IMF Working Paper

System Priors:

Formulating Priors about DSGE Models' Properties

Michal Andrle and Jaromir Benes

IMF Working Paper

Research Department

System Priors: Formulating Priors about DSGE Models' Properties[1]

Prepared by Michal Andrle and Jaromir Benes

Authorized for distribution by Benjamin Hunt

December 2013

Abstract

This paper proposes a novel way of formulating priors for estimating economic models. System priors are priors about the model's features and behavior as a system, such as the sacrifice ratio or the maximum duration of response of inflation to a particular shock, for instance. System priors represent a very transparent and economically meaningful way of formulating priors about parameters, without the unintended consequences of independent priors about individual parameters. System priors may complement or also substitute for independent marginal priors. The new philosophy of formulating priors is motivated, explained and illustrated using a structural model for monetary policy.

Author's E-Mail Address: mandrle@imf.org, jbenes@imf.org

Contents   Page

Tables

Figures

I. INTRODUCTION

This paper proposes a novel approach to elicit individual parameter priors using views on dynamic models' properties – system priors. The Bayesian estimation paradigm uses the likelihood principle and updates prior beliefs about the parameters with the information conveyed by the data. The prior beliefs about many individual structural model parameters are often difficult to formulate. When priors about individual parameters are specified as independent, the implied prior distributions for the model's properties as a system may be rather infeasible or take an unintended form. It can, thus, be more transparent or even easier to express prior views about the properties of the model directly and back-out the implied priors for parameters. This may entail, for instance, priors on properties such as impulse response functions, conditional correlations or the size of a cumulative output loss after a permanent disinflation (the sacrifice ratio). Our new method of devising priors is about just that – formulating priors on individual parameters via prior restrictions on system's properties. We label these priors "system priors" and use this term henceforth.

There are many benefits of employing system priors with very little cost of implementation. System priors introduce complex dependence among structural parameter priors, and allow for imposition of intuitive and economic priors about the aggregate behavior of the model. They also avoid unintended consequences on model's properties frequently associated with marginal independent priors. In technical terms, the implementation is just a straightforward extension of standard Bayesian methods. The main cost of implementation remains formulating relevant, transparent, and economic meaningful beliefs about the model's behavior. This paper provides guidance on all the issues mentioned.

The contribution of the paper lies mainly in the proposal to change the philosophy of formulating priors, not in technical implementation details, though these are covered as well. We argue for more careful design of priors, reflecting a priori restrictions on the system properties of macroeconomic models, where the deep nature of behavioral parameters may not be as deep as often believed, with identification issues and complex interactions among parameters contributing to the complexity of the problem. We also argue, following Geweke (2010), for much greater emphasis on the prior predictive analysis of models and null-hypothesis, so it is clear whether it is the data or the model, a priori, that is responsible for the final estimates of system properties. Note, this information cannot be trivially obtained from looking at marginal posterior and prior distributions of structural coefficients. To give an example, reporting a prior density of impulse response functions is quite rare in the literature.

Formulating parameter priors using prior views on selected model properties introduces complex dependence among parameters. Imposing a prior restriction on selected properties of the model, e.g. impulse response function, frequency response function, or conditional correlations, introduces restrictions on a joint distribution of structural parameters. This feature is quite distinct from the case of independent marginal priors about individual parameters, which introduce none. Furthermore, individually very reasonable parameter priors can result in rather unreasonable aggregate model properties, different from the researcher's beliefs, due to the nonlinear mapping of parameters into model's properties. In contrast, a prior about system properties of the model creates direct stochastic restrictions on the combinations of parameters. For instance, the parameters driving price and wage rigidities, monetary policy response, or interest rate sensitivity of output will be closely linked when the priors on the horizon of monetary authority's reaction to inflation-forecast deviation from the target are put in place.

Priors expressed in terms of model properties seem useful and effective, as sometimes the relevant information cannot be obtained from the likelihood function of the model. Although the likelihood function of the model fully describes its properties with respect to observed data, such information may not always be sufficient for parameterizing models to be used for policy analysis. There are requirements on the behavior of the model that reach beyond sample data, but resonate well with economic theory and intuition. Disinflation policies, for instance, are rarely observed in the macroeconomic data frequently, but the experience from other countries and historical episodes can discipline parameter estimation of the model using a sacrifice ratio prior.

System priors may facilitate model comparison and comparison of priors among economists. Models designed to explain a common dataset may have very different structures or set of structural parameters. The priors about the model's properties, system priors, are higher-level priors and may be much more coherent across models than the parameter values themselves. When Faust (2009) gives an example of two prior beliefs economists may have, (i) a consumption growth insensitive to transitory changes in short-term interest rates and (ii) a prior about long and variable lags of monetary policy, system priors allow researches to directly implement those. This is a significant difference from the standard literature where largely arbitrary priors, in Faust's terms, may not guarantee that the economics of priors is satisfied.

Translating system priors to individual parameter priors is relatively straightforward. We express higher-level system priors as a set of stochastic restrictions on parameters. Together

with a set of marginal priors, the system priors works as a penalty for the likelihood function of the model. The composite prior is usually not analytically tractable but can be easily obtained using standard sampling techniques. Since dynamic stochastic general equilibrium (DSGE) models are frequently estimated using Monte Carlo techniques, incorporating system priors amounts to adding only one additional step to the estimation – evaluating system priors on top of evaluating marginal priors and the likelihood function.

Our contribution can be related to some existing literature. We build on and enrich the literature dealing with Bayesian estimation of dynamic economic models following An and Schorfheide (2007) and others. We extend the analysis of Geweke (2010), Faust and Gupta (2012) or Faust (2009), who point out the importance of prior and posterior predictive analysis of the model. We are, however, directly shaping our parameter priors around the model implied properties. On a more operational level, our computations are are close to Del Negro and Schorfheide (2008), who demonstrate how to formulate priors based on pre-sample data and long-run stylized facts of the economy, but our philosophy is very distinct from theirs. More recently, Jarociński and Marcet (2010) use a prior on output growth expectations at the beginning of the sample to elicit priors for an autoregressive model. Our approach is more general, it encompasses the methods indicated, and we focus strongly on the economics of the prior elicitation.

The outline of the paper is as follows: The second section of the paper discusses the motivation of our approach and its formal implementation in greater detail. The third section uses a simple New-Keynesian model to illustrate how system priors induce individual parameters' priors. Then we conclude.

II. SYSTEM PRIORS – FORMULATING PRIORS ABOUT MODELS' PROPERTIES

A. Motivation

Building an economic model useful for policy analysis requires both art and science, and is very much about beliefs. Prior beliefs determine the structure of the model, selection of variables and, of course, the range and prior distribution of parameter values. The process of model building and parameterization should be as transparent as possible. This is not always the case, as documented by Leamer (1978) or Faust (2009), among others. The Bayesian approach to estimating dynamic economic models is, for some, a way to make the process

more transparent. For others, it can be a pragmatic way to regularize numerical difficulties associated with the estimation problem.

