impulse response functions to the shock we want to consider in the test, for the set of horizons $\mathcal{H} = \{1,\dots,h\}$, in model j. The weighting matrix $\Omega_2(\widehat{\theta}_T)$ is the covariance matrix of the impulse response function $I_2(\widehat{\theta}_T,\mathcal{H})$ and is deduced from the covariance matrix of $\widehat{\theta}_T$ as

$$\Omega_2(\widehat{\theta}_T) = \left(\left.\frac{\partial I_2(\theta,\mathcal{H})'}{\partial\theta}\right|_{\theta=\widehat{\theta}_T} V(\widehat{\theta}_T) \left.\frac{\partial I_2(\theta,\mathcal{H})}{\partial\theta'}\right|_{\theta=\widehat{\theta}_T} \right)^{-1}. \quad (3)$$

[3]This is actually no surprise as a whole strand of the literature has used these delays to reconcile models with the data (see Burnside et al. [1993], Burnside and Eichenbaum [1996], Christiano et al. [1997] ...)

The matrix $\Omega_2(\widehat{\theta}_T)$ may not be defined when $\dim(\widehat{\theta}_T) < \dim(I_2(\theta,\mathcal{H}))$. As in the case of the weighting matrix W_T at the estimation stage (see equation (1)) in practice we use instead the inverse of the diagonal of $(\partial I_2(\theta,\mathcal{H})'/\partial\theta)V(\widehat{\theta}_T)(\partial I_2(\theta,\mathcal{H})/\partial\theta')$. Once again, with this choice, the structural parameters θ of model $\mathcal{M}_1$ are selected so that $I_1(\theta,\mathcal{H})$ lies as much as possible inside the confidence interval of $I_2(\hat{\theta}_T,\mathcal{H})$.

An attractive feature of this procedure is that it allows to test immediately for the equivalence between the two models. Indeed, one can conduct an over–identification test *à la* Hansen [1982], which is distributed as a chi–square with $nh - \dim(\theta)$ degrees of freedom, where n is the number of variables for which we want to match the impulse response functions. When the null hypothesis is rejected, there does not exist a value for θ for which the model without timing restrictions is observationally equivalent to the constrained model. Information sets matter. However, when the null hypothesis is rejected, one still cannot conclude that information restrictions have no consequences on the dynamics of the model. More precisely, one must test whether $\widetilde{\theta}_T = \widehat{\theta}_T$ before concluding that information sets do not matter. This can be formally tested using the following Wald statistics

$$\mathcal{W}_1(\widetilde{\theta}_T,\widehat{\theta}_T) = (\widetilde{\theta}_T - \widehat{\theta}_T)'V(\widehat{\theta}_T)^{-1}(\widetilde{\theta}_T - \widehat{\theta}_T) \tag{4}$$

Under the null of equality, the statistics is distributed as a chi–square with $\dim(\theta)$ degrees of freedom. When the null is not rejected, the two models are observationally equivalent. Information restrictions do not matter for the estimation of the deep parameters. On the contrary, when the null is rejected, $\widetilde{\theta}_T \neq \widehat{\theta}_T$, the structural parameters, when plunged into model $\mathcal{M}_1$, have to be distorted in order to match the dynamics implied by the model with information lags. The information structure then matters. But this questions the ability of the econometric technique, which uses short–run restrictions, to treat structural parameters as deep parameters — i.e. parameters invariant to the information structure.

One can go one step further by rebuilding some counterfactual dynamics for the model with information lags. Indeed, the last value of the set of parameters, $\widetilde{\theta}_T$, (when statistically different from $\widehat{\theta}_T$) can be used to feed model $\mathcal{M}_2$. This then allows us to test whether $\mathcal{M}_2$ is affected by this change in θ. This can be formally tested using the Wald test

$$\mathcal{W}_2(\widetilde{\theta}_T,\widehat{\theta}_T) = \left(I_2(\widetilde{\theta}_T,\mathcal{H}) - I_2(\widehat{\theta}_T,\mathcal{H})\right)\Omega_2(\widehat{\theta}_T)\left(I_2(\widetilde{\theta}_T,\mathcal{H}) - I_2(\widehat{\theta}_T,\mathcal{H})\right)' \tag{5}$$

where $\Omega_2(\widehat{\theta}_T)$ is defined in equation (3). Under the null, $\mathcal{W}_2(\widetilde{\theta}_T,\widehat{\theta}_T)$ is distributed as a chi–square with nh degrees of freedom. When the null of equality cannot be rejected, a change in the parameter does not alter the dynamics properties of model $\mathcal{M}_2$. This reveals a severe identification problem of the structural parameters in the model, as highlighted by Canova and Sala [2005].

2 A Fully—Fledged DSGE Model

This section develops a fully-fledged business cycle model including a variety of real frictions. These features reinforce the propagation mechanisms and therefore allow for a formal test of the model.

2.1 The Model

We consider an extended version of the RBC model in which we allow for habit formation and investment adjustment costs. Both mechanisms have proven useful in accounting for the dynamics of macroeconomic variables in particular in terms of their persistence properties (see e.g. Boldrin, Christiano and Fisher [2001] and Christiano et al. [2005a]).

We assume that intertemporal consumption choices are not time separable and that the flows of consumption services are a linear function of current and lagged consumption decisions. Labor is assumed to be indivisible as in Hansen [1985]. More precisely, the intertemporal expected utility function of the representative household is given by

$$\mathscr{E}\left[\sum_{s=0}^{\infty}\beta^s\left\{\log\left(C_{t+s}-bC_{t+s-1}\right)-\chi N_{t+s}\right\}\middle|\mathscr{I}_t\right], \tag{6}$$

where $\beta \in (0,1)$, $b \in [0,1)$ and $\chi > 0$. In what follows, we consider two versions of this model, differentiated according to the timing decision of hours. $\mathscr{E}(\cdot|\mathscr{I}_t)$ denotes the expectation operator conditional on the information set $\mathscr{I}_t$.

The representative firm produces the homogeneous final good Y_t by means of capital, K_t, and labor, N_t, using a constant returns–to–scale technology represented by the following Cobb–Douglas production function

$$Y_t = K_t^{\alpha}\left(Z_t N_t\right)^{1-\alpha},$$

where $\alpha \in (0,1)$. Z_t is a shock to total factor productivity and is assumed to follow a random walk process with drift of the form

$$\log(Z_t) = \gamma_z + \log(Z_{t-1}) + \sigma_z \varepsilon_{z,t},$$

where $\sigma_z > 0$ and $\varepsilon_{z,t}$ is *iid* with zero mean and unit variance. The constant γ_z is the drift term and accounts for the deterministic component of the growth process. The homogenous good can be either used for consumption and investment purposes. Capital accumulation is governed by the law of motion

$$K_{t+1} = (1-\delta)\,K_t + \mathcal{F}\left(I_t, I_{t-1}\right)$$

where $\delta \in (0,1)$ is the constant depreciation rate. The function $\mathcal{F}(\cdot,\cdot)$ accounts for the presence of adjustments costs in the capital accumulation. $\mathcal{F}(\cdot,\cdot)$ indeed rewrites as

$$\mathcal{F}(I_t, I_{t-1}) = \left[1 - \mathcal{S}\left(\frac{I_t}{I_{t-1}}\right)\right] I_t$$

where $\mathcal{S}(\cdot)$ reflects the presence of adjustment costs. We assume that $\mathcal{S}(\cdot)$ satisfies *(i)* $\mathcal{S}(\gamma_z) = \mathcal{S}'(\gamma_z) = 0$ and *(ii)* $\xi = \mathcal{S}''(\gamma_z)\gamma_z^2 > 0$. It follows that the steady state of the model does not depend on the parameter ξ while its dynamic properties do. Notice that following Christiano et al. [2005a], Christiano and Fisher [2003] and Eichenbaum and Fisher [2005], we adopt the dynamic investment adjustment cost specification. In this environment, it is the growth rate of investment which is penalized when varied in the neighborhood of its steady state value. In contrast, the standard specification penalizes the investment-to-capital ratio. The dynamic specification for adjustment costs is a significant source of internal propagation mechanisms as it allows for a hump-shaped response of investment to various shocks.

