Working Paper

Christian Schoder[1]

Estimating Keynesian models of business fluctuations using Bayesian Maximum Likelihood*

January 23, 2016

Abstract

An empirical approach to model estimation and evaluation based on Bayesian Maximum Likelihood is introduced to the post-Keynesian literature. To illustrate the method, it is applied to a neo-Kaleckian type of model of Euro Area business cycle fluctuations including endogenous fiscal and monetary policy as well as endogenous wage formation. To evaluate its empirical performance, the marginal likelihood and impulse-responses conditional on the proposed model are contrasted to those conditional on the corresponding Bayesian vector auto-regression models after relaxing the theory-implied cross-coefficient restrictions. The estimated parameter distributions are broadly in line with the empirical literature. Yet, a Bayesian vector auto-regression with loose theory-implied restrictions on the prior outperforms the neo-Kaleckian model considerably indicating misspecification. Further, a baseline Dynamic Stochastic General Equilibrium model is superior in terms of the marginal likelihood. Comparative impulse-response analysis indicates a failure of the neo-Kaleckian model to satisfyingly capture the fiscal and monetary policy transmission mechanisms.

Keywords: Post-Keynesian economics, Bayesian Maximum Likelihood, Bayesian Vector Auto-Regression, model estimation, model evaluation

JEL Classification System: E12, E32

Estimating Keynesian models of business fluctuations using Bayesian Maximum Likelihood.[*]

Christian Schoder[†]

January 23, 2016

Abstract

An empirical approach to model estimation and evaluation based on Bayesian Maximum Likelihood is introduced to the post-Keynesian literature. To illustrate the method, it is applied to a neo-Kaleckian type of model of Euro Area business cycle fluctuations including endogenous fiscal and monetary policy as well as endogenous wage formation. To evaluate its empirical performance, the marginal likelihood and impulse-responses conditional on the proposed model are contrasted to those conditional on the corresponding Bayesian vector auto-regression models after relaxing the theory-implied cross-coefficient restrictions. The estimated parameter distributions are broadly in line with the empirical literature. Yet, a Bayesian vector auto-regression with loose theory-implied restrictions on the prior outperforms the neo-Kaleckian model considerably indicating misspecification. Further, a baseline Dynamic Stochastic General Equilibrium model is superior in terms of the marginal likelihood. Comparative impulse-response analysis indicates a failure of the neo-Kaleckian model to satisfyingly capture the fiscal and monetary policy transmission mechanisms.

Keywords: Post-Keynesian economics, Bayesian Maximum Likelihood, Bayesian Vector Auto-Regression, model estimation, model evaluation
JEL Classification: E12, E32

1

1 Introduction

In recent years, considerable advances have been made in estimating Dynamic Stochastic General Equilibrium (DSGE) models using Bayesian methods (cf. An and Schorfheide 2007).[1] An entire research agenda has been devoted to improve the empirical performance of this model class (cf. Smets and Wouters 2003, Smets and Wouters 2007, Christoffel et al. 2008, Ratto et al. 2009). The forecasting performance made remarkable progress compared to Vector Auto-Regression (VAR) models even though it is still low (cf. Edge and Gurkaynak 2010). More importantly, the continuous evaluation of the model's ability to explain the data enhanced the theory within the neo-classical paradigm.[2] This endeavor's ultimate aim is to provide policy makers at governments and central banks with a theoretically rigorous and empirically well-performing macroeconomic model which can be used for policy prescription and evaluation (cf. Christiano et al. 2010).

Rejecting the general-equilibrium conception of a market economy and emphasizing the disequilibrium nature of unemployment, the post-Keynesian (PK) cousins of the DSGE models span an attractive framework for macroeconomic analysis.[3] In terms of maturity as well as suitability for policy prescription and evaluation, however, PK models are considerably lagging behind DSGE models, even though the core of PK theory may be more sound than the core of neoclassical theory.

Why do policy authorities not use PK models for policy analysis? The reasons are certainly manifold and shall not be discussed here. Yet, one crucial part of the story may be that the vast majority of PK models proposed is primarily pedagogical. Their aim is to explain the transmission of a shock rather than assess its impact quantitatively.[4] Hence the proposed models are typically highly stylized and an insufficient effort is made to evaluate their overall empirical performance.

---

[1]For an overview and applied introduction to Bayesian Maximum Likelihood see Fernández-Villaverde (2009) and Guerron-Quintana and Nason (2012).

[2]Consumption habit formation, capital adjustment costs, variable capital utilization, price and wage rigidities, rule-of-thumb households, financial frictions, job search and matching frictions are prominent examples for market imperfections that have been introduced to the standard Real Business Cycle model to improve its empirical performance (cf. Christiano et al. 2010).

[3]Behavioral assumptions particularly regarding the formation of output, investment, prices and wages, differ between the Kaleckian, Kaldorian, Harrodian and Marxian tradition. Yet, all post-Keynesian (PK) schools of thought share the perception of unemployment being a disequilibrium phenomenon: There is no market mechanism whatsoever that adjusts the nominal wage such that the labor market clears, even in the absence of rigidities. Moreover, it is ambiguous whether the wage should rise or fall to reduce unemployment. Since the economy does not fully utilize its resources in the steady state, output is demand determined, an implication known as the principle of effective demand.

[4]Among many others, Taylor (1985), Blecker (1989), Skott (1989), Bhaduri and Marglin (1990), Shaikh (2009), Schoder (2014b) and Carvalho and Rezai (2014) propose model variants to describe the effects of a shock to income distribution. Lavoie (1995) and Hein (2007) are examples of models that provide a simple theory of the monetary policy transmission. Stockhammer (2005-06), Meirelles and Lima (2006), Lima and Meirelles (2007), Hein and Schoder (2011) and Skott and Ryoo (2008) propose models to study the effects of a phenomenon referred to as *financialization* on economic activity.

This is not to say that PK models do not lend themselves to empirical analysis. Applying a crude classification, one branch in the literature has attempted to make empirical statements by estimating single equations taken from or motivated by an underlying economic model (cf. Lavoie et al. 2004, Barbosa-Filho and Taylor 2006, Naastepad and Storm 2006-07, Stockhammer et al. 2009, Hein and Schoder 2011, Skott and Zipperer 2012, Papadimitriou et al. 2013).[5] Another branch has relied on estimating Structural Vector Auto-Regression (SVAR) models where merely the selection of variables and the scheme of identification comes from theory (cf. Stockhammer and Onaran 2004, Taylor et al. 2012, Schoder 2014a). To the author's best of knowledge, however, there has not been an attempt to estimate a complete PK model and use the obtained results for policy analysis.