The priors for dynamic economic models often are relatively arbitrary and designed bottom-up. The priors for vector autoregressive (VAR) models or dynamic stochastic general equilibrium (DSGE) models in the literature often are rather arbitrary. Priors for VAR models tend to be mechanistic, individually set, with possibly counter-intuitive results for models' properties, say for its steady state.[1] In the case of DSGE models, where individual parameters usually have at least some structural interpretation, the situation is better. However, the independent marginal priors about bits and pieces of the model may lead, and often do, to unintended consequences for the properties of the model implied by the prior – e.g. steady-state values, slope of the Phillips curve, cross-correlations among variables, or implied horizon of the monetary policy reaction and effectiveness.

Many economists have strongly held beliefs about the aggregate behavior of the model, yet less so for arcane, model-specific constructs and parameters. We, economists, have priors about monetary policy horizon, conditional cross-correlation of consumption and investment, sensitivity of consumption to changes in short-term interest rate, or a cumulative loss of output after a disinflation, for instance. In terms of these priors we can communicate across models irrespective of their structure, state- or time-dependent pricing, or labor market structure in the model, for instance. The individual structural parameters sometimes are not so structural after all. For instance, a key parameter in an often used time-dependent pricing scheme due to Calvo (1983) has a very stylized interpretation and its structure does not enable researches to identify real and nominal rigidities separately, see e.g. Coenen, Levin, and Christoffel (2007). What matters is just the slope of the Phillips curve and the resulting model dynamics.[2] The prior across models about the Calvo price rigidity parameter is thus not comparable or transferable. When building the prior bottom-up, using individual independent parameters only, it is hard to implement the high-level system priors researchers have in their minds.[3]

It is essential to understand all implications of the priors for the model's behavior, although this is all but ignored in applied work. It is hardly possible to come up with good parameter

---

[1]The two-step prior by Villani (2009) about the model steady-state growth is a rare exception to the rule. Note that a constant term of a VAR(p) model is not its steady-state.

[2]For instance, with the assumption of firm-specific capital, the slope of the Phillips curve features an additional parameter that appears nowhere else in the model. The interpretation of the Calvo price parameter then changes dramatically, without any change in the model's dynamics.

[3]Our experience with models seem to suggest that vague priors lead to unintended consequences on model's properties, whereas overly tight priors do not allow data to speak enough. Using system priors allows us to delineate boundaries within which data can speak quite strongly.

priors without a thorough understanding of the model at hand. Sensible individual-parameter distributions convolute in a non-linear way into aggregate model properties, which can be unintuitive, unintended, and often go unchecked. While there is no substitute for a careful examination of model properties, the *prior-predictive analysis* is one of the tools that allow analysts to inspect the distribution of any system property chosen. Disregarding the prior-implied distribution of model properties is hazardous; at least a prior-mode or mean properties (impulse-response function, second moments, etc.) of the model should be contrasted to a posterior result. Plots of marginal prior and posterior parameter distributions just won't do.

High-level system priors– priors about model's properties– are an effective way of expressing economically meaningful priors directly. A prior about system properties ties together individual parameter combinations that satisfy the stochastic constraints imposed. For example, what is the prior view about the maximum length of the inflation deviation from the inflation target after a demand shock? For a given parameterization of the model, a simulation is performed, a test criterion evaluated and assessed using the prior distribution for the test statistic, the duration of the inflation response. The analyst puts forth a direct prior about the system's property. It will not be the case that the duration of the inflation cycle implied by bottom-up independent marginal priors, possibly without the analyst's intention, restricts the response in such a way the data cannot change. A prior predictive analysis would detect the issue, though would not provide a method to implement the prior. System priors do.

After realizing the usefulness of system priors, a variety of sensible priors that economists can devise and test open up, which are discussed in the next section. The range of priors economists exercise in their calibration exercises and estimation 'specification searches' is wide. With system priors' transparency of formulating the prior, the situation is much clearer than with 'standard' bottom-up priors.[4] In effect, most economists estimating DSGE models entertain a particularly dogmatic system prior already – the condition of saddle-path stable, or Blanchard-Khan stable, solution of the model.[5] Only a subset of parameters implied by the prior is feasible then. System priors about the model's properties work as a non-dogmatic version of the same principle, introducing cross-dependence among parameters to satisfy the set of stochastic constraints.

---

[4]One may not agree with other one's priors, but at least these are clearly stated and exposed to criticism and discussion.

[5]For instance, Geweke (2007) points out the fact that the prior distribution for DSGE models is truncated to the subset of the parameter space corresponding to a saddle-path stable solutions of the model and its implications. Note also that the Blanchard-Kahn stability breaks the marginal prior independence and thus can affect identification of marginally unidentified parameters, see Poirier (1998) and our discussion below.

B. Candidates for System Priors for DSGE Models

Candidates for properties that can be used to formulate system priors on model parameters abound. The system properties used for the estimation will follow from the specific goals of the model, the prior beliefs of the analyst, and the particular application or test at hand. Any statistic implied by the model is, in principle, amenable to be employed as a system prior.

The following is a selective list of model-implied features that are worthy of consideration as system priors. This list is by no means exhaustive and aims to stimulate further discussion.

- **Steady-state values of the model variables:** It is necessary to assure that the prior guarantees a sensible steady-state of the model, in terms of level or growth rates of variables.[6]

- **Conditional or unconditional moments of the model:** Cross-correlations of variables, conditioned on a subset of shocks, are an important candidates for system priors. For instance, the co-movement of investment with private consumption after a demand shocks can be judged positive a priori.

- **Prominent policy scenarios:** Counter-factual policy scenarios informed by experience from other countries, microsimulation models, or expert judgement are prime candidates for system priors. Key policy experiments we identified are (i) permanent disinflation and (ii) anticipated delayed response of monetary policy for several periods.

- **Characteristics of impulse-response functions:** – Prior beliefs about the peak impacts of shocks, duration of recessions, or the horizon of monetary policy effectiveness, among others, can be expressed in explicit terms using impulse-response functions and model simulations.

- **Frequency response function and spectral characteristics:** Spectral characteristics of the model and the implied filter are complex functions of the structural parameters. Assumptions about signal-to-noise ratios, power of shocks at particular frequencies for selected variables, or coherence between macro variables can be used to specify an important class of system priors. Frequency-domain system priors are particularly useful for thinking about trend, cycles, and high-frequency dynamics of the model.

---

[6]This point is especially acute with respect to time-series models, namely VARs.

The system priors can be constructed both in relative or absolute terms but should be robust to trivial changes in scaling. In the process of translating the prior beliefs into numerical characterizations, either of a scalar form (sacrifice ratio) or a vector form (impulse response path) can be used. Our experience with system priors suggests it is easier to work with a single summary statistic in most cases. We've also learned that the need to formalize and quantify one's prior beliefs leads to much refined and careful analysis.