Finally the market clearing condition on the good market writes

$$Y_t = C_t + I_t.$$

We consider two versions of this model. In the first version, denoted $\mathcal{M}_1$, we follow the standard approach and assume that all decisions are taken after the realization of the technology shock at t ($\mathscr{I}_t = \{Z_{t-i}; i = 0, \dots, \infty\}$). In the second version, denoted $\mathcal{M}_2$, we follow Christiano et al. [2005b] and assume that labor supply decisions are taken prior to observing the technology shock ($\mathscr{I}_t = \{Z_{t-i}; i = 1, \dots, \infty\}$). Consumption and investment are however chosen after the shock is observed.

The model is then deflated for the stochastic trend component[4] and log–linearized around the deterministic steady state. The approximated dynamic system is then solved using standard techniques.

2.2 Benchmark Calibration

In order to provide quantitatively meaningful results, we need to assign values to the deep parameters. We split the set of structural parameters in two subsets. The first set consists of parameters held fixed across all experiments. This set consists of $\{\alpha, \beta, \delta, \gamma_z\}$ and the assigned values are reported in Table 1. The parameter α is set such that the labor share is 0.64. The depreciation rate, δ, is set such that capital depreciates at the annual rate of 10%. The growth factor γ_z is set to 1.005, so as to mimic the average growth rate of US output over the period 1948:I-2002:IV. Finally, the discount factor, β, is set to 0.99.

[4]Output, consumption and investment are divided by Z_t, and the capital stock is divided by Z_{t-1}

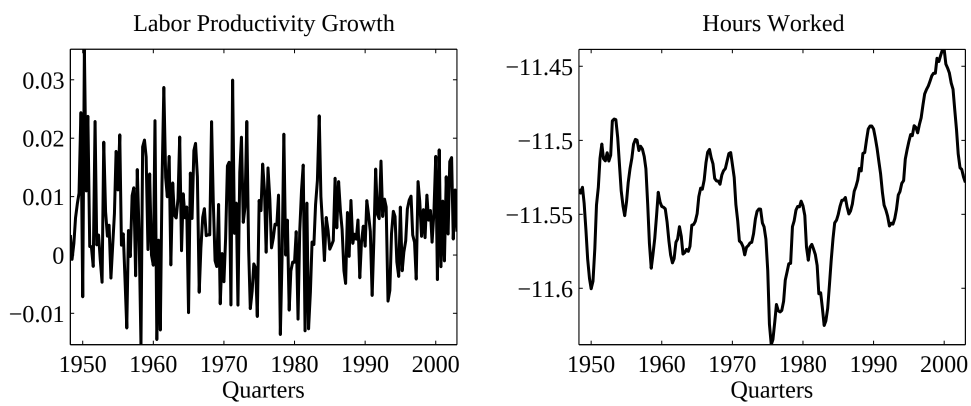
Table 1: Baseline parametrization

γ_z	α	δ	β
1.005	0.360	0.025	0.990

The second subset, θ, consists of the habit persistence parameter b, the adjustment costs parameter, ξ, and the standard deviation of the technology shock, σ_z. We investigate two cases. In the first one, we estimate b, ξ, and the volatility of the shock so as to reproduce the response of hours worked and output to a technology shock in the US economy as obtained from a SVAR. In the second case, both b and ξ are set to zero implying that the model corresponds to the canonical frictionless RBC model and is similar to the case considered by Chari et al. [2005] and Christiano et al. [2005b].

Structural parameters are estimated using the MDE approach. We therefore estimate a SVAR for the US economy. Following Galí and Rabanal [2004], we use quarterly U.S. data on the growth rate of labor productivity in the nonfarm business–sector, and hours of all persons[5] in this sector for the period 1948:I–2002:IV. The data are reported in Figure 1. Hours are taken

Figure 1: Data



in level, as in Christiano, Eichenbaum and Vigfusson [2004]. In practice, the VAR is estimated with four lags.

Usually, technology shocks are identified in SVARs as the only shocks that exert a long-run effect on the level of labor productivity, e.g. Galí [1999]. Using such an identifying restriction, we obtain results reported in Figure 2. The figure graphs the impulse response function of output and hours worked after a positive technology shock as identified by the long–run (plain line) identification assumption. The shaded area corresponds to the 95% confidence interval of the responses obtained relying on numerical integration. As can be seen on the graph, output

[5] Hours are expressed in per–capita terms using the civilian noninstitutional population aged 16 and over.

Figure 2: Impulse Response to a Technology Shock (Actual Data)

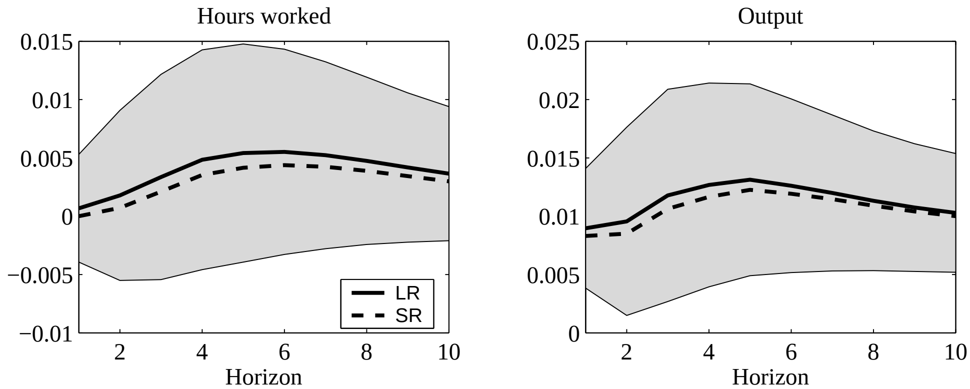
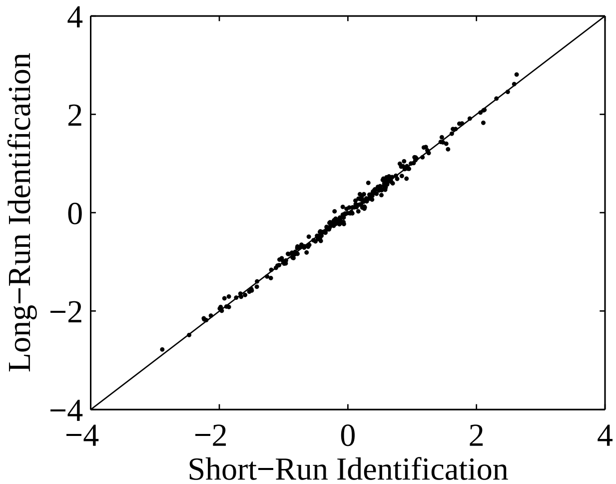