The aim of the present paper is to advance the status quo of PK model estimation and evaluation. This is done by introducing the method of Bayesian Maximum Likelihood to the empirical PK literature. In particular we present a PK model of the business cycle which is simple but sufficiently rich to allow us to make empirical statements. It includes fiscal and monetary policy, and endogenous wage formation. For robustness checks, we also consider models which exclude these features by imposing zero-restriction on the respective parameters. The model specifications are then estimated employing quarterly data of the Euro Area from 1970 to 2014. The models are evaluated using Economic Theory Vector Auto-Regressions (ET-VARs) which have been proposed under the name of DSGE-VARs by Del Negro and Schorfheide (2004) and Del Negro et al. (2007). An ET-VAR(λ) is a Bayesian VAR approximation of the economic model where the hyperparameter λ controls the weight which is put on the cross-coefficient restrictions implied by economic theory. With $\lambda \to 0$, the VAR is unrestricted, with $\lambda \to \infty$ the VAR converges to the economic model. Model evaluation is then based on comparing the marginal likelihoods of the data as well as the impulse-responses to shocks for different λ, i.e. for VAR approximations of the economic model with different weights put on the theory-implied restrictions. Finally, we address the question which structural shocks drive the variation in the data in average and over time by conducting a decomposition of the forecast error variance.

We obtain the following results. First, the posterior means of the estimated parameters are of the expected size and the diagnostic tests indicate convergence of the posterior simulator. Surprisingly, the model fit to the data does not assign an important role to the labor market feedback on wage formation not lending much support for Goodwin (1967) wage dynamics. Since this mechanism is the main channel through which the supply side affects the demand side in our economy, we do not find productivity and wage bargaining power shocks to affect aggregate demand and its components much. Second, despite the robust estimation results and good diagnostic properties of the estimation, the PK model's

---

[5]As a side note, none of these contributions is entirely satisfying as there are good reasons to expect the estimation results of each of these papers to be biased. This is because the core assumption required for the Ordinary Least Squares (OLS) estimator to yield valid results, i.e. a zero covariance between the regressors and the disturbances also known as the exogeneity of the regressors, is likely not to be met in the specifications. The common practice of removing the contemporaneous value of the regressor does not eliminate the endogeneity problem if the model is dynamic or if the residuals exhibit serial correlation.

empirical performance is rather poor. The marginal likelihood of the data increases considerably if we loosen the theory-implied restrictions on the prior of a Bayesian VAR model approximating the PK model which indicates misspecification of the economic model proposed. Moreover, the benchmark Dynamic Stochastic General Equilibrium model proposed by Smets and Wouters (2003) performs considerably better. Third, comparing the impulse-response functions of Bayesian VARs which differ only by the tightness of the theory-implied cross-coefficient restrictions of the prior suggests that the misspecification of the estimated PK model outlined in this paper originates in the failure to replicate satisfyingly the cyclical response to shocks and the fiscal and monetary transmission mechanisms. Fourth, a change in income distribution triggered by a bargaining power shock does not unambiguously affect economic activity. Hence, there is no clear indication of wage-led or profit-led demand.

Estimating PK models by Bayesian Maximum Likelihood is appealing on many levels. First, perceiving economic models as asserted multivariate distributions for the data is a more natural and methodologically sound way of empirically evaluating economic theory than assembling single-equation OLS estimates (Haavelmo 1944). Second, the Bayesian approach to statistical inference which can be interpreted as a combination of calibration and estimation makes the influence of economic theory on econometric analyses very transparent. There is no need for data mining. Third, the Maximum Likelihood approach which relies on the Kalman filter allows us to estimate models even though not all variables are observed. Fourth, in the Bayesian framework a natural way of evaluating competing models is to compare how well they describe the data. Fifth, the distribution over impulse-response functions allows us to analyze the empirical macroeconomic effects of disturbances such as fiscal policy shocks, monetary policy shocks and wage policy shocks. Sixth, a forecast error variance decomposition allows us to study which shocks have been driving the business cycle in average and over time. Finally, from a practical perspective, Bayesian model estimation and evaluation can easily be conducted by using the software package *Dynare*.

The remainder of the paper proceeds as follows: Section 2 presents the simple PK stock-flow consistent model to be estimated for the Euro Area. Section 3 introduces the concepts of Bayesian Maximum Likelihood and ET-VARs, applies these methods to our PK model and discusses the estimation results, marginal likelihoods, impulse-response functions as well as variance decompositions. Section 4 concludes the paper.

2   A simple PK model

2.1   Preliminaries

This section presents a PK stock-flow consistent model of a neo-Kaleckian flavor. Note the following: First, for simplicity we consider a model for business cycle fluctuations and, therefore, assume a deterministic steady growth path in the long run which is assumed to be consistent with the firm's investment behavior when all targets are met, the sum of

productivity and population growth as well as the growth rate of government spending (cf. Skott 1989, Duménil and Lévy 1999, Foley and Michl 1999, Lavoie 1996, Shaikh 2009, Taylor 2012, Allain 2015). All growing variables are normalized by this trend. Hence, shocks will have level-effects on aggregate demand and its components but not growth effects. Kaleckians may want to slightly adapt the model to allow for growth effects by changing the investment specification below. Yet, this requires additionally to normalize all variables by the stochastic trend arising from the evolution of the capital stock.

Second, in contrast to the convention in the PK literature we will state the behavioral assumptions in non-linear form such that the linear model can be obtained by applying a first-order approximation around the steady state. This is because we estimate not only behavioral equations but the entire model including budget constraints which are non-linear by nature as they include the real interest rate. The linearized model will then be fed into the Bayesian Maximum Likelihood estimation procedure.

Third, all linearized behavioral relations can be interpreted as aggregated linear approximations to the first-order conditions of the corresponding inter-temporal optimization problems (cf. Schoder 2015b). The parameters of the aggregative model can be seen as convolutes of deep parameters as well as variables fixed at their steady state values. If one opposes the idea of micro-foundations based on inter-temporal optimization, the behavioral relations can also be interpreted as non-linear versions of the stylized relationships known from the conventional PK literature. Further, we will assume adaptive expectation formation (which is not inconsistent with inter-temporal optimization as argued by Schoder (2015b)).

2.2 Economic setting

The supply side of the goods market is characterized by a fixed-coefficient production technology according to which non-depreciating and non-traded capital and labor are combined in order to produce a homogeneous output good used for both consumption and capital investment. Labor has to be compensated by a nominal wage per working hour which is assumed to be a collective policy variable and the same for anyone providing labor. Accounting implies that aggregate sales equal aggregate expenditures as there are no inventories. *Keynesian* unemployment prevails, i.e. there is always an abundance of underutilized labor.

We follow Hein (2007) and Hein and Stockhammer (2010) by assuming the economy to be populated by rentier households, worker households and firms. Workers supply labor and consume all of their income. Rentiers save out of their profit income and do not provide labor services. The supply-side of the external finance market is determined by the rentier's saving. The demand-side is determined by the firm's investment decision and dividend decision the latter of which pins down retained earnings and required bank loans. Firm's follow a Kaleckian (1971) type of price setting and Steindlian (1952) type of investment rule adjusted for short-run analysis and including the interest rate. Further, there is a policy authority setting the interest rate as well as government consumption and taxes. The rate of nominal wage inflation is determined by a bargaining process between workers and firms.