C. Examples of System Priors

This section gives examples of particular system priors that are found are simple and powerful ones to use for estimation of DSGE models. The examples below cover (i) the sacrifice-ratio prior, (ii) a delayed response of monetary policy, (iii) duration priors about the model dynamics, and (iv) frequency domain prior about measurement errors. Obviously, a great variety of system priors is available and our lists and examples are not exhaustive. These are our favorites that we have tested in practice in several projects.

1. Sacrifice-Ratio Priors

It is the great policy relevance, sensitivity with respect to key parameters, and limited information in the sample data that make the sacrifice ratio a prime candidate for system priors. The sacrifice ratio measures the cumulative loss of output after a permanent, unannounced 1% permanent reduction in the central bank's target rate of inflation. The sacrifice ratio is one of the key statistics and 'smell tests'[7] for a monetary model and the role of monetary policy. For an exemplar use of the concept, see e.g. FED (1996), Laxton and Pesenti (2003), or Fuhrer (1994), among others. The sacrifice ratio marries both theoretical rigor and great policy relevance.[8] Another interesting aspect of the sacrifice-ratio statistic is that disinflations are rare or not present in the estimation sample. The likelihood function of the model is thus not very informative about this model property. The prior, however, can be elicited using the experience from historical episodes in other countries, periods not covered by the sample, or simply using expert judgment and caution.

---

[7]Our use of the term is broadly defined by the question of (Solow, 2010, pp. 2): "Does this really make sense?"

[8]In his discussions of Fuhrer (1994), Gregory Mankiw acknowledges the importance of the sacrifice ratio analysis: "It is a rare pleasure to read a paper about the sacrifice ratio written by someone under the age of 50. The sacrifice ratio is one of those subjects in macroeconomics that is at the heart of many practical policy discussions but, at the same time, rarely finds its way into serious academic publications. It is good to see someone trying to be both practical and serious at the same time." (Fuhrer, 1994, pp. 70)

The sacrifice-ratio prior is a highly non-linear function of important dynamic structural parameters. The sensitivity of a system prior with respect to key parameters is of course important. In the example application below, it is demonstrated that the sacrifice ratio can be very responsive to changes in the parameterization. Depending on the hyper-parameters of the system prior, e.g. the distribution, mean and variance of the sacrifice ratio, any parameterization deviating far from the prior distribution is appropriately penalized.

From the global identification point of view, the sacrifice-ratio system prior can be very powerful. The effort to implement system priors is not undertaken in order to render those completely uninformative. In our empirical work, an informative prior about the sacrifice ratio has proven to be an important identification assumption. It is able to select between different parameterizations with very similar likelihood and dynamics, but having dramatically different policy implications in terms of the sacrifice ratio, notably in open economy model.

2. Delayed policy response

One of the frequent questions in policy institutions, namely central banks, is about the effects of a delayed change in a policy instrument. The scenario may entail a demand shock followed by keeping the policy rate unchanged for k periods. The longer the markets expect no change in the policy instrument, the more accommodative the policy is. A statistic of interest can be how much more the inflation increases compared to a standard reaction of the policy when it follows a variant of a Taylor rule, for instance. Such a statistic is very nonlinear function of parameters– a statistic that goes beyond the properties of the model that can be directly observed in the data. The simulation entails anticipated actions and thus uses both stable and unstable dynamics of the model, allowing for richer identification options than the likelihood, as explained below.

3. Duration Priors

Duration system priors proved useful in the model estimation to prevent a priori unreasonable dynamics. We, economists, seem to often have strongly held priors about the duration or timing of economic phenomena. We have priors about the duration of business cycles, or the adjustment of the economy to particular shocks. For instance, inflation-targeting central banks usually operate within a 'monetary policy horizon' – a horizon where monetary policy is effective and during which inflation should return to its target after the initial shock. A

prior suggesting that no significant deviation of inflation from its target is to be found after the fourth year would seem like a rather consensual prior. It may work as an upper bound on the response of inflation.

Scale-invariant duration priors are relatively easy to implement. Our suggestion is to compute the integral of the particular variable's response in absolute value up to horizon T and $T + h$ and express the prior about the relative contribution of the h-period horizon, that is as a ratio of these two numbers. No dynamics between periods T and $T + h$ implies that the statistic of interest is one. Such a definition of dynamics is invariant to the size of the shock in a linear model; in non-linear models a variety of modifications are possible. If needed, the measure of persistence of the inflation response can be related to the persistence of the output expansion.

4.   Frequency-domain Priors

System priors based on statistics computed in the frequency domain are abundant and offer great flexibility and control over the estimation process. Using marginal independent priors about structural parameters to limit an property of the model a priori is often extremely hard, or impossible. The equivalence between time domain and frequency domain allows researchers to use whichever tool that seems most appropriate for the purpose at hand. When the contribution of trends, cycles, or measurement errors is of a concern, the frequency domain is the intuitive and correct approach. We discuss some examples of system priors in the frequency domain below.[9]

Measurement-error priors can be refined by restricting a share of variance or periodicity over which these should not affect the observed variable. Standard priors on measurement error's variance do not restrict the role of measurement errors in a sophisticated way. In the case of an uncorrelated measurement error process, its spectral density is flat and the contribution to the estimated variance of the observed endogenous variable can be easily constrained a priori. For correlated measurement errors and more general processes, some a priori information up to what frequencies the measurement error can contribute to is also easy to implement. For each parametrization of the model, the contribution of measurement error to population or sample variance can be evaluated and penalized for a deviation from the values implied by the prior distribution of the statistic.

---

[9]See e.g. Hamilton (1994) or Koopman (1974) for elements of frequency-domain analysis of time series. Familiarity with the concept of spectral density, a distribution of variance across cycles of different periodicity is, is useful for this section.

The relationship among variables at cyclical frequencies and the nature of their trends often requires a priori information. The issue of trends specification or de-trending is a big one in the case of DSGE models, see e.g. Canova (2012) for a lucid discussion. The system priors can be associated with the frequency domain representation of the model, or with the frequency response function of the Wiener-Kolmogorov or Kalman filter implied by the structural model. The properties of the implied filter are fully described by the model's parameterization and hyper-parameters determining the treatment of initial conditions. For instance, the mid-point of the filter's implied gain function for particular variable to another is a feasible option for setting up a system prior on the model.

In simple setups, one can parameterize the model such that an analytical prior on signal-to-noise ratio is feasible, otherwise a system prior needs to be specified. A prior on the signal-to-noise ratio may be needed, for instance, to guard the estimation from the pile-up problem, see e.g. Stock (1994) or Laubach (2001).[10] In more complex models with endogenous trends, or in trend-cycle models, as in Andrle and others (2009) or Canova (2012), system prior on implied signal to noise ratio and frequency response function of the model provide the only and efficient way to express a priori views on trend properties, despite having no closed-form solution.