Figure 3: Plot of Structural Shocks



permanently and positively increases after a technology improvement, while hours increase in the short–run and only gradually. Interestingly, the impact response of hours is almost zero. This suggests the alternative identification restriction of technology shocks that consists in assuming that these shocks have no contemporaneous effect on hours worked (dashed line in Figure 2). This actually corresponds to the information structure of model $\mathcal{M}_2$. Interestingly, the dynamics of output and hours obtained using a short–run identification scheme are very similar to those obtained using long–run restrictions.[6] This is illustrated in Figure 2. The figure shows that the short–run identification statistically delivers the same technology component as the long–run identification. This is confirmed by Figure 3 which reports the plot of the technology shock as identified relying on a short–run identification against the technology shock as identified relying

[6]Extending the VAR to a broader set of variables by including, for example, the consumption–output ratio does not alter both the impact effect and the shape of the IRF to a technology shock. In particular, the impact effect is not significantly different from zero. The results are not reported in this version of the paper to save space but are available from the authors upon request.

on a long–run identification scheme. A test of the slope of the regression line of the two structural shocks indicates that a unitary slope cannot be rejected at the 5% conventional significance level (p–value : 55.14%). The latter result suggests that imposing short-run restrictions in our VAR model does not distort the estimated impulse responses of output and hours. Moreover, the confidence intervals are pretty narrow (see Figure 4), suggesting that these responses are very informative when used as an item to be matched in the MDE procedure.

The responses of output and hours to a technology shock identified through a short–run restriction are the objects to be matched by the DSGE model with information lags (model $\mathcal{M}_2$). More precisely, the vector of structural parameters $\theta = (b, \xi, \sigma_z)$ is estimated by matching the 11 responses of output and the 10 free responses of hours (the first point is excluded as its distribution is degenerated). We estimate the parameters using the MDE method (section 1.1). More precisely, we minimize the loss function defined in equation (1). The results are reported in Table 2. They first indicate that the full model cannot be rejected by the data at conventional

Table 2: A canonical experiment: Estimation Results

	b	ξ	σ_z	J–stat
Full Model	0.7253	3.3942	0.0126	9.2817
	(0.4337)	(1.7215)	(0.0006)	[95.30]
RBC Model	–	–	0.0102	102.6938
			(0.0004)	[0.00]

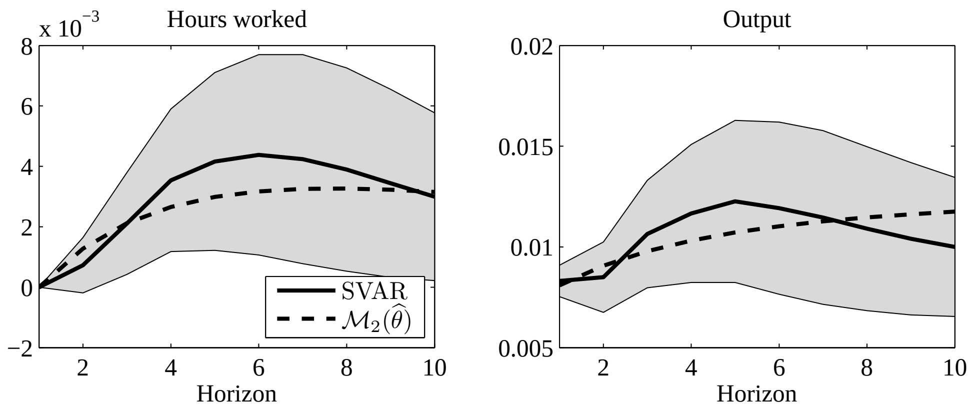
Note: Standard deviation into parenthesis, p–values (%) into brackets.

significance level (p–value=95.30). This is illustrated in panel (a) of Figure 4 which plots both the theoretical and the actual impulse responses of hours worked and output. The figure shows that the theoretical impulse responses lie within the 95% confidence interval of the corresponding impulse response found in the data, so that the model can be thought of as a good representation of the dynamics of the data. This result obtains because the model exhibits significant habit persistence and investment adjustment costs. For instance, the parameters ruling the size of these two phenomena are significantly estimated. A constrained version of the model — the standard RBC model — where both b and ξ are restricted to be zero is strongly rejected by the data, as can be seen from panel (b) of Figure 4. The failure of this model originates mainly from its inability to reproduce gradual increases in hours and output after a technology shock. Therefore the quasi likelihood ratio test (see Newey and West [1987]) of the joint significance of b and ξ strongly rejects the nullity of these two parameters.

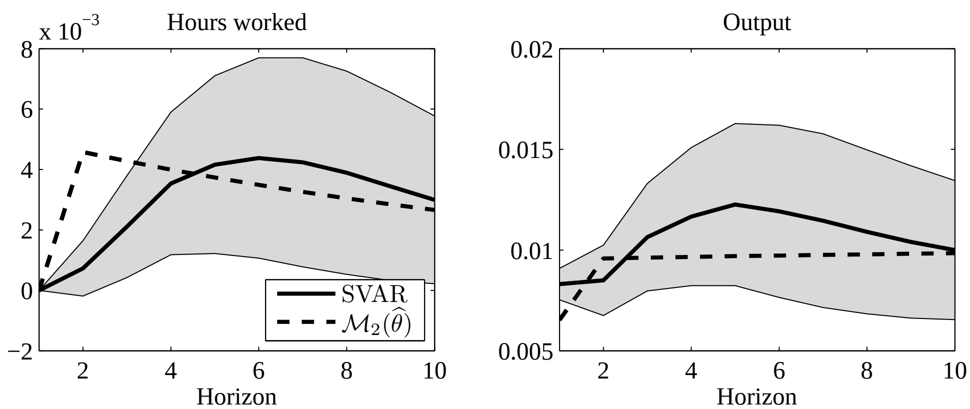
It is also worth noting that the estimated value of both parameters are within the range of values reported in earlier studies (see e.g. Christiano et al. [2005a], Boldrin et al. [2001], Christiano and

Figure 4: Impulse Response to a Technology Shock

(a) Full Model



(b) RBC Model



Fisher [2003]). A direct consequence of this result is that the model possesses strong propagation mechanisms that enhance its ability to account for the persistence of aggregates. For example, the second order serial correlation of output growth found in the data is 0.19. The model can account for 75% of it as it generates a second order autocorrelation of 0.14.

3 Short-Run Restriction as an Identification Device

In this section, we quantitatively assess the ability of short-run restrictions to identify the deep parameters and closely follow the approach described in Section 1.2.