2.3 Model equations

Rentier and worker households. We assume a type of Cobb-Douglas consumption function which is related to a dynamic, restricted and non-linear version of the Modigliani (1986) consumption function, which, in its traditional form, is a popular choice in the stock-flow consistent literature (cf. Godley and Lavoie 2012). The representative rentier's real consumption $C_{r,t}$ depends on real profit income net of taxes $(1 - t_z)Z_{r,t}$ with t_z denoting a proportional tax rate on distributed profits and steady-state real wealth B_r.[6] In particular,

$$\frac{C_{r,t}}{B_r} = \left((1 - t_z)\frac{Z_{r,t}}{B_r}\right)^{\phi_c} V_{c,t} \tag{1}$$

where ϕ_c and $1 - \phi_c$ are the elasticities to consume out of income and wealth which are generally larger than the respective marginal propensities to consume as both $Z_r/C_r > 1$ and $Z_r/B_r > 1$ in the neighborhood of the steady state. The assumption that the elasticities sum up to unity is made to exclude scale effects on consumption. $V_{c,t}$ is a first-order autoregressive shock process to be specified below. Its mean will be restricted such that output is normalized to unity at the steady state. Note that we use the end-of-period notation for stocks. Real wealth evolves according to

$$B_{r,t} = \frac{R_t}{\Pi_{p,t+1}}\frac{1}{\Gamma}\left(B_{r,t-1} + (1 - t_z)Z_{r,t} - (1 + t_c)C_{r,t}\right) \tag{2}$$

where R_t, $\Pi_{p,t}$, Γ and t_c denote the gross interest rate, the gross price inflation rate, the gross rate of economic growth and the consumption tax rate, respectively. Note that growth is induced by the firm's investment decision discussed below.

The representative worker's real consumption expenditures $(1 + t_c)C_{w,t}$ are simply their wage income net of taxes

$$(1 + t_c)C_{w,t} = (1 - t_w)\omega_t L_t \tag{3}$$

where ω_t, L_t and t_w are the real wage, the employed labor hours and the wage income tax rate, respectively. The supply of labor hours $N_{w,t}$ is assumed to evolve according to

$$\frac{N_{w,t}}{N_w} = V_{n,t} \tag{4}$$

N_w is the steady-state number of hours supplied and $V_{n,t}$ is a first-order auto-regressive shock process. N_w will be calibrated such that the steady-state unemployment rate matches the average in the data.

---

[6]Mind the assumption of steady-state wealth instead of contemporaneous wealth affecting investment. It is made because, given the simplicity of the remainder of the model, a contemporaneous wealth effect would imply a one-time shock to the interest rate to be persistently expansionary due to the level increase in wealth. A similar assumption will be made for the effect of the capital stock on investment. Future extensions of the model could consider contemporaneous wealth and capital effects on consumption and investment, respectively.

Firms. The fixed-coefficient production function combined with the assumption of abundant labor implies a proportional relationship between labor demand L_t and output Y_t, where the proportionality factor is subject to a first-order auto-regressive productivity process $V_{a,t}$. If the fixed-coefficient production function is interpreted as a Cobb-Douglas production function with a *constant* steady-state output-capital ratio ψ, we get

$$L_t = \frac{Y_t}{V_{a,t}} \psi^{\frac{\alpha}{1-\alpha}} \tag{5}$$

where $\psi = Y/K$ and α is the capital elasticity of output.[7] Note that PK models typically assume a full utilization of labor, also over the cycle. Following this practice, we abstract from labor hording which will be soaked up by $V_{a,t}$ in our empirical analysis. With a constant steady-state capital-output ratio, the marginal costs φ_t are equal to the average costs which depend on productivity and the real wage. If we assumed the firm to maximize expected inter-temporal profits subject to a Cobb-Douglas production function with a fixed output-capital ratio, its labor choice would imply

$$\varphi_t = \frac{\omega_t}{V_{a,t}} \frac{1}{1-\alpha} \left(\frac{\psi}{V_{a,t}} \right)^{\frac{\alpha}{1-\alpha}} \tag{6}$$

which, from a Kaleckian perspective, can also be interpreted as a proportional relation between average costs and wage costs (cf. Schoder 2015b). In this view, prices would then be set with a mark-up on these average costs. In the Kaleckian (1971) theory of distribution, the mark-up fully determines the real wage and, for a given productivity, the wage share. Here we assume that adjusting prices is costly such that the mark-up decreases with an increase in inflation. An implication of this is that the bargaining over nominal wage inflation will have an impact on the real wage as firms do not simply pass on higher costs to higher prices. Interpreting the Kaleckian *degree of monopoly* as the inverse of the elasticity ϵ by which purchasers can substitute between the products of the different firms, we get

$$(1 - \varphi_t) \frac{1}{\epsilon} - \varphi_t = \tau \left(\Pi_{p,t}^e - \Pi_{p,t} \right) \tag{7}$$

where τ, $\Pi_{p,t}^e$ and $\Pi_{p,t}$ are the scaling parameter of quadratic price adjustment costs, the $t+1$-gross inflation rate as expected in t, and the period t gross rate of price inflation, respectively.[8]

Real investment I_t is assumed to respond to output and the interest rate. In particular, we specify

$$\frac{I_t}{K} = (\Gamma - (1-\delta)) \left(\frac{Y_t^e}{Y} \right)^{\phi_{iy}} \left(\frac{R_t}{R} \right)^{-\phi_{ir}} V_{i,t} \tag{8}$$

---

[7]In a neo-classical use of the production function, α determines the capital share in income. This is not the case in our PK model as we do not apply the marginalist theory of distribution.

[8]Note that, in the absence of price adjustment costs, i.e. when $\tau = 0$, (7) collapses to the standard Kaleckian price equation $1 = (1 + \epsilon)\omega_t/a_t$ with $a_t = V_{a,t}(1-\alpha)(1/V_{a,t}\psi)^{-\alpha/(1-\alpha)}$

where ϕ_{iy} is the capacity utilization elasticity of investment, and ϕ_{ir} is the interest elasticity of investment. $\Gamma - (1 - \delta)$ is the secular rate of gross capital accumulation with δ denoting the rate of capital depreciation. $(\Gamma - (1 - \delta))V_{i,t}$ may be interpreted as *animal spirits*. Since we assume capacity output to be constant, the ratio between output and steady-state output is equal to the ratio between capacity utilization and the normal rate which is equivalent to (Steindl 1952). The interest rate term is included in the investment function for two reasons: First, Schoder (2015b) has argued that the target utilization rate should decrease with the interest rate. Instead of endogenizing the target utilization rate, we add the interest rate to the investment function for the sake of simplicity, which is equivalent up to the first order. Second, without a negative interest elasticity of investment, monetary policy would not have any economic effects.[9]

Note that we normalize investment by the steady-state capital stock rather than the actual capital stock evolving according to

$$\Gamma K_t = I_t + (1 - \delta) K_{t-1}. \tag{9}$$

This is because otherwise we would introduce a unit root to investment and all other demand components. A one-time shock to investment, for instance, would have a level effect on the capital stock and hence on investment. From a Kaleckian perspective, this may be a desirable property of the model but would require us to transform the model equations in order to make the model variables stationary. Since we are interested in explaining the short-run fluctuations of the business cycle and filter the data accordingly, we stick to our model. Note that temporary shocks still have level effects due to the deterministic growth rate assumed which all variables are normalized by.