D. Implementing System Priors

The method for implementing system priors can be interpreted both in the coherent Bayesian or classical framework. In the classical framework, the system priors are equivalent to *penalty function,* or regularization, methods.[11] The following discussion is centered around the proper and coherent Bayesian view, with no loss of generality. Given the model $\mathcal{M}$, the Bayesian approach to parameter estimation consists of updating the prior beliefs, $\tilde{p}(\theta|\mathcal{M})$, about the parameters, θ, with the evidence using the data, Y^o, and the model's likelihood, $L(Y^o|\theta, \mathcal{M})$, i.e.

$$p(\theta|Y^o, \mathcal{M}) = \frac{L(Y^o|\theta, \mathcal{M}) \times \tilde{p}(\theta|\mathcal{M})}{p(Y^o|\mathcal{M})} \propto L(Y^o|\theta, \mathcal{M}) \times \tilde{p}(\theta|\mathcal{M}), \tag{1}$$

---

[10]Essentially, the pile-up problem concerns the estimator putting too little weight on the variance of permanent component.

[11] It is well known that many regularization functions have a Bayesian interpretation. For instance, quadratic penalty function corresponds to the Gaussian distribution prior on the feature, or the Lasso method can be understood as the Laplace prior on the feature.

The expression (1) is completely general, without the type of the prior distribution function $\tilde{p}(\theta|\mathcal{M})$ specified. In the case of DSGE models, the prior information is virtually always formulated in terms of independent marginal priors on components of θ. In our case, the prior distribution is elicited, also, with the help of *system priors*.

To implement the system priors, the composite prior is built by first specifying marginal distributions for individual parameters, and then augmenting it with system priors in the second step. Backing out the individual parameter priors using the system priors boils down to an *inverse problem.* We do not require the system priors to identify all parameter priors by themselves; one can start with a set of independent marginal priors on structural parameters, which are further updated using non-sample information in the form of the system prior. There is an analogy to dummy observations priors, see Theil and Goldberger (1961). Although there are cases where a system prior can be implemented using a change-of-variables formula or other analytic calculations, we propose relying fully on numerical specification and Monte Carlo integration techniques, which fit well into the standard estimation setup of DSGE models.

Specification of the independent marginal priors on parameters is a feasible starting point for the composite prior. The first restriction, a prior, on a parameter is expressed as a joint distribution composed of marginal distributions of elements in θ, i.e.

$$p(\theta|\mathcal{M}) = p(\theta_1) \times \cdots \times p(\theta_k). \tag{2}$$

This is the most common structure of a prior specification in the DSGE literature, see e.g. Canova (2007). Individual parameters are considered independent and marginal distributions $p(.)$ determine range and plausible location of individual parameter values. The independence assumption is unrealistic and is used here for convenience, as it will be updated with the system prior.

To update the prior formulation, a set of model properties, Z, with its distributional assumptions needs to be specified. Model properties $Z = h(\theta)$ are endowed with a complete probabilistic model $Z \sim D(Z^s)$, where D is a distribution function with hyper-parameters Z^s. With little loss in generality, we consider Z^s to describe the mean and variance of the prior properties distribution. The likelihood function for non-sample information $Z = h(\theta)$, given Z^s, is denoted $p_S(Z^s|\theta, h, \mathcal{M})$ – a system prior. The combination of the marginal priors with the priors on the system properties of the model results in the final, composite joint prior distribution $\tilde{p}(\theta|Z^s, h, \mathcal{M}) \propto p_S(Z^s|\theta, h, \mathcal{M}) \times p(\theta|\mathcal{M})$.

The posterior distribution is obtained by combining the composite prior information and available sample observations. The posterior of the model parameters is then proportional to the likelihood of the model, given the observed sample updated with the prior, conditioned on the extra model features information Z^s, that is

$$
\begin{aligned}
p(\theta|Y^o, Z^s, \mathcal{M}) \quad &\propto \quad L(Y^o|\theta, \mathcal{M}) \times p_S(Z^s|\theta, h, \mathcal{M}) \times p(\theta|\mathcal{M}) \quad &(3) \\
&\propto \quad L(Y^o|\theta, \mathcal{M}) \times \tilde{p}(\theta|Z^s, h, \mathcal{M}). \quad &(4)
\end{aligned}
$$

The posterior distribution in (3) can be analyzed by popular Monte Carlo integration techniques. The use of Monte Carlo integration techniques, namely, variants of a Random-Walk Metropolis algorithm, has become widespread. The analysis of a posterior distribution using the augmented prior does not differ much from the current practice, as described in An and Schorfheide (2007) or Canova (2007), for instance. All components of the prior distribution and the likelihood function can be evaluated at particular θ, drawn from a suitable proposal distribution.[12] A Random-Walk Metropolis algorithm, using the $N(\theta^{(i)} - \theta^{(i-1)}, c\Sigma)$ as a proposal distribution is a perfectly feasible option. Here Σ can be a derived from the Hessian of the posterior mode $\theta^\star$ obtained by maximizing the criterion function (3) and c, a scaling parameter. The evaluation of $p(.)$ and likelihood $L(.)$ is unchanged, with an additional step of evaluating the system prior $p_S(.)$. For most system priors, the model does not need to be re-solved, since the solution of the model is also required for the likelihood function evaluation as well and is thus already available.

Investigation of the composite prior distribution will usually require sampling techniques. Although an analytical expression for $p_S(Z^s|\theta, \mathcal{M})$ might exist in some cases, most often it will not and sampling from the distribution to obtain θ will not be possible. An application of Rejection Sampling or Importance Sampling algorithm may work well for smaller problems.[13] A Random-Walk Metropolis algorithm applied to (3) with the likelihood function 'switched-off' and set to a constant, is our preferred way of analysis, mostly due to its ease of implementation. The necessity of two sampling steps to evaluate posterior and prior distributions separately is the only minor drawback of using system priors.

Unlike the independent marginal prior, the system properties prior will invoke a complex interdependence among parameters. The prior on system properties ties several or all parameters together to some extent. These restrictions are global, as they hold for all admissible val-

---

[12] The interpretation of (3) as a penalized maximum-likelihood estimation is possible and encouraged.