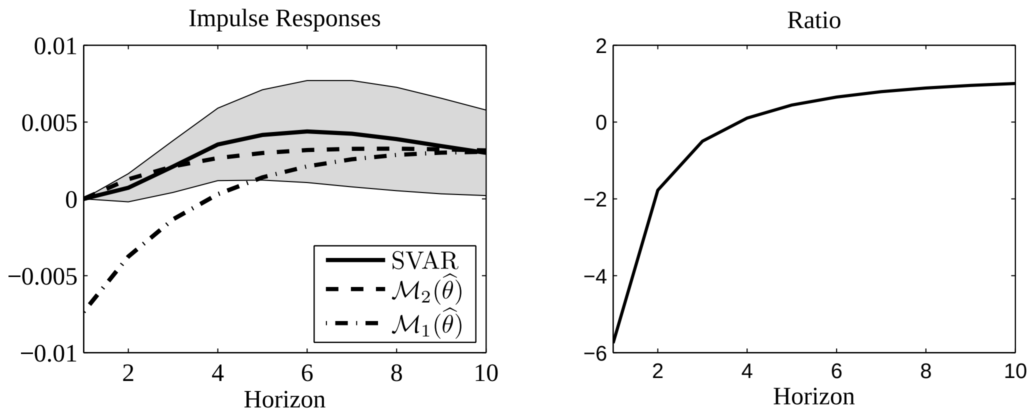
3.1 Assessing the Role of Timing Restrictions

Figure 5 superimposes the empirical SVAR model, fitted model $\mathcal{M}_2(\widehat{\theta}_T)$ and the model $\mathcal{M}_1(\widehat{\theta}_T)$. Let us recall that $\mathcal{M}_2(\widehat{\theta}_T)$ corresponds to the model where labor supply is decided prior to observing the shocks and is therefore compatible with the identifying restriction used in the SVAR, and $\mathcal{M}_1(\widehat{\theta}_T)$ is the model where labor can freely adjust to the technology shock. The first part of the figure essentially replicates Figure 4 and illustrates that the model is supported by the data. We then add the impulse response of hours in model $\mathcal{M}_1(\widehat{\theta}_T)$. Panel (a) of the figure indicates that a version of the model where hours are decided after the shocks are observed is rejected by the data. Not only is the impact effect of a technology shock different from zero — which should have been expected — but more interestingly the dynamics of hours worked implied by models $\mathcal{M}_1(\widehat{\theta}_T)$ and $\mathcal{M}_2(\widehat{\theta}_T)$ significantly differ at medium-run horizons. This is illustrated by the right part of panel (a) which reports the impulse response ratio (for the relevant innovation) for hours worked. The graph shows that the effects of information restrictions propagate over time in the model.

The mechanisms at work are simple. In model $\mathcal{M}_1(\widehat{\theta}_T)$, consumption adjusts only gradually and increasing investment is costly. Therefore, the labor supply drops sharply due to a strong income effect. However, after a few quarters, the intertemporal substitution effect dominates, and thus hours increase. In contrast, model $\mathcal{M}_2$ delivers a very different picture. In this case, hours do not respond on impact (by construction) and increase in all subsequent periods. Indeed, following a positive technology shock, households mildly increase consumption and investment on impact. Due to habit formation and dynamic investment adjustment costs, agents maintain high consumption and investment levels in subsequent periods. This leaves no other choice but to increase labor supply in order to sustain these plans. In equilibrium, hours persistently increase after a permanent technology shock. This result is confirmed by examining the autocorrelation function of hours worked, as reported in panel (a) of Figure 6. The persistence of hours worked is

Figure 5: Hours worked

(a) Full Model



(b) RBC Model

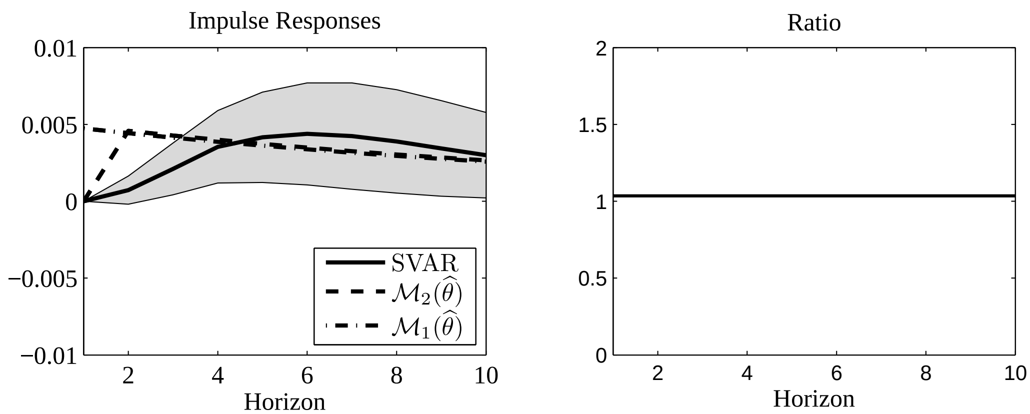
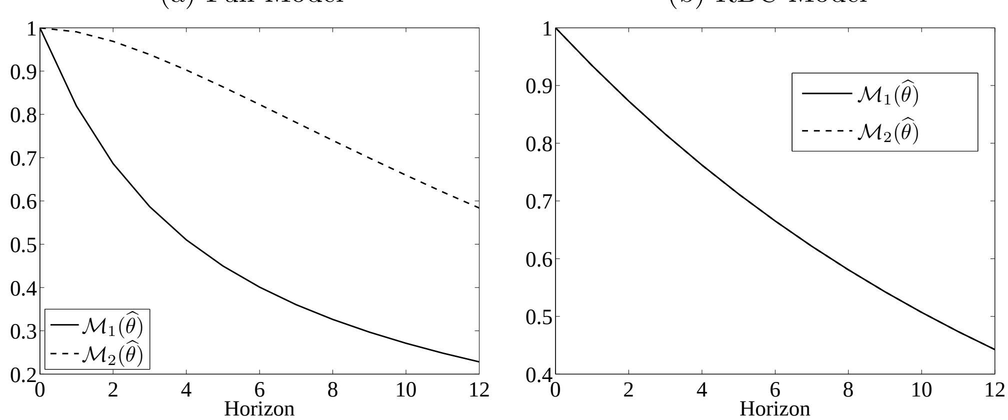


Figure 6: Autocorrelation function of Hours worked

(a) Full Model (b) RBC Model



enhanced in model $\mathcal{M}_2$ compared to $\mathcal{M}_1$, therefore illustrating that, contrary to the simple RBC model, the observability of shocks has a long–lasting impact on the labor supply decision. This result fundamentally originates from the existence of strong internal propagation mechanisms in the model.

In order to get a better understanding of this result, panel (b) of Figure 5 also reports the impulse responses of hours worked in the standard RBC model ($b = \xi = 0$), which is known to possess very weak internal propagation mechanisms (see Cogley and Nason [1995]). The striking feature that emerges from the graph is that except for the impact effect, the response of hours to a technology shock displays the same profile in both models. This is confirmed by the examination of the right panel of the figure that reports the ratio of the impulse response of hours at the time of the innovation with and without information restriction. The figure indicates that the dynamics of hours worked are the same in both model once properly lagged. Indeed, the simple RBC model incorporates no other real friction than the observability restriction. The latter only matters for one period. In the next period the dynamics of the model is back to that of the standard model, up to a scale effect. In other words, the observability restriction does not propagate over time and exerts no effect whatsoever on the dynamic properties of hours worked once the initial period has passed. This is confirmed by looking at the autocorrelation function of hours worked, as reported in panel (b) of Figure 6. The autocorrelation function of hours as obtained from the two versions of the model are the same, therefore illustrating that the information restriction does not affect the dynamics of hours worked in the model. In other words, shutting down propagation mechanisms implies that hours worked share the same dynamic properties in the two versions of the model. Information lags do not play any role for hours worked as long as propagation mechanisms are weak.[7]