Apart from the price and investment decisions, the firm chooses the share of investment to be financed externally, ξ. We assume the firm to maintain a constant debt-capital ratio.[10]

Policy. Turning to policy, we assume total tax revenues to comprise wage taxes, distributed profit taxes and consumption taxes,

$$T_t = t_w \omega_t L_t + t_z Z_{r,t} + t_c(C_{r,t} + C_{w,t}). \tag{10}$$

---

[9]Note that a one-time rise in the interest rate will have a highly persistent expansionary effect on private wealth. If consumption were affected by actual wealth rather than steady-state wealth, a rise in the interest rate would be expansionary.

[10]We have also considered a Minskian (1976, 1985) extension of the model by assuming that firms increase leverage when output is rising and reduce leverage when the interest rate is below normal, i.e.

$$\frac{\xi_t}{\xi} = \left(\frac{Y_t}{Y_{t-1}}\right)^{\phi_{\xi y}} \left(\frac{R_t}{R}\right)^{\phi_{\xi r}}$$

where $\phi_{\xi y}$ and $\phi_{\xi r}$ are the elasticities of external finance with respect to output and the interest rate. Yet, the estimation procedure has attributed the wrong signs to the elasticities. Hence, we leave the Minskian mechanism out in the current model.

t_w will be calibrated such that, for given t_z and t_c, the structural budget is balanced. The deficit responds in the short run to the output gap, which implies

$$\frac{G_t}{G} = \left(\frac{Y_t}{Y}\right)^{\phi_{gy}} V_{g,t} \tag{11}$$

where ϕ_{gy} is the output gap elasticity of government spending and $V_{g,t}$ is a first-order auto-regressive shock process. Government debt $D_{g,t}$ evolves according to

$$D_{g,t} = \frac{R_t}{\Pi_{p,t+1}} \frac{1}{\Gamma} \left(D_{g,t-1} + G_t - T_t\right) \tag{12}$$

Monetary policy follows a Taylor rule by setting the interest rate in response to the deviations of inflation and output from the steady state. In particular,

$$\frac{R_t}{R} = \left(\frac{\Pi_{p,t}}{\Pi_p}\right)^{\phi_{r\pi}} \left(\frac{Y_t}{Y}\right)^{\phi_{ry}} V_{r,t} \tag{13}$$

where $V_{r,t}$ is a first-order auto-regressive shock process.

The gross rate of nominal wage inflation $\Pi_{w,t}$ is the outcome of an institutionalized bargaining process between the worker and the firm which seek to maximize the long-run real wage and the long-run profit rate, respectively. As shown by Schoder (2015b) the solution of the bargaining game is

$$1 = \left(1 - \frac{1}{\nu_t}\right) \frac{\omega(\Pi_{w,t})}{r(\Pi_{w,t})} \frac{r'(\Pi_{w,t})}{\omega'(\Pi_{w,t})} \tag{14}$$

where $\omega(\cdot)$ and $r(\cdot)$ are the real wage and the profit rate evaluated at the steady state and where ν_t is the worker's bargaining power. The solution to the bargaining problem implies a positive relationship between the rate of wage inflation and the workers' bargaining power. We assume ν to depend on employment, i.e.

$$\frac{\nu_t}{\nu} = \left(\frac{L_t}{L}\right)^{\phi_{\nu l}} V_{\nu,t} \tag{15}$$

where $\phi_{\nu l}$ is the employment elasticity of the worker's bargaining power and $V_{\nu,t}$ is a first-order auto-regressive shock process. The main property of this specification is that the wage inflation rate moves pro-cyclically if $\phi_{\nu l} > 0$. Since the firms do not pass on higher wage costs to higher prices completely, the real wage also moves pro-cyclically with a lag. This is highly consistent with a Goodwin (1967) reserve army of labor specification of the real wage.

Macroeconomic balance and definitions. Summing over the household's budget constraint, the firm's profit equation and the governments budget constraints, leads to the macroeconomic balance condition

$$Y_t = C_{r,t} + C_{w,t} + I_t + G_t + \frac{\tau}{2} \left(\Pi_{p,t} - \Pi_p\right)^2. \tag{16}$$

9

The rentier's real income is

$$Z_{r,t} = Y_t - \omega_t L_t - (1 - \xi_t) I_t - \frac{\tau}{2} (\Pi_{p,t} - \Pi_p)^2. \tag{17}$$

We can define the rate of unemployment as

$$u_t = 1 - \frac{L_t}{N_{w,t}}. \tag{18}$$

The growth of the real wage is connected to the rates of price and wage inflation as

$$\frac{\omega_t}{\omega_{t-1}} - 1 = \Pi_{w,t} - \Pi_{p,t}. \tag{19}$$

Expectations. The expectations of prices and output evolve adaptively as

$$\frac{\Pi_{p,t}^e}{\Pi_{p,t}} = \left(\frac{\Pi_{p,t-1}^e}{\Pi_{p,t}} \right)^{\rho_{\pi e}} \tag{20}$$

and

$$\frac{Y_t^e}{Y_t} = \left(\frac{Y_{t-1}^e}{Y_t} \right)^{\rho_{ye}}, \tag{21}$$

respectively. Note that this type of backward-looking expectation formation may reinforce macroeconomic shocks when a change in expectations causes, *ceteris paribus*, the realization to move in the same direction.

Shock processes. The first-order auto-regressive shock processes are

$$V_{g,t} = (V_{g,t-1})^{\rho_g} \exp(\sigma_g \varepsilon_{g,t}), \tag{22}$$

$$V_{a,t} = (V_{a,t-1})^{\rho_a} \exp(\sigma_a \varepsilon_{a,t}), \tag{23}$$

$$V_{r,t} = (V_{r,t-1})^{\rho_r} \exp(\sigma_r \varepsilon_{r,t}), \tag{24}$$

$$\frac{V_{c,t}}{V_c} = \left(\frac{V_{c,t-1}}{V_c} \right)^{\rho_c} \exp(\sigma_c \varepsilon_{c,t}), \tag{25}$$

$$V_{n,t} = (V_{n,t-1})^{\rho_n} \exp(\sigma_n \varepsilon_{n,t}), \tag{26}$$

$$V_{i,t} = (V_{i,t-1})^{\rho_i} \exp(\sigma_i \varepsilon_{i,t}), \tag{27}$$

$$V_{\nu,t} = (V_{\nu,t-1})^{\rho_\nu} \exp(\sigma_\nu \varepsilon_{\nu,t}), \tag{28}$$

where $\varepsilon_{x,t} \sim n.i.d.(0,1)$ is a disturbance with scale σ_x on variable x.