[13] In complex problems importance sampling can be preferred due to its easy of parallelization and the need to investigate thoroughly the support of the composite prior distribution.

ues of θ. This is desirable, as it may sharpen the identification and dissect the parameter space into feasible regions, similar to the requirement of saddle-path stability of the model that is usually imposed. It is important to realize the limitations of marginal posterior and prior distribution plots for individual elements of θ, not showing the joint probability of parameters now in both cases. For proper understanding of the prior information imposed on the model, a prior predictive analysis and prior distribution of model system properties (impulse-response function, for instance) is of utmost importance. We are sure that nobody can understand from a sequence of Beta, Inverted-Gamma, or Normal distribution what these parameter priors imply for the model's dynamics.[14]

E. Identification Issues

The specification of system priors may affect the posterior distribution of even purely unidentified parameters. As is well known, the prior distribution regularizes the likelihood even for non-identified or weakly identified parameters. See e.g. Canova and Sala (2009) for a discussion of identification in DSGE models. Poirier (1998) also shows that even a posterior distribution of the unidentified parameter can be different from its marginal prior distribution if the parameter is not independent of others. System priors provide the prior parameter interdependence, as well as the requirement of Blanchard-Kahn stability does. Both restrictions affect global identification.[15]

System priors have implications both for local and global identification. They may be a powerful mode-selection device that is easy to understand thanks to its transparency and economic foundations. For instance, in the case of a bi-modal likelihood, the nature of the system prior on sacrifice ratio may navigate the estimation to a particular mode. Further, using system priors with multiple conditions restricts feasible parameterizations to a great extent due to the 'curse of dimensionality' – or a blessing of dimensionality, in our case. It may be possible to find many parameterizations satisfying sacrifice-ratio priors, but fewer also satisfying IRF duration priors or others.

Locally, the posterior-mode Hessian or Fischer Information Matrix are useful tools for understanding the criterion function. By virtue of the log-additive nature of the posterior

---

[14]Even the case of Normal priors is complex, since the mapping from the parameters to solution and model's properties is highly non-linear even for linear rational-expectations models

[15]For instance, parameters of the monetary policy rule can be locally weakly identified or not identified at all. The condition of Blanchard-Kahn stability, however, may provide only a subset of admissible combinations of the parameters, further sharpened by system priors.

functional (3), the Hessian evaluated at $\theta^\star$ can be decomposed as

$$H(\theta^\star) = H_Y(\theta^\star) + H_S(\theta^\star) + H_P(\theta^\star), \tag{5}$$

i.e. as a sum of components due to log-likelihood and composite prior – system properties priors and independent marginal priors. All components reveal a great deal of local information about parameter interdependence. The marginal prior component H_p is diagonal, regularizing the likelihood component H_Y and, akin to a ridge regression, guaranteeing the non-singularity of $H_Y + H_p$. The component due to system priors H_S will not be diagonal in general, revealing identification patterns among parameters consistent with the system properties priors.

Globally, system priors may contribute to mode-selection. System priors, e.g. the sacrifice ratio, are very nonlinear function of parameters, not identified by the likelihood. One interesting option is identifying information via the solution of the model featuring *anticipated events*, see e.g. Blanchard and Kahn (1980) for the forward expansion, or the intuitive discussion in Turnovsky (2000). Forward expansions are dependent also on unstable eigenvalues of the system, unlike the standard solution. Hence, a system prior based on an anticipated N-period delayed interest rate reaction to a demand shock may be a strong identifying criterion a priori, since there is no information on it in the data.[16]

III. ILLUSTRATION OF THE METHOD

A. Simple Monetary Policy Model

We set up a fairly standard closed economy monetary policy model to illustrate the workings of system priors. The model is simple and rather stylized but it is no straw man; variants of the model have been used for policy analysis and projections.

---

[16]Without forward expansion for anticipated events, standard recursive solutions of linear rational expectation models retain dynamics of stable eigenvalues and operate on their saddle path after an initial jump in reaction to new information. In the case of a pre-announced event, the solution up to the implementation phase is off the model's saddle path and depends also on unstable eigenvalues of the system. To illustrate the point using an extreme example, in the purely forward-looking model in Canova and Sala (2009) the interest-rate rule parameter, $a5$, drops from the solution altogether. It is not weakly identified, it simply is not there. However, it can be shown that it does not drop from the forward expansion for anticipated events. Thus, even in this extreme case, the scenario of delayed interest rate response to a demand shock –shock v_1 in Canova and Sala (2009)– results in different interest rate reactions depending on the parameter $a5$ of the policy rule.

The model follows a typical New-Keynesian closed economy model with price rigidities.
Inflation, π_t, is driven by output in excess of its trend or equilibrium value, using a forward-looking Phillips curve. The output cycle, $\hat{y}$, is determined by an output equation derived from consumption smoothing and is interest sensitive. The monetary policy authority sets the short-term nominal interest rate, i_t, via an inflation-forecast based rule, weighting the expected deviation of year-on-year inflation from its target and the output gap.

$$\hat{y}_t = \alpha_1 y_{t+1|t} + \alpha_2 y_{t-1} + \alpha_3 (rr_t - \overline{rr}_t) + \varepsilon_t^y \tag{6}$$

$$\pi_t^c = \lambda_1 \pi_{t+1|t}^c + (1-\lambda_1)\pi_{t-1}^c + \lambda_2 \hat{y}_t + \varepsilon_t^\pi \tag{7}$$

$$i_t = \gamma_1 i_{t-1} + (1-\gamma_1) \times \left[(\overline{rr}_t + \overline{\pi}_t) + \gamma_2 (\pi_{t+3|t}^{y/y} - \overline{\pi}_{t+3}) + \gamma_3 y_t \right] + \varepsilon_t^i \tag{8}$$

$$rr_t = i_t - \pi_{t+1|t} \tag{9}$$

$$\overline{\pi}_t = \rho_{\overline{\pi}}\overline{\pi}_{t-1} + \varepsilon_t^{\overline{\pi}} \tag{10}$$

Despite its small size and simplicity, the model can display nontrivial dynamics in response to structural shocks. It is driven by eight parameters $\theta = \{\alpha_{1,2,3}, \lambda_{1,2}, \rho_i, \gamma_{1,2,3}\}$ and four standard deviations for structural shocks. Inflation target is assumed to follow a random-walk, $\rho_\pi = 1$. There cannot be a smaller model than this one. Namely, there exist parameterizations featuring stark differences in the coefficient values that deliver very similar properties of the model, differing in just a few aspects, which may not always be easily identified from the observed data.

The model is semi-structural but can be derived from behavioral assumptions. A derivation of the model is carried out in the appendix, under the assumption of GHH preferences, habit formation, flexible wages, labor as a single input into production, and a one sector economy. The model based on explicit behavioral assumptions is a restricted version of the semi-structural model, embodying very stylized assumptions. We use, however, the semi-structural version of the model as it is satisfactory for illustrating the main point of the paper.

There are several key prior restrictions that even the semi-structural form of the model strictly adheres to. First, the Phillips curve is homogeneous and independent of the inflation target level. The inflation target is set by the monetary authority and it is the role of the policy rule to make sure it is achieved. Second, the sum of the coefficients on lead and lag in the output equation need not sum to unity, in contrast to the models' price dynamics. This can be easily seen from the derivation of the model in the Appendix.