To see this, we propose to gauge the effect of using the unrestricted model $\mathcal{M}_1$ in place of the restricted model $\mathcal{M}_2$ at the estimated value of the structural parameters $\widehat{\theta}_T$. Figure 7 plots the evolution of the rates of growth of output, labor productivity, consumption and investment in model $\mathcal{M}_1(\widehat{\theta}_T)$ as a function of the corresponding quantities in model $\mathcal{M}_2(\widehat{\theta}_T)$ when both models are fed with the technology shocks identified from the SVAR model. Table 3 reports the slope of the relation between the two sets of quantities as well as the p–value associated to the test of a unitary slope. The results indicate that the differences found in hours worked transmit to the other variables of the economy. In none of the graph are the two time series aligned on the 45° degrees line — which would be the case if the models were observationally equivalent. This is confirmed by examining the slope of the relationship between the two quantities. As indicated in Table 3, the slope is never found to be equal to one — the p–values associated to

[7] In the appendix, we consider a small DSGE model with habit persistence and no capital accumulation. This exercise illustrates analytically the role of internal persistence in the propagation of information restrictions over time.

the test are all very close to zero.

Figure 7: $\mathcal{M}_1(\widehat{\theta}_T)$ versus $\mathcal{M}_2(\widehat{\theta}_T)$

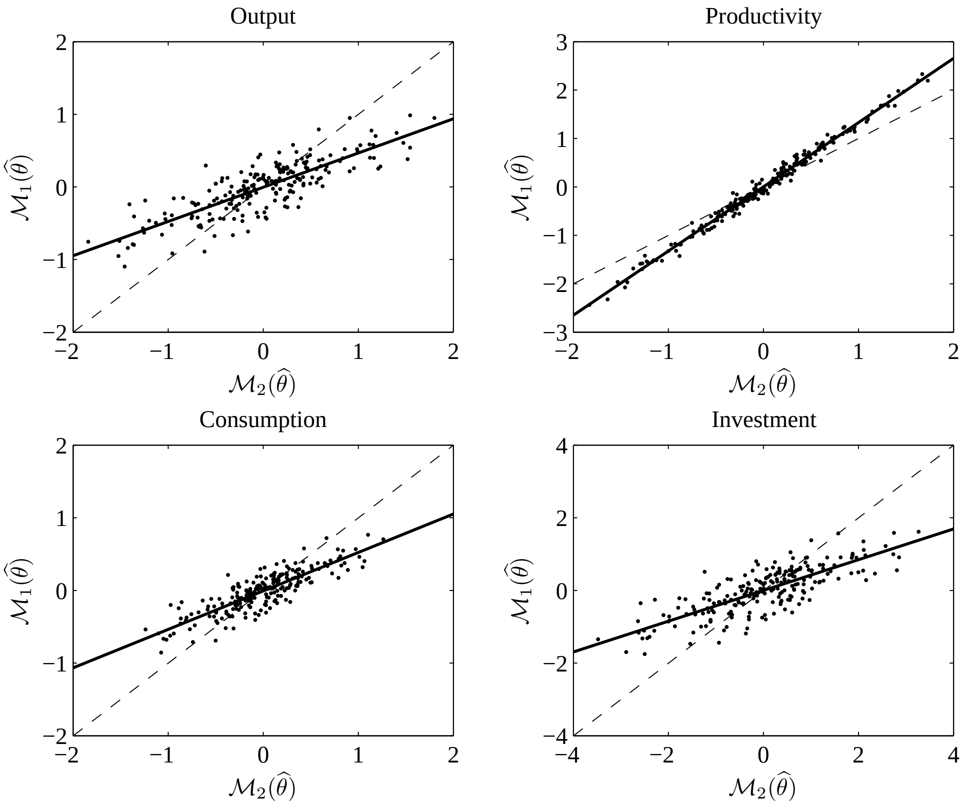


Table 3: Regression line $(\mathcal{M}_1(\widehat{\theta}_T), \mathcal{M}_2(\widehat{\theta}_T))$

	Output	Productivity	Consumption	Investment
Slope	0.4718	1.3268	0.5295	0.4233
	[0.00]	[0.00]	[0.00]	[0.00]

Note: p–values (%) into brackets.

This experiment illustrates that restricting the ability of agents to react to shocks — by assuming a particular short–run identification scheme — actually puts a lot of structure on the data and is far from being innocuous as soon as the model possesses internal mechanisms that ought to propagate information restrictions over time. In other words, beyond putting some identifying restrictions on the dynamic system, imposing short–run restrictions on the VAR model amounts to selecting a theoretical model. Beside this obvious observation, the structure of information may matter for the other variables, and therefore have important consequences on the estimation of the deep parameters.

3.2 Counterfactual Experiments

In order to get a better quantitative sense of the role of timing restrictions on the model, we now run the counterfactual experiments previously described in Section 1.2. We begin by obtaining a new value, $\widetilde{\theta}_T$, for the vector of structural parameters that makes the unrestricted version of the model ($\mathcal{M}_1$) observationally equivalent to $\mathcal{M}_2(\widehat{\theta}_T)$. This is achieved by minimizing the loss function defined in equation (2). An important aspect of this exercise is that it offers a quantitative assessment of the innocuousness of short run identifying restrictions. Indeed, were the estimation method efficient and immune to identification problems, it should be able to deliver the same estimation for the deep parameters, especially so if the objects to be matched are the responses of variables which are not constrained in the short–run (*i.e.* C, I, Y, Y/N).

We investigate this issue on output, labor productivity, consumption and investment, and run the experiment for the parameters pertaining to habit persistence, b, investment adjustment costs, ξ and the size of the shock σ_z. Note that only the first two parameters are deep parameters and should be left unaffected by the timing of the decisions. The last parameter is estimated in order to maximize the ability of the model to mimic the impulse response of each variable without affecting the deep parameters. Results are reported in Table 4. First of all, the results