Measurement equations. Finally, we add seven measurement equations to our model which define the variables corresponding to our data set:

$$d_{y,t} = 100 \ln \left(\frac{Y_t}{Y} \right) \tag{29}$$

$$d_{c,t} = 100 \ln \left(\frac{C_{r,t} + C_{w,t}}{C_r + C_w} \right) \tag{30}$$

$$d_{i,t} = 100 \ln \left(\frac{I_t}{I} \right) \tag{31}$$

$$d_{r,t} = 100(R_t - R) \tag{32}$$

$$d_{\pi p,t} = 100(\Pi_{p,t} - \Pi_p) \tag{33}$$

$$d_{\pi w,t} = 100(\Pi_{w,t} - \Pi_w) \tag{34}$$

$$d_{u,t} = 100(u_t - u) \tag{35}$$

Despite the simplicity of the model, it could be used to study a range of policy issues. What is the multiplier effect of a fiscal expansion? What is the effect of a monetary policy shock? What is the effect of a strengthening of the workers' bargaining power causing the real wage to go up? Obviously the answer to these questions depends crucially on the parameterization of the model. Hence, we want to estimate them in order to address these questions for the Euro Area.

The proposed model is very simplistic abstracting from many aspects of the economy which will be relevant for explaining the data such as the foreign sector, public infrastructure investment, and capital gains. We leave it to future research to incorporate these features to post-Keynesian models to improve their empirical performance. The current model should be seen as the starting point for this line of research.

Consistent with Haavelmo's (1944) philosophy of econometrics, our economic model can be interpreted as a probability model that asserts a joint probability distribution of the observed variables for a given set of parameters and a given joint probability distribution of the innovations to the system. Confronting this probability model with the observed data allows us to infer on the joint and marginal distributions of unobserved parameters and variables, which is the essence of Bayesian Maximum Likelihood discussed in the next section. Before we can bring our model to the data, however, a few steps of simplification are necessary. So far, our economic model is stated as a *dynamic, backward-looking, non-linear* system of equations. The next sub-section briefly explains how to obtain a linear approximation of the non-linear model around its steady state. Afterwards, the linearized model is rearranged into the so-called state-space form which is a convenient representation for computing joint probability distributions. Then, based on the sample of observed macroeconomic data for the Euro Area, we can use the *dynamic, backward-looking, linear* probability model to make inference on the marginal probability distributions of the parameters of interest as well as on the probability distribution of observing the data we do.

2.4 Steady state and linearization

Since our estimation procedure requires a linear model, it has to be linearized. This is a strong simplification and future research should consider estimation techniques that can handle non-linear models along the lines of Fernández-Villaverde et al. (2010) and Fernández-Villaverde et al. (2011). Nevertheless, bear in mind that in most post-Keynesian models core equations are stated in linear form in the first place which is equally simplistic.[11] If we need to linearize, it will be done best around the steady state.

The steady state is the solution of the static model, i.e. when no transition is left and hence $x_t = x_{t-1}$ for any variable x with steady state value x^{ss}. Evaluated at the steady state our model consists of 35 equations including the same number of static endogenous variables.

The steady state of our model can easily be solved for analytically. The recursive computation of the steady state for each variable is reported in Appendix A. To linearize our model, we apply a log-linear approximation to the equations around the steady sate. The full set of linearized model equations is derived in Appendix B. Note that the linearized model looks very reminiscent of a standard PK model.

2.5 The linearized model in state space form

To express the model in more compact form, let us first collect all estimated as well as all calibrated parameters in vectors. We define the parameter vectors as

$$\Theta_e \equiv \begin{bmatrix} \tau & \Theta'_{e\phi} & \Theta'_{e\rho} & \Theta'_{e\sigma} \end{bmatrix}'$$

with

$$\Theta_{e\phi} \equiv \begin{bmatrix} \phi_c & \phi_{iy} & \phi_{ir} & \phi_{gy} & \phi_{r\pi} & \phi_{ry} & \phi_{\nu l} \end{bmatrix}'$$

$$\Theta_{e\rho} \equiv \begin{bmatrix} \rho_g & \rho_a & \rho_r & \rho_c & \rho_n & \rho_i & \rho_\nu & \rho_{\pi e} & \rho_{ye} \end{bmatrix}'$$

$$\Theta_{e\sigma} \equiv \begin{bmatrix} \sigma_g & \sigma_a & \sigma_r & \sigma_c & \sigma_n & \sigma_i & \sigma_\nu \end{bmatrix}'$$

where τ and the ϕ's are structural parameters affecting both steady state and propagation of the model, the ρ's are structural parameters capturing only the persistence of shocks, and the σ's are the standard deviations of the innovations, as well as

$$\Theta_c \equiv \begin{bmatrix} \Theta'_{c1} & \Theta'_{e2} \end{bmatrix}'$$

including the calibrated parameters with

$$\Theta_{c1} \equiv \begin{bmatrix} \Gamma & \psi & \alpha & \epsilon & \delta & \xi & t_c & t_z & t_w \end{bmatrix}'$$

$$\Theta_{c2} \equiv \begin{bmatrix} Y & C_r & C_w & I & G & R & \Pi_w & \Pi_p & N_w & K & \nu & D_g & B_r & u & L \end{bmatrix}'$$

---

[11]Abstracting away from non-linear model dynamics is a shortcoming that is common to most empirical analyses of economic models, be it Bayesian estimation of DSGE models, single-equation estimation of SFC models or estimation of SVAR models with theory-driven identification schemes.

where $\boldsymbol{\Theta}_{c1}$ involves structural parameters and $\boldsymbol{\Theta}_{c2}$ parameters which additionally turn out to equal the steady states of the respective variables. The observant reader may have noticed that V_c is neither part of $\boldsymbol{\Theta}_e$ nor $\boldsymbol{\Theta}_c$. This is because in the estimation procedure V_c will be perceived as a (constant) variable. Why is this necessary? In the economic model V_c is obviously a parameter. Yet, we will restrict it such that, without loss of generality, $Y = 1$ at the steady state. The functional form of this restriction is implied by the model above evaluated at the steady state and involves only parameters.[12] The problem is that under the restriction V_c is a function of an estimated parameter, i.e. ϕ_c. The estimation procedure will draw different values for ϕ_c to simulate the posterior distributions. The restriction then requires that V_c adjusts accordingly. Hence, V_c will be interpreted as a variable, and not be calibrated or estimated.

Further calibration choices and parameter restrictions involve the following: $\Gamma = 1.0042$ is roughly the sample quarterly growth rate of real GDP, $\psi = 0.1$ is close to the quarterly output-capital ratio of industrialized countries (cf. OECD Economic Outlook database). We set $\alpha = 0.25$, $\epsilon = 0.35$ and $\delta = 0.025$ which are close to the values for the capital elasticity of output, the price mark-up and the capital depreciation rate as typically assumed in the empirical literature (cf. Martins et al. 1996, and Jean and Nicoletti 2002, Smets and Wouters 2003, Christoffel et al. 2008, Ratto et al. (2009) and Schoder 2015a). We choose not to estimate them as they are only weakly identified. The share of external finance is calibrated as $\xi = 0.2$. We set $t_z = 0.15$ and $t_c = 0.15$ and restrict $t_w = 0.156$ in order for the structural budget to be balanced, i.e. (12) to hold at the steady state for $D_g = 2.4$ which corresponds to an annual debt-to-GDP ratio of 60%.