B. Computational Experiment

To illustrate the idea of system priors, the example below demonstrates how a single system prior induces individual parameter priors. We believe that this type of exercise is very informative about the philosophy of system priors. The example uses the model with very diffuse marginal independent priors on individual parameters, which is updated to form a composite prior by a single *system prior.* An interesting aspect of the exercise is to see how and for which parameters the prior distribution gets changed. In principle, if the property of the model used as a system prior is affected by the parameter, the individual parameter's prior distribution will get influenced by introducing the system prior.

The effect of a system prior on an individual parameter's prior distribution depends on the global identification of the parameter with respect to the system prior. If the parameter is not affecting the statistic within its whole support, the parameter is not globally identified with respect to a system prior. Note that the mapping from parameters in θ to the rational-expectations solution is highly nonlinear and that there may be multiple parameterizations consistent with the desired value of the system prior. This, actually, is the key principle, otherwise the prior would collapse into calibration. This multiplicity may lead to the existence of *iso-parametric paths*.

In our example, the system prior used is specified in terms of a distribution of the sacrifice ratio. Strictly speaking, there are two system priors, since we truncate the whole parameter space to regions that are Blanchard-Kahn stable (a dogmatic system prior, in essence). The sacrifice ratio is defined as a cumulative loss in output after a permanent decrease of the central bank's inflation target by 1 percentage point. In our simple linear model, the disinflation from 3% to 2% will produce the same loss of output as disinflation from 15% to 14% would. In a more complex, possibly nonlinear models, this would not hold and the scope for employing the sacrifice-ratio prior would be even larger. The sacrifice ratio is a key statistic for a policy model's evaluation, see e.g. FED (1996) or Laxton and Pesenti (2003).

The system prior is specified such that the sacrifice ratio is assumed to be distributed as N(-0.8,0.05). This constitutes relatively informative but very transparent prior information about the model. Normal distribution is employed for simplicity only; one can opt for any distribution of interest. For each parameter, a rather diffuse marginal independent prior is chosen, relying on truncated Normal distribution. In particular, it is interesting to see how a relatively diffuse multivariate Normal prior distribution of all parameters gets translated to a prior distribution of the sacrifice ratio before the system priors are taken on board.

Figure 1 demonstrates how the prior about the sacrifice ratio induces priors across individual parameters. The blue dashed line is the marginal independent prior, the red solid line is the marginal prior distribution of the composite prior, after the system prior about the sacrifice ratio is applied. Some parameters are affected more then others, some seem essentially irrelevant for the size of the sacrifice ratio. The coefficients in the output and inflation equations are very much affected by the assumptions about the range of the system prior. It should also be understood that only marginal distributions are plotted, whereas for the estimation, the full joint prior distribution will be exploited. Interestingly it shows that the least identified coefficients in our example are the coefficients of the monetary policy rule.

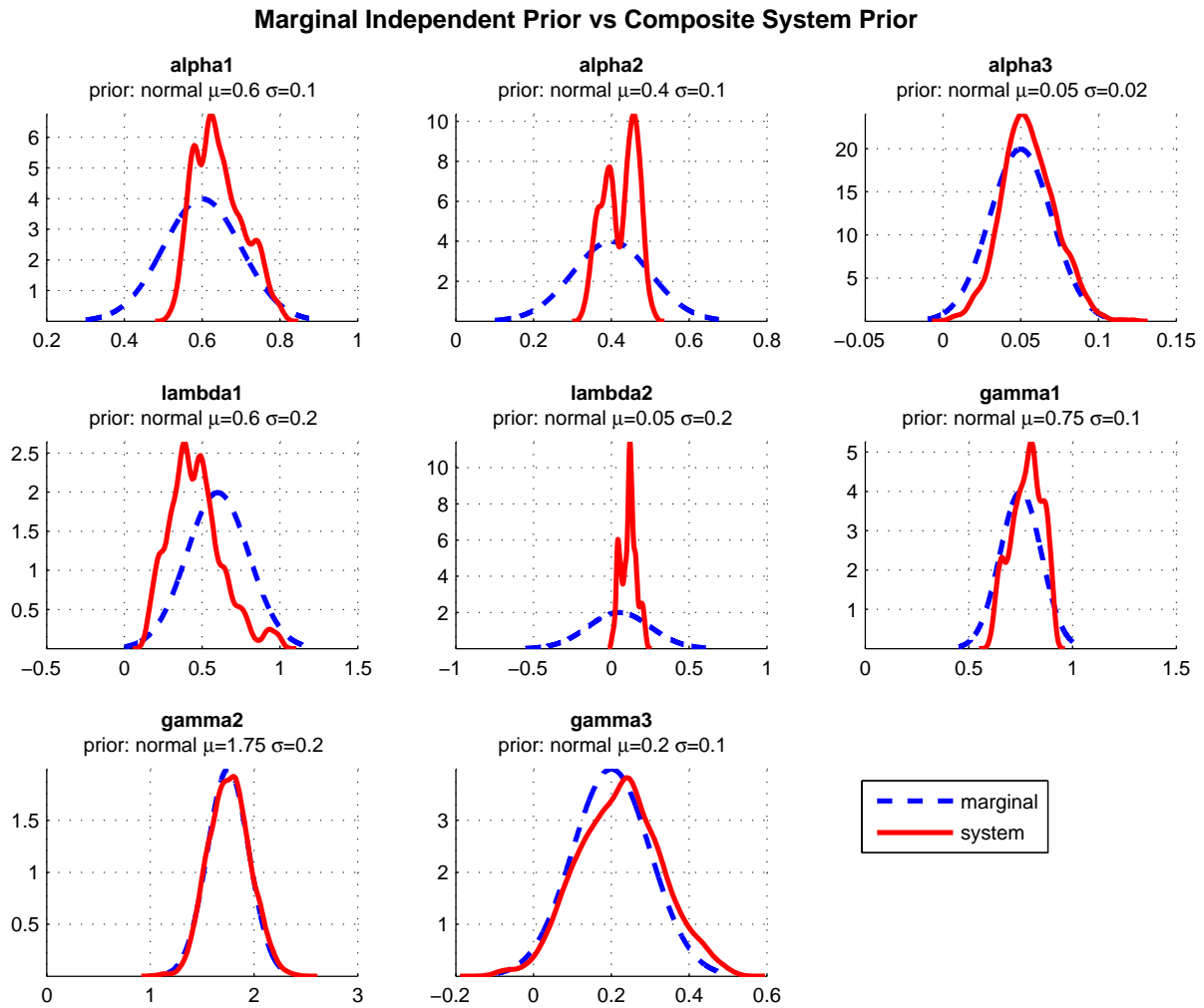**Figure 1. Priors induced by a sacrifice ratio prior**