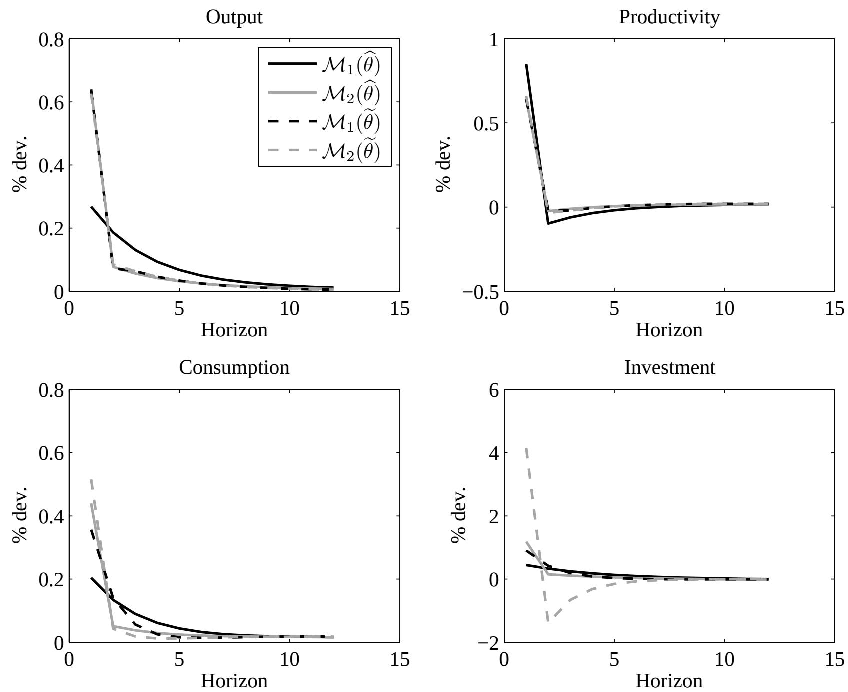
Table 4: Counterfactual Estimates

	ξ	b	σ_z	J–stat	$\mathcal{W}_1(\widetilde{\theta}_T, \widehat{\theta}_T)$
Benchmark	3.3942	0.7253	0.0126		
Output	0.9750	-0.0404	0.0124	0.1479	6.8710
	(0.3996)	(0.0929)	(3.4541e-4)	[100.00]	[7.60]
Productivity	1.0566	-0.1101	0.0130	0.2373	8.0250
	(0.6757)	(0.1778)	(3.6125e-4)	[99.99]	[4.50]
Consumption	1.3330	0.4950	0.0128	0.04392	2.9478
	(3.5999)	(0.5748)	(9.0620e-4)	[100.00]	[39.90]
Investment	0.2875	0.9855	0.0225	0.0677	616.5846
	(0.9353)	(0.0540)	(0.0422)	[100.00]	[0.00]
All	1.4061	-0.0414	0.0127	5.73815	5.7266
	(0.2972)	(0.0770)	(1.4832e-4)	[100.00]	[12.60]

Note: Standard deviation into parenthesis, p–values (%) into brackets.

indicate that no matter the macroeconomic variable under consideration, there always exists a parameterization of $\mathcal{M}_1$ that is statistically equivalent to $\mathcal{M}_2(\widehat{\theta}_T)$, as the J–stat test never leads to reject model $\mathcal{M}_1(\widetilde{\theta}_T)$. At first glance, this may be considered as good news since imposing short–run restrictions to identify shocks does not preclude the use and relevance of alternative theories, i.e. with or without information lags. This however also highlights that imposing short–run restrictions in a SVAR does not necessarily give enough information to guarantee full identification of the model. The results indicate that model $\mathcal{M}_1(\widetilde{\theta}_T)$ contains relevant information to account for model $\mathcal{M}_2(\widehat{\theta}_T)$.

Figure 8: Counterfactual IRFs (Growth rates)



Note: This graph depicts the dynamics of the growth rate of output, productivity, consumption and investment following a 1% postive technology shock. $\mathcal{M}_1(\widehat{\theta})$ denotes the unrestricted model fed with θ as estimated from the data using $\mathcal{M}_2$. $\mathcal{M}_2(\widehat{\theta})$ denotes the restricted model fed with θ as estimated from the data using $\mathcal{M}_2$. $\mathcal{M}_1(\widetilde{\theta})$ denotes the unrestricted model fed with θ as estimated so as to match $\mathcal{M}_2(\widetilde{\theta})$. Finally, $\mathcal{M}_2(\widetilde{\theta})$ denotes the restricted model fed with $\widetilde{\theta}$.

Let us first investigate the case where the variable used to perform the experiment is output. The first striking feature that emerges from Table 4 is that, as indicated by the p–value associated to the $\mathcal{W}_1(\widetilde{\theta}_T,\widehat{\theta}_T)$ test, the vector of parameter needed to match the output dynamics implied by $\mathcal{M}_2(\widetilde{\theta}_T)$ using $\mathcal{M}_1(\widetilde{\theta}_T)$ is significantly different from the initial parameter vector (see column 5 of the table). In particular, much lower adjustment costs are needed, and more importantly the model requires weak durability rather than habit persistence to match the dynamics. This triggers a drastic theoretical change in the model. Information sets therefore matter a lot. Regarding the dynamics of output, $\mathcal{M}_1$ is more informative than $\mathcal{M}_2$. Indeed, $\mathcal{M}_1$ is able to generate a variety of impact response of hours, depending on ξ. For example, when $\xi = 0$, the response is positive, while for very large ξ, the response becomes negative. Thus, there exists an intermediate value for ξ such that the impact response of hours is almost zero. As a consequence, under such a parameterization, the dynamics of output in $\mathcal{M}_1$ is similar to that generated by $\mathcal{M}_2$. Thus, a conclusion that can be drawn from this test is that a model with information lags on hours can be interpreted as a model with modest investment adjustment costs with unrestricted hours.

Table 5: Identification test ($\mathcal{W}_2(\widetilde{\theta}_T, \widehat{\theta}_T)$)

Output	Productivity	Consumption	Investment
1.5570	1.6873	0.1736	56.9746
[99.67]	[99.55]	[99.99]	[0.00]

Note: p–values (%) into brackets.

More worrying is that, once plugged back into $\mathcal{M}_2$, the new vector of parameters delivers the same output dynamics as with the initial vector of parameters. The p–value associated to the $\mathcal{W}_2(\widetilde{\theta}_T, \widehat{\theta}_T)$ test is 99.70%, reported in Table 5, indicating that one cannot reject that the two parameterization of model $\mathcal{M}_2$ generate the same output dynamics. This is first illustrated in Figure 8 which reports the impulse response of output growth in each version of each model. As can be seen from the figure, $\mathcal{M}_2(\widetilde{\theta}_T)$ and $\mathcal{M}_2(\widehat{\theta}_T)$ share the same dynamic properties. This is confirmed by examining the upper left panel of Figure 9 which plots the time series of output growth as implied by model $\mathcal{M}_2(\widetilde{\theta}_T)$ (in ordinate) against that in $\mathcal{M}_2(\widehat{\theta}_T)$ (in abscissae) when the models $\mathcal{M}_2(\widehat{\theta}_T)$ and $\mathcal{M}_2(\widetilde{\theta}_T)$ are fed with the technology shocks obtained from the VAR. As can be seen from the graph all points are aligned along the 45° line.

Table 6: Regression line ($\mathcal{M}_2(\widehat{\theta}_T), \mathcal{M}_2(\widetilde{\theta}_T)$)