The steady state parameters are calibrated as follows: We normalize $Y = 1$ without loss of generality (and restrict V_c such that $Y^{ss} = Y$). $G = 0.2$ is roughly the sample share of government consumption in total demand, $R = 1.007$ is the inputed sample average nominal interest rate assuming the average quarterly inflation rate has been $\Pi_p = 1.005$ which is close to the ECB's inflation target. $N_w = 0.511$ is restricted such that $u^{ss} = u = 0.091$ which is the average unemployment rate in the sample. Note that we can simply calibrate N_w accordingly as the restriction does not involve estimated parameters. ν is restricted such that the steady-state bargaining solution $\Pi_w^{ss} = \Pi_w = \Pi_p = 1.005$ using (14) evaluated at the steady state. Finally, we have $K = Y/\psi = 10$. The remaining parameters C_r, C_w, I, B_r and L are calibrated at the corresponding variables' steady state values. Note that none of these parameters depend on estimated parameters.

It will be convenient to represent the linearized model in state space form as

$$\mathbf{S}_t = \mathbf{F}\mathbf{S}_{t-1} + \mathbf{Q}\varepsilon_t \quad \varepsilon_t \sim n.i.d.(\mathbf{0}_n, \mathbf{I}_n) \tag{36}$$

$$\mathbf{M}_t = \mathbf{H}\mathbf{S}_t \tag{37}$$

---

[12]In particular, $V_c = \frac{Y^{ss} - C_w^{ss} - I^{ss} - G^{ss}}{((1-t_z)Z_r^{ss})^{\phi_c}(B_r)^{1-\phi_c}}$ where, as argued above, the steady states of the variables are expressions in only parameters.

where $\mathbf{S}_t$ is the vector of unobserved or state variables,

$$\mathbf{M}_t = \begin{bmatrix} d_{y,t} & d_{c,t} & d_{i,t} & d_{r,t} & d_{\pi_p,t} & d_{\pi_w,t} & d_{u,t} \end{bmatrix}'$$

is the vector of observed variables,

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{g,t} & \varepsilon_{a,t} & \varepsilon_{r,t} & \varepsilon_{c,t} & \varepsilon_{n,t} & \varepsilon_{i,t} & \varepsilon_{\nu,t} \end{bmatrix}' \sim n.i.d.(\mathbf{0}_n, \mathbf{I}_n)$$

is a vector of innovations with n being the number of innovations, and $\mathbf{F} \equiv \mathbf{F}(\mathbf{\Theta})$, $\mathbf{Q} \equiv \mathbf{Q}(\mathbf{\Theta})$, and $\mathbf{H} \equiv \mathbf{H}(\mathbf{\Theta})$ are parameter matrices with their elements being functions of the model parameters. Eq. (36) is referred to as state equation as it controls the evolution of the unobserved states over time. Eq. (37) is called measurement equation as it relates the observed variables to the unobserved states.[13]

3 Empirical model evaluation

So far, we have set up the non-linear economic model, substituted out any forward-looking elements by assuming adaptive expectation formation, linearized the model around the steady state and squeezed it into state-space form (36) and (37). Simple theorems of probability allow us to infer from the observed data and our state of prior beliefs on various probability distributions: For instance, we might be interested in the distribution of the income elasticity of consumption, ϕ_c, conditional on only the observed data (and the economic model) but independent of the realization of any other parameter such as the sales elasticity of investment, $\phi_{i,y}$. Moreover, we might want to compute the probability that a specific economic model generates the data which we observe independent of any specific realization of a parameter, which is referred to as the marginal likelihood of the data or the unconditional probability of the data. As explained below, the method of Bayesian Maximum Likelihood allows us to simulate these distributions.

3.1 Bayesian methods of model estimation and evaluation

Preliminaries. We seek to estimate the probability density functions $\mathcal{P}(\theta_{i,e}|\mathcal{M}_T, \mathbf{\Theta}_c)$ for all elements $\theta_{i,e}$ of $\mathbf{\Theta}_e$ conditional on only the observed data $\mathcal{M}_T \equiv \{\mathbf{M}_0, \dots, \mathbf{M}_T\}$ and the

---

[13]Note that representing a linear model in state space form can be a tricky task and may require redefining a variable x_{t-1} as a variable y_t in order to stay within the first-order auto-regressive structure. The state vector may also include 1 as a variable to allow variables to be constant over time. Note further that our model involves the same number of shocks as observed variables which is a requirement for the ET-VAR(λ) approximation of the economic model which we will be using below (cf. Del Negro and Schorfheide 2004, Del Negro et al. 2007). In contrast, Bayesian Maximum Likelihood estimation of the economic model requires the number of observed variables to be not greater than the number of shocks (cf. An and Schorfheide 2007).

vector of calibrated parameters $\mathbf{\Theta}_c$.[14] Basic principles of probability theory imply

$$\mathcal{P}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c) = \frac{\mathcal{L}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)\mathcal{P}(\mathbf{\Theta}_e|\mathbf{\Theta}_c)}{\mathcal{P}(\mathcal{M}_T|\mathbf{\Theta}_c)}$$
$$\propto \mathcal{L}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)\mathcal{P}(\mathbf{\Theta}_e|\mathbf{\Theta}_c)$$

which is also known as Bayes' rule. The joint posterior distribution of the parameters $\mathcal{P}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)$ conditional on the data and the calibrated parameters is proportional to the joint likelihood function of the parameters $\mathcal{L}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)$ conditional on the data and the calibrated parameters multiplied by the prior distribution $\mathcal{P}(\mathbf{\Theta}_e|\mathbf{\Theta}_c)$. For inference on $\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c$, the unconditional likelihood $\mathcal{P}(\mathcal{M}_T|\mathbf{\Theta}_c)$ can be disregarded as it is invariant to changes in the parameters to be estimated and does not affect the properties of the posterior distribution which can always be scaled to represent a proper probability density function.

The idea of conventional Maximum Likelihood estimation is to choose the set of parameters $\mathbf{\Theta}_e$ which maximizes the joint likelihood $\mathcal{L}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)$ conditional on our linearized model and assumed distributions of the structural innovations (cf. Fernández-Villaverde 2009 and Guerron-Quintana and Nason 2012). If the model includes unobserved variables, the Kalman filter which is explained in detail in Appendix C is typically employed to evaluate the likelihood function.

The traditional Maximum Likelihood approach performs poorly for large macroeconomic models with limited data. Parameters tend to be only weakly identified and the likelihood function tends to have local maxima in regions of the parameter space which are not plausible from an economic theory perspective. Hence we follow a Bayesian approach and weigh the likelihood function with our prior beliefs of a plausible parameter distribution. This adds curvature to the otherwise possibly rather flat or wavy likelihood function and makes parameters in the off less likely.