Marginal Independent Prior vs Composite System Prior

Figure 2 illustrates two prior distributions of the sacrifice ratio – the marginal independent prior and the composite prior. The value of the sacrifice ratio at the mode (and by that token, at the mean) implied by the marginal independent prior is -0.5, indicated by a red vertical line on both panels. As can be seen from the left panel, the diffuse Gaussian priors result in highly non-Gaussian prior distribution for the sacrifice ratio. This is to be expected. The mapping from the structural parameters to the sacrifice ratio is highly nonlinear and must satisfy the saddle-path stability of the model. The right panel depicts the distribution of the sacrifice ratio, after the system prior of N(-0.8,0.05) has been implemented.[17]

Figure 2. Prior predictive distributions: sacrifice ratio
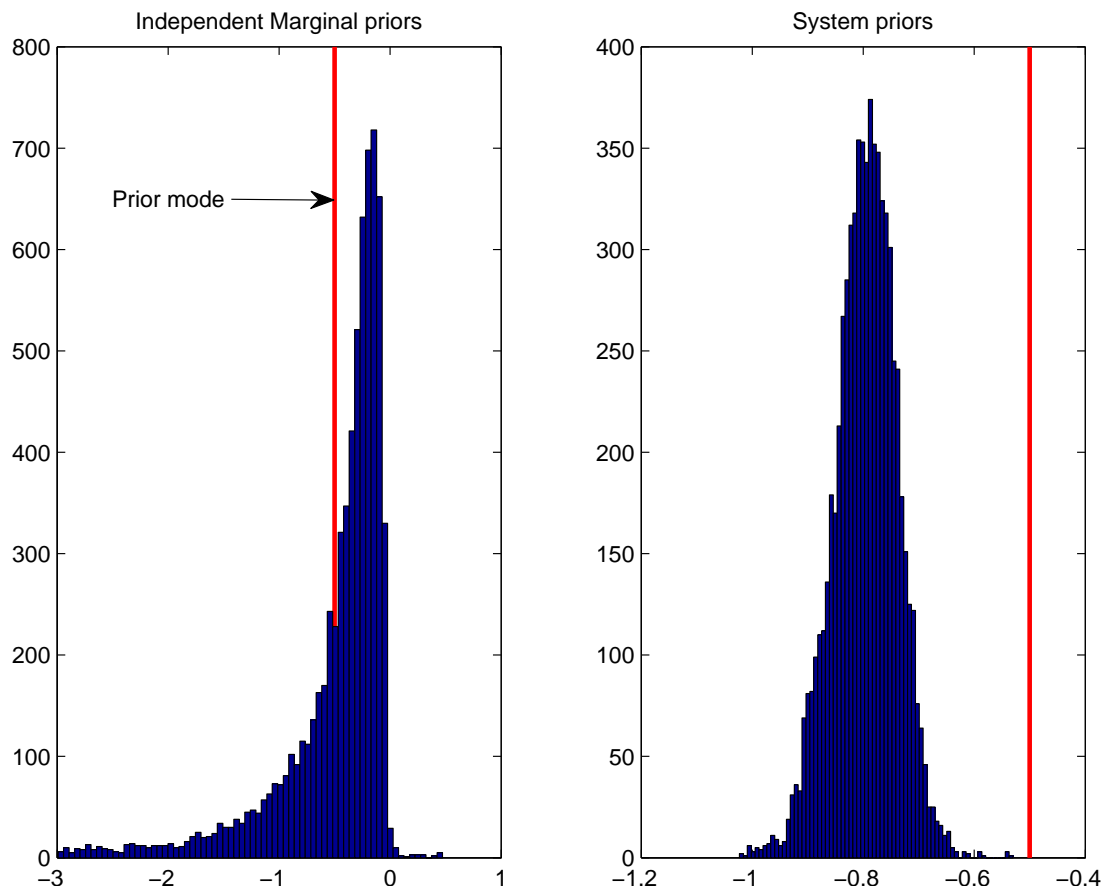
CONCLUSIONS

This paper demonstrated a novel approach how to formulate priors for economic models. System priors are priors directly related to the structural system properties of the model, that allow one to back out implied complex joint prior distribution for individual parameters. System priors allow for eliciting transparent and economically meaningful priors for the estimation of structural models, without unintentional consequences as is often the case with independent marginal parameter priors.

System priors –no matter what they are– are transparent and easy to communicate. With system priors, it is perfectly acceptable to have economic priors and it is likewise acceptable to have quite informative priors. System priors foster transparency and it is their transparency and economically meaningful nature that allow the priors to be discussed and debated. After all, these are subjective priors that need to be explained. Although several examples of system priors were discussed, these are just the tip of the iceberg. System priors free analysts from 'prior specification fishing,' chasing for the combination of marginal independent priors that roughly convey their more often vague ideas about model behavior, if they have any. On the other hand, system priors force researches to carefully articulate their priors.

System priors can serve both as a complement and substitute for individual independent marginal parameter priors. There is no doubt that having priors directly about structural parameters is the first best approach in many cases. System priors are not mutually exclusive with individual parameter priors in principle; better, system priors may impose further structure on those, as is clear from the computational procedure suggested.

References

An, S., and F. Schorfheide, 2007, "Bayesian analysis of DSGE Models," *Econometric Reviews*, Vol. 26, No. 2–4, pp. 113–172.

Andrle, M., Ch. Freedman, R. Garcia-Saltos, D. Hermawan, D. Laxton, and H. Munandar, 2009, "Adding Indonesia to the Global Projection Model," Working Paper 09/253, International Monetary Fund, Washington DC.

Blanchard, O., and Ch.M. Kahn, 1980, "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, Vol. 48, pp. 1305–1312.

Calvo, G.A., 1983, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, Vol. 12, pp. 383–398.

Canova, F., 2007, *Methods for Applied Macroeconomic Research* (Princeton, New Jersey: Princeton UP).

———, 2012, "Bridging DSGE Models and the Raw Data," Working Paper 635, Barcelona Graduate School of Economics, Barcelona, Spain.

Canova, F., and L. Sala, 2009, "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, Vol. 56, No. 4, pp. 431–449.

Coenen, G., A.T. Levin, and K. Christoffel, 2007, "Identifying the Influences of Nominal and Real Rigidities in Aggreagate Price-Setting Behavior," *Journal of Monetary Economics*, Vol. 54, pp. 2439–2466.

Del Negro, M., and F. Schorfheide, 2008, "Forming Priors for DSGE Models (and How it Affects the Assesment of Nominal Rigidities)," Working Paper w13741, National Bureau of Economic Research, Cambridge, MA.

Faust, J., 2009, "The New Macro Models: Washing Our Hands and Watching for Icebergs," *Economic Review (Sveriges Riksbank)*, Vol. 1, pp. 45–68.

Faust, J., and A. Gupta, 2012, "Posterior Predictive Analysis for Evaluating DSGE Models," Working Paper w17906, National Bureau of Economic Research, Cambridge, MA.

FED, 1996, "A Guide to FRB/US," Working Paper version 1.0, Federal Reserve Board, Board of Governors, Washington, D.C.