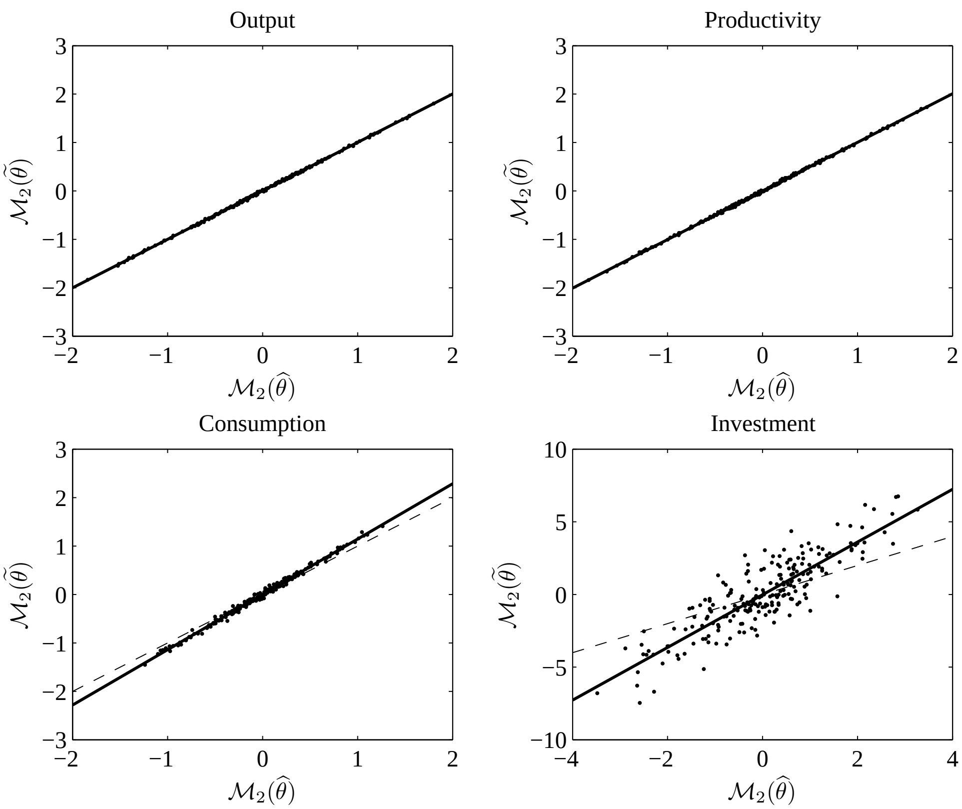
	Output	Productivity	Consumption	Investment
Slope	1.0014	1.0033	1.1417	1.8134
	[34.50]	[4.50]	[0.00]	[0.00]

Note: p–values (%) into brackets.

Table 6 then reports the slope of the regression line corresponding to the graph as well as the p–value associated to the test for a unitary slope. The test does not reject the unitary slope hypothesis, implying that the two time series are observationally equivalent. This actually points to a severe identification problem of DSGE models already highlighted in a simple framework by Canova and Sala [2005]. Although fundamentally different — one suggesting the existence of large investment adjustment costs and habit persistence, the other mitigating the former and suggesting (non significant) durability in consumption decisions — the two parameterization, once used in the same model, deliver statistically the same output dynamics.

This problem also arises when labor productivity is used as once again the p–value associated to $\mathcal{W}_2(\widetilde{\theta}_T, \widehat{\theta}_T)$ in Table 5 is 99.60%. A first glance at the response of consumption apparently

Figure 9: $\mathcal{M}_2(\widehat{\theta}_T)$ versus $\mathcal{M}_2(\widetilde{\theta}_T)$



suggests that consumption might deliver some useful information as the impact effect of a technology shock differs across models. However, a look at the standard errors associated to each coefficient in Table 4 reveals a strong identification problem as the standard deviations are very large. Furthermore, varying the initial conditions for the algorithm always leads to a situation were optimization delivers different outcome close the what the algorithm is fed with. This occurs because consumption, in the benchmark model, is so smooth that it does not provide enough variability to identify parameters. Therefore using consumption the econometrician would face severe identification problems. Investment dynamics, on the contrary, carry a lot of information as plugging the new parameters back into model $\mathcal{M}_2$ generates a fundamentally different dynamics. This is confirmed both by looking at the impulse responses reported in Figure 8 or the comparison plot in Figure 9. For instance, the test of a unit slope in the regression line associated to model comparison leads to strong rejection with a p–value close to zero. In addition, Table 5 shows that the p-value associated to $\mathcal{W}_2(\widetilde{\theta}_T, \widehat{\theta}_T)$ is zero. This indicates that $\mathcal{M}_2(\widehat{\theta}_T)$ and $\mathcal{M}_2(\widetilde{\theta}_T)$ fundamentally differ from each other and reveals another critical problem: the method does not really treat the structural parameters as deep parameters. Indeed, should the method be able to estimate deep parameters, these parameters should be statistically the same (deliver the same investment dynamics) in both model. What the lower right panel of the figure reveals is that it is definitely not the case and that great caution should be used when applying the method.

4 Concluding remarks

This paper investigates the quantitative implications of restricting the information sets conditional on which decisions are taken in DSGE models. To do so, we propose a variety of simple statistical tools designed to assess the role of timing restrictions on aggregate dynamics. We then consider an example of a fully-fledged DSGE model which we formally take to US data using the MDE approach based on short-run exclusion restrictions.

Our results indicate that restricting the information set is of no substantial consequences if one is dealing with a version of our DSGE model with weak internal propagation mechanisms, such as the basic Real Business Cycle model. However, in a DSGE model supported by the data and featuring a large number of real frictions that enhance its persistence properties, we obtain a very different picture. More precisely, we show by way of example that the Impulse Response Functions of certain variables can be modified in a marked way, i.e. in terms of sign, amplitude and persistence properties. More important, we obtain that the structural parameters estimated via the MDE approach are very sensitive to the timing of decisions, questioning the "depth" of such estimation techniques.

References

Altig, D., L. Christiano, M. Eichenbaum, and J. Linde, *Firm–Specific Capital, Nominal Rigidities and the Business Cycle*, Working Paper 11034, NBER 2005.

Altug, S., Time–to–Build and Aggregate Fluctuations: Some New Evidence, *International Economic Review*, 1989, *30*, 889–920.

Blanchard, O.J. and D. Quah, The Dynamic Effects of Aggregate Supply and Demand Disturbances, *The American Economic Review*, 1989, *79* (4), 655–673.

Boivin, J. and M. Giannoni, *Has Monetary Policy Become More Effective?*, Working Paper 9459, NBER 2005.

Boldrin, M., L.J. Christiano, and J.D.M. Fisher, Habit Persistence, Asset Returns and the Business Cycle, *American Economic Review*, 2001, *91*, 149–166.

Burnside, C. and M. Eichenbaum, Factor Hoarding and the Propagation of Business Cycle Shocks, *The American Economic Review*, 1996, *86* (5), 1154–1176.

——, ——, and S. Rebelo, Labor Hoarding and the Business Cycle, *Journal of Political Economy*, 1993, *101* (2), 245–273.

Canova, F. and G. de Nicolò, Monetary disturbances matter for business fluctuations in the G-7, *Journal of Monetary Economics*, 2002, *49* (6), 1131–1159.

—— and J. Pina, Monetary Policy Misspecification in VAR models, in C. Diebolt and C. Krystou, editors, *New Trends In Macroeconomics*, Springer Verlag, 2005.

—— and L. Sala, *Back to Square One: Identification Issues in DSGE Models*, mimeo, CREI and Universitat Pompeu Fabra 2005.

Chari, V.V., P.J. Kehoe, and E.R. McGrattan, *A Critique of Structural VARs Using Real Business Cycle Theory*, Working Paper 631, Federal Reserve Bank of Minneapolis 2005.

Christiano, L.J., Modelling the Liquidity effect of a Money Shock, *Federal Reserve Bank of Minneapolis Quarterly Review*, 1991, *Winter 91*, 3–34.

—— and J.D.M. Fisher, *Stock Market and Investment Goods Prices: Implications for Macroeconomics*, Working Paper 10031, NBER 2003.

—— and M. Eichenbaum, Liquidity Effects and The Monetary Transmission Mechanism, *American Economic Review*, 1992, *82* (2), 346–53.

——, ——, and C.L. Evans, Sticky Price and Limited Participation Models of Money: A Comparison, *European Economic Review*, 1997, *41*, 1201–1249.