The priors reflect our beliefs regarding the distribution of the estimated parameters before seeing the data. We need to choose the distribution and calibrate its moments, the so-called hyperparameters. Our confidence regarding our prior beliefs will affect the choice of the variance around the selected mean. A weak prior will have a large variance allowing a wide range of values to be likely realizations. In contrast, a strong prior with a low variance reflects high confidence that the posterior mode should lie in the close neighborhood of the prior mean. Choosing the prior distribution is a delicate task. We do not want to the prior to be too weak running danger that the posterior mode or mean end up in an implausible region of the parameter space. Yet, we do not want the prior to be too strong to dominate the likelihood. We will conduct a sensitivity analysis and check the dependence of our results on the choice of priors.

Bayes' rule allows us to obtain the joint posterior distribution $\mathcal{P}(\mathbf{\Theta}_e|\mathcal{M}_T, \mathbf{\Theta}_c)$. Yet, we are ultimately interested in the marginal posterior distribution of a specific parameter $\theta_{i,e}$,

---

[14]See Fernández-Villaverde (2009) and Guerron-Quintana and Nason (2012) for introductions to Bayesian Maximum Likelihood estimation of DSGE models.

which can be obtained by integrating out the other parameters from the joint posterior, i.e.

$$\mathcal{P}(\theta_{i,e}|\mathcal{M}_T, \Theta_c) = \int_0^\infty \mathcal{P}(\Theta_e|\mathcal{M}_T, \Theta_c)d\theta_{1,e}\ldots d\theta_{i-1,e}d\theta_{i+1,e}\ldots d\theta_{n,e}$$

which is the posterior distribution of a parameter independent of the realization of any other estimated parameter. Due to the non-linear dependence of the likelihood from the parameter vector, we are unable to compute $\mathcal{L}(\Theta_e|\mathcal{M}_T, \Theta_c)$ and, hence, $\mathcal{P}(\Theta_e|\mathcal{M}_T, \Theta_c)$ analytically. We rely on numerical methods to simulate the likelihood as well as the marginal posterior distributions of the parameters. In particular, we employ the Metropolis-Hastings Markov-Chain-Monte-Carlo (MH-MCMC) simulator. Before discussing the intuition of this algorithm for obtaining the posterior distributions, we briefly cover the choice of priors.

Obtaining the posterior distribution. Once we have specified the priors for our estimated parameters, we can feed all necessary information, i.e. the priors, the model and the data, into the MH-MCMC algorithm to simulate the posterior distribution of our parameters. Here, we focus on the intuition of this procedure. A treatment of the simulator in greater detail can be found in Appendix D.

The MH-MCMC algorithm generates a series of parameter vectors $\{\hat{\Theta}_{e,l}\}_{l=1}^H$ where H is the number of iterations. The nature of the algorithm implies that the simulated series $\{\hat{\theta}_{i,e,l}\}_{l=1}^H$ is drawn from a distribution that converges to the marginal posterior of interest $\mathcal{P}(\hat{\theta}_{i,e}|\mathcal{M}_T, \Theta_c)$. Hence, we can interpret the distributions of the simulated parameters as an approximation of the posterior distributions.

The core idea of the MH-MCMC algorithm is to have the parameter vector Θ_e move across the different iterations of the algorithm as a type of random walk. The parameter vector $\Theta_{e,l}$ is then stored for each iteration l to get $\{\hat{\Theta}_{e,l}\}_{l=1}^H$. Why can we perceive this series as being drawn from $\mathcal{P}(\hat{\theta}_e|\mathcal{M}_T, \Theta_c)$? This is because of the peculiar nature of the random walk process which the parameter vector follows over the iterations. In general, a random walk implies the *realization* in the current iteration to be the sum of the realization in a *reference* iteration (typically the previous one) and a random shock. The random walk process assumed here is peculiar regarding the reference. The realization of the parameter vector in l is not necessarily the reference in $l+1$. A decision rule determines if a realization in l qualifies as the reference value for the realization in $l+1$ or if the old reference parameter vector from $l-1$ should be kept.[15] It can be shown that the distribution of the simulated parameters converges to $\mathcal{P}(\hat{\theta}_e|\mathcal{M}_T, \Theta_c)$ if the acceptance rate of the decision rule is around 0.24.

---

[15]This decision rule compares the likelihood associated with the realization in l with the likelihood associated with the reference in l whose evaluation requires the Kalman filter discussed in Appendix B. Only if the likelihood of the realization exceeds the likelihood of the reference, will the realization be the next iteration's reference. If not, a draw from a uniform distribution between zero and one will be compared to the ratio of the realization's and reference's likelihood (which, then, is lower than one). If and only if this ratio exceeds the draw, the realization will be accepted as the next iteration's reference.

The random walk process includes a *jump* parameter $\bar{w}$ which controls the effect of a unit innovation to the realization and, hence, by how much the parameter vector jumps from one iteration to the other. We set this hyper parameter to 0.5 which ensures the acceptance ratio to lie between 0.23 and 0.30.

The number of iterations H should be chosen sufficiently large to ensure convergence of the distribution of the proposals $\hat{\Theta}_{e,l}$ to $\mathcal{P}(\hat{\Theta}_e | \mathcal{M}_T, \Theta_c)$. In our case, 50.000 iterations are sufficient for the distributions to converge.

Finally, the posterior distribution simulated by the MH-MCMC algorithm may depend on the initialization of the parameter vector. To reduce this dependency, we run 2 chains of iterations each initializing the parameter vector as the mean of a burn-in stage which, in turn, starts from the mode of the posterior distribution obtained by a MCMC based optimization routine. Each additional chain starts the burn-in from the next best mode. For all estimations reported in the present paper, the Brooks and Gelman (1998) diagnostics indicate convergence of the posterior distributions across chains and time.

Economic Theory Vector Auto-Regression. A useful tool for evaluating the empirical performance of an economic model and for spotting potential sources of misspecification has been developed by Del Negro and Schorfheide (2004) and Del Negro et al. (2007). It has been referred to as DSGE-VAR. Yet, since its application is not restricted to this type of model, we will refer to it as Economic Theory Vector Auto-Regression (ET-VAR). The underlying idea is the following: The model represented by (36) and (37) can be approximated by a VAR. The economic model implies cross-coefficient restrictions on the parameters of this VAR. Misspecification can then be measured by the extent to which the estimated VAR parameters deviate from the theory-implied restrictions. More specifically, a Bayesian VAR approximation of the economic model is estimated with the prior means of the VAR parameters being obtained by a mapping from the model parameters $\hat{\Theta}$, given the prior $\mathcal{P}(\Theta_e | \Theta_c)$. The tightness of the theory-implied prior of the VAR parameters is scaled by a hyperparameter λ. The larger λ, the stronger is the confidence in the cross-coefficient restrictions of the economic model. In particular, $\lambda = 0$ and $\lambda = \infty$ correspond to an unrestricted VAR and a fully restricted VAR, respectively. We interpret the marginal likelihood function of λ, i.e. $\mathcal{P}(\mathcal{M}_T | \lambda)$ as an overall measure of fit. It is the probability of the observed data given only the model and λ and independent of any specific parameter value. It can be numerically approximated using the output of the MH-MCMC simulator. Denote its mode by

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmax}} \ \mathcal{P}(\mathcal{M}_T | \lambda).$$

A large value of $\hat{\lambda}$ and a likelihood ratio (of $\lambda = \hat{\lambda}$ versus $\lambda = \infty$) close to 1 can be interpreted as evidence in favor of the economic theory restrictions.