Fuhrer, J.C., 1994, "Optimal Monetary Policy and the Sacrifice Ratio," *Conference Series, Federal Reserve Bank of Boston*, Vol. 38, No. June, pp. 44–69.

Gali, J., 2008, *Monetary Policy, Inflation, and the Business Cycle* (Princeton,New Jersey: Princeton UP).

Geweke, J., 2007, "Comment," *Econometric Reviews*, Vol. 26, No. 2–4, pp. 193–200.

———, 2010, *Complete and Incomplete Econometric Models* (Princeton, NJ: Princeton University Press).

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

Jarociński, M., and A. Marcet, 2010, "Autoregressions in Small Samples, Priors about Observables and Initial Conditions," Working Paper 1263, European Central Bank, Frankfurt am Main, Germany.

Koopman, L.H., 1974, *The Spectral Analysis of Time Series* (San Diego, CA: Academic Press).

Laubach, T., 2001, "Measuring the NAIRU: Evidence from Seven Economies," *The Review of Economics and Statistics*, Vol. 83, pp. 218–231.

Laxton, D., and P. Pesenti, 2003, "Monetary rules for small, open, emerging economies," *Journal of Monetary Economics*, Vol. 50, No. 5, pp. 1109–1146.

Leamer, E.E., 1978, *Specification Searches: Ad Hoc Inference with Nonexperimental Data* (New York: John Wiley & Sons, Inc.).

Poirier, D.J., 1998, "Revising Beliefs in Nonidentified Models," *Econometric Theory*, Vol. 14, pp. 483–509.

Solow, R., 2010, "Building a Science of Economics for the Real World," Techn. rep., Prepared statement for House Committee on Science and Technology Subcommittee on Investigation and Oversight, July 20.

Stock, J.H., 1994, "Unit Roots, Structural Breaks and Trends," *Handbook of Econometrics*, Vol. 4, pp. 2739–2841.

Theil, H., and A. Goldberger, 1961, "On Pure and Mixed Statistical Estimation in Economics," *International Economic Review*, Vol. 2, No. 1, pp. 65–78.

Turnovsky, Stephen J., 2000, *Methods of Macroeconomic Dynamics – 2nd Edition* (Cambridge: MIT Press).

Villani, M., 2009, "Steady-state Priors for Vector Autoregressions," *Journal of Applied Econometrics*, Vol. 24, No. 4, pp. 630–650.

APPENDIX

IV. BEHAVIORAL FOUNDATIONS OF A SIMPLE MONETARY MODEL

The standard New Keynesian Model, see e.g. Gali (2008), amounts to a system similar to our simple model, with the restriction $\alpha_2 = 0$.

We assume a single sector economy where monopolistic intermediate firms produce goods using labor only, i.e.:

$$Y_t(j) = Z_t L_t(j)^\gamma, \tag{11}$$

where Z_t is aggregate technology shock. We assume that intermediate firms' prices are sticky, which can be easily operationalized using a version of Rotemberg or Calvo pricing. We assume Calvo (1983) pricing with the parameter θ determining the probability of changing the price in the current period. Firms that do not re-optimize prices in the current period follow a rule of thumb and index their prices by previous period change of the average prices level in the industry. These assumption lead to a Phillips Curve of the form:

$$\pi_t^y = \frac{\beta}{1+\beta} \pi_{t+1|t}^y + \frac{1}{1+\beta} \pi_{t-1}^y + \frac{(1-\theta)(1-\theta\beta)}{\theta(1+\beta)} r\hat{m}c_t + \varepsilon_t^\mu, \tag{12}$$

where $r\hat{m}c = \hat{w}_t - \hat{p}_t - \hat{z}_t$ is a log-deviation of real marginal costs from its steady-state, as determined by the inverse markup. The 'cost-push shock' ε_t^μ follows from a time variation in the elasticity of substitution among intermediate goods.

There is a mass of consumers who supply their labor services, L_t, and consume final goods, C_t, at a price P_t. All consumers can use a deposit account at a perfectly competitive bank with the rate of interest R_t^B. They also receive proportional share of intermediate firms profits, Π_t and transfers, T_t. The consumer's problem can be stated as:

$$\max \sum_{t=s}^{\infty} \beta^t \left((C_t(i) - hC_{t-1}) - \chi \frac{L_t^{1+\eta}}{1+\eta} \right)^{1+\sigma} \times \frac{1}{1+\sigma} \tag{13}$$

subject to a period budget constraint

$$P_t C_t(i) + B_t(i) = R_{t-1}^B B_{t-1}(i) + W_t L_t(i) + \Pi_t(i) + P_t T_t(i). \tag{14}$$

The bank is perfectly competitive, refinancing itself at a central bank. For simplicity we assume that transformation of credit and deposits is costless, but subject to stochastic shocks ϕ_t. The

market interest rate is thus equal to the policy rate, augmented for the banking premium, i.e.
$R_t^B = R_t \phi_t^B$.

We assume a cashless-limit economy, for simplicity. The monetary policy authority operates under flexible inflation targeting regime and sets the nominal policy rate according to an interest rate rule,

$$\widehat{R}_t = \rho \widehat{R}_{t-1} + (1-\rho)\left[\gamma_\pi(\pi_{t+3}^{y/y} - \bar{\pi}_{t+3}) + \gamma_y \hat{y}\right] + \varepsilon_t^i. \tag{15}$$

Here $\widehat{R}_t$ denotes annualized deviation of the nominal policy rate from its trend (equilibrium) value $\bar{R}S_t$.

The log-linearized model can be written down using the following set of equations:

$$\hat{y}_t = \frac{1}{1+h}y_{t+1|t} + \frac{h}{(1+h)(1+\iota_L)}\hat{y}_{t-1} - \frac{1-h}{1+h}\frac{\zeta_L}{\sigma}(\hat{R}_t - \pi_{t+1}) + \hat{\phi}_t^B \tag{16}$$

$$\pi_t^y = \frac{\beta}{1+\beta}\pi_{t+1}^y + \frac{1}{1+\beta}\pi_{t-1}^y + \left(\frac{(1-\theta)(1-\theta\beta)}{\theta(1+\beta)}\right)\frac{\eta}{\gamma}(\hat{y}_t - \hat{z}_t) + \varepsilon_t^\mu \tag{17}$$

$$\widehat{R}_t = \rho \widehat{R}_{t-1} + (1-\rho)\left[\gamma_\pi(\pi_{t+3}^{y/y} - \bar{\pi}_{t+3}) + \gamma_y \hat{y}\right] + \varepsilon_t^i. \tag{18}$$

where $\zeta_L \equiv (1 - \iota_L/[(1+\eta)(1-h)])$ and $\iota_L \equiv \chi L^\eta(L/C)/(1+\eta)$.

It is clear that the simple structural model is a rather restricted version of the semi-structural model estimated in the main body of the paper.