——, ——, and ——, Monetary Policy Shocks: What Have we Learned and to What End?, in M. Woodford and J. Taylor, editors, *Handbook of Macroeconomics*, North-Holland, 1999, chapter 3.

——, ——, and ——, Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy*, 2005, *113* (1), 1–45.

——, ——, and R. Vigfusson, *What Happens After a Technology Shock?*, mimeo, Northwestern University 2004.

——, ——, and ——, *VARs as a Guide to Estimating Dynamic General Equilibrium Models*, mimeo, Northwestern University 2005.

Cogley, T. and J.M. Nason, Output Dynamics in Real–Business–Cycle Models, *The American Economic Review*, 1995, *85* (3), 492–511.

Eichenbaum, M. and J.D.M. Fisher, Fiscal Policy in the Aftermath of 9/11, *Journal of Money, Credit and Banking*, 2005, *37* (1), 1–22.

Erceg, C., L. Guerrieri, and C. Gust, Can Long-Run Restrictions Identify Technology Shocks?, *Journal of the European Economic Association*, 2005, *3* (5), Forthcoming.

Galí, J., Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?, *American Economic Review*, 1999, *89* (1), 249–271.

—— and P. Rabanal, *Technology Shocks and Aggregate Fluctuations: How Well Does the RBS Model Fit Postwar U.S. Data?*, Working Paper 10636, NBER 2004.

Giannoni, M. and M. Woodford, Optimal Inflation-Targeting Rules, in B. Bernanke and M. Woodford, editors, *The Inflation Targeting Debate*, Chicago: University of Chicago Press, 2004, pp. 93–162.

Hansen, G., Indivisible Labor and the Business Cycles, *Journal of Monetary Economics*, 1985, *16* (3), 309–27.

Hansen, L.P., Large Sample Properties of Generalized Method of Moment Models, *Econometrica*, 1982, *50*, 1269–1286.

Ireland, P., A Method for Taking Models to the Data, *Journal of Economic Dynamics and Control*, 2004, *28* (6), 1205–1226.

Newey, W.K. and K.D. West, Hypothesis Testing with Efficient Method of Moments Estimation, *International Economic Review*, 1987, *28* (3), 777–787.

Obstfeld, M. and K. Rogoff, Exchange Rate Dynamics Redux, *Journal of Political Economy*, 1995, *103* (3), 624–660.

Rotemberg, J. and M. Woodford, An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy, in B. Bernanke and J. Rotemberg, editors, *NBER Macroeconomics Annual*, Cambridge (MA): MIT Press, 1997, pp. 297–346.

Schorfheide, F., Loss Function Based Evaluation of DSGE Models, *Journal of Applied Econometrics*, 2000, *15* (6), 645–670.

Sims, C., Macroeconomics and Reality, *econometrica*, 1980, *48* (1), 1–48.

Smets, F. and R. Wouters, An estimated stochastic dynamic general equilibrium model of the euro area, *Journal of European Economic Association*, 2003, *1*, 1123–1175.

Uhlig, H., What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics*, 2005, *52* (2), 381–419.

Woodford, M., *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press, 2003.

A The Role of Propagation Mechanisms: A Simple Model

This appendix shows, using a simple example, how internal propagation mechanisms are critical to propagate information lags. Notations and information assumptions on model $\mathcal{M}_1$ and $\mathcal{M}_2$ are the same as is in the main text. Let us consider a simple DSGE model with habit formation in consumption and a permanent technology shock. Preferences are represented by the following lifetime utility function

$$\mathscr{E}\left[\sum_{s=0}^{\infty}\beta^s\left\{\log(C_{t+s}-bC_{t+s-1})-\chi N_t\right\}\middle|\mathscr{I}_t\right] \tag{7}$$

The parameter $\beta \in (0,1)$ is the subjective discount factor and $b \in [0,1)$ governs the evolution of consumption habits. Finally, $\chi > 0$ is the marginal disutility of labor. $\mathscr{E}$ denotes the expectation operator conditional on the information set $\mathscr{I}_t$. The household is subject to the simple budget constraint $C_t \leqslant W_t N_t$ where W_t is the real wage.

The representative firm produces a homogenous good with a constant returns to scale technology represented by the simple production function $Y_t = Z_t N_t$. Z_t denotes the level of aggregate technology, and is assumed to follow an exogenous stochastic process modeled as a random walk. The rate of growth of technology,$\gamma_{z,t} \equiv \log(Z_t/Z_{t-1})$ is assumed to evolve according to

$$\gamma_{z,t} = \rho\gamma_{z,t-1} + (1-\rho)\gamma + \sigma\varepsilon_{z,t} \qquad |\rho| < 1\ ,\ \sigma > 0$$

where $\varepsilon_{z,t}$ is *iid* with zero mean and unit variance.

In equilibrium, we have $Y_t = C_t = Z_t N_t$. Deflating for the stochastic growth component and log–linearizing the economy around its deterministic steady state, it is simple to obtain the decision rule for labor decisions in both model

$$\mathcal{M}_1:\ \widehat{n}_t = \frac{b}{\gamma}\widehat{n}_{t-1} - \frac{b\mu}{\gamma}\widehat{\gamma}_{z,t} \qquad \text{and} \qquad \mathcal{M}_2:\ \widehat{n}_t = \frac{b}{\gamma}\widehat{n}_{t-1} - \rho\frac{b\mu}{\gamma}\widehat{\gamma}_{z,t-1}$$

where $\mu \equiv \frac{\gamma(1-\beta\rho)}{\gamma-\beta\rho b} > 0$.

The dynamics of output growth is then simply given by

$$\mathcal{M}_1 \quad : \quad (1-\rho L)\left(1-\frac{b}{\gamma}L\right)\Delta\hat{y}_t = \left(1-\frac{b\mu}{\gamma}\right)\left(1-\frac{b(1-\mu)}{\gamma-b\mu}L\right)\varepsilon_{z,t}, \tag{8}$$

$$\mathcal{M}_2 \quad : \quad (1-\rho L)\left(1-\frac{b}{\gamma}L\right)\Delta\hat{y}_t = \left(1-\frac{b}{\gamma}(1+\mu\rho)L+\frac{b\mu\rho}{\gamma}L^2\right)\varepsilon_{z,t}. \tag{9}$$

It is important to note that the solution for output growth is fundamental in both model. Therefore $\varepsilon_{z,t}$ is indeed the innovation of output growth. In the first model, output only partially reacts to the shock because of the initial decline in hours. In contrast, in the second model, because the firm cannot adjust hours in the first period, the output responds one for one to a technology shock. As is clear from equations (8) and (9), models $\mathcal{M}_1$ and $\mathcal{M}_2$ share the same AR structure. But their MA components are fundamentally different. Model $\mathcal{M}_1$ possesses an MA(1) component while model $\mathcal{M}_2$ exhibits a MA(2) component. This therefore affects the dynamic behavior of output growth and therefore shows up in both its impulse response and autocorrelation function. On the contrary, when the latter are shut down — i.e. when $b = 0$ — both models deliver the exact same output dynamics for any $\rho \in (-1,1)$. Indeed in such a case, output dynamics reduces to

$$(1-\rho L)\,\Delta\hat{y}_t = \varepsilon_{z,t}$$

in both model, and information lags do not matter anymore. This reflects the fact that when the models do not have any internal propagation mechanisms, information sets do not matter.