Table 1: Estimation results

Parameters	Specifications 1-3			Specification 1				Specification 2				Specification 3			
	prior dist.	prior mean	prior std.	post. mode	mode std.	post. mean	90% HDP IV	post. mode	mode std.	post. mean	90% HDP IV	post. mode	mode std.	post. mean	90% HDP IV
τ	$\mathcal{G}$	50.00	10.00	44.33	3.11	43.65	[30.31 , 59.39]	44.68	4.36	44.63	[30.24 , 60.03]	44.35	6.31	44.84	[28.86 , 59.39]
ϕ_c	$\mathcal{B}$	.50	.20	.04	.03	.05	[.01 , .10]	.15	.08	.18	[.02 , .32]	.15	.09	.18	[.03 , .32]
ϕ_{iy}	$\mathcal{G}$	1.00	.50	.42	.16	.49	[.13 , .85]	1.81	.19	1.80	[1.40 , 2.22]	1.81	.25	1.82	[1.43 , 2.21]
ϕ_{ir}	$\mathcal{G}$	1.00	.50	.35	.11	.44	[.12 , .75]	.60	.21	.80	[.21 , 1.34]	.61	.29	.79	[.25 , 1.33]
ϕ_{gy}	$\mathcal{U}$	.00	2.31					-4.00	.19	-3.76	[-4.00 , -3.44]	-4.00	.25	-3.74	[-4.00 , -3.41]
$\phi_{r\pi}$	$\mathcal{U}$	.00	1.15					.12	.05	.12	[.04 , .19]	.12	.05	.12	[.04 , .19]
ϕ_{ry}	$\mathcal{U}$	.00	.58					.09	.02	.10	[.06 , .13]	.09	.02	.10	[.06 , .14]
$\phi_{\nu l}$	$\mathcal{U}$	.00	.58									.08	.13	.08	[-.14 , .29]
ρ_g	$\mathcal{B}$	.50	.20	.81	.04	.81	[.73 , .88]	.80	.05	.80	[.72 , .88]	.80	.05	.80	[.72 , .87]
ρ_a	$\mathcal{B}$	.50	.20	.76	.05	.76	[.67 , .85]	.76	.05	.75	[.67 , .85]	.76	.05	.75	[.66 , .84]
ρ_r	$\mathcal{B}$	.50	.20	.86	.03	.86	[.79 , .93]	.84	.05	.84	[.75 , .91]	.84	.05	.84	[.76 , .92]
ρ_c	$\mathcal{B}$	.50	.20	.74	.06	.74	[.65 , .83]	.75	.05	.74	[.66 , .83]	.75	.06	.75	[.66 , .84]
ρ_n	$\mathcal{B}$	.50	.20	.82	.04	.82	[.75 , .89]	.82	.04	.82	[.75 , .90]	.82	.04	.82	[.75 , .90]
ρ_i	$\mathcal{B}$	.50	.20	.86	.04	.86	[.79 , .92]	.82	.05	.82	[.75 , .90]	.82	.04	.82	[.74 , .90]
ρ_ν	$\mathcal{B}$	.50	.20	.09	.05	.11	[.02 , .19]	.09	.05	.11	[.02 , .18]	.08	.05	.11	[.02 , .19]
ρ_{re}	$\mathcal{B}$	.50	.20	.02	.00	.02	[.01 , .03]	.02	.00	.02	[.01 , .03]	.02	.00	.02	[.01 , .03]
ρ_{ye}	$\mathcal{B}$	.50	.20	.71	.09	.71	[.50 , .93]	.34	.07	.37	[.23 , .50]	.35	.08	.37	[.23 , .52]
σ_g	$\mathcal{IG}$	.01	.05	.01	.00	.01	[.01 , .01]	.03	.00	.03	[.02 , .03]	.03	.00	.03	[.02 , .03]
σ_a	$\mathcal{IG}$	.01	.05	.00	.00	.00	[.00 , .00]	.00	.00	.00	[.00 , .00]	.00	.00	.00	[.00 , .00]
σ_r	$\mathcal{IG}$	.01	.05	.00	.00	.00	[.00 , .00]	.00	.00	.00	[.00 , .00]	.00	.00	.00	[.00 , .00]
σ_c	$\mathcal{IG}$	.01	.05	.02	.00	.02	[.02 , .02]	.02	.00	.02	[.02 , .02]	.02	.00	.02	[.02 , .02]
σ_n	$\mathcal{IG}$	.01	.05	.00	.00	.00	[.00 , .01]	.00	.00	.00	[.00 , .01]	.00	.00	.00	[.00 , .01]
σ_i	$\mathcal{IG}$	.01	.05	.01	.00	.01	[.01 , .01]	.01	.00	.01	[.01 , .01]	.01	.00	.01	[.01 , .01]
σ_ν	$\mathcal{IG}$	.01	.05	.01	.00	.01	[.01 , .01]	.01	.00	.01	[.01 , .01]	.01	.00	.01	[.01 , .01]

Notes: $\mathcal{B}$, $\mathcal{U}$, $\mathcal{G}$ and $\mathcal{IG}$ denote the Beta, Uniform, Gamma and Inverse Gamma distributions, respectively. τ is the price adjustment cost scaling parameter, ϕ_c is the income elasticity of rentier's consumption, ϕ_{iy} is the expected sales elasticity of investment, ϕ_{ir} is the interest rate elasticity of investment, ϕ_{gy} is the output elasticity of government consumption, $\phi_{r\pi}$ is the inflation elasticity of the interest rate, ϕ_{ry} is the output elasticity of the interest rate, ϕ_w is the unemployment elasticity of the wage inflation rate, ρ_x is the persistence in the evolution of variable x after a shock ε_x to variable x, and σ_x is the standard deviation of shock ε_x.

3.2 Estimation results

Estimation results are reported in Table 1. Three specifications are considered. The baseline abstracts from endogenous fiscal and monetary policy as well as endogenous wage formation by setting $\phi_{gy} = \phi_{r\pi} = \phi_{ry} = \phi_{\nu l} = 0$. The second specification includes endogenous fiscal and monetary policy but not wage formation. The third specification finally involves the entire model.

The second to fourth columns in Table 1 report the prior distribution, mean and standard deviation for each of the estimated parameters. Regarding the distributional form of the priors we choose an Inverse Gamma distribution for the standard deviations of the shocks, a Beta distribution for parameters that should lie between zero and one (ϕ_c,ρ_g,ρ_a,ρ_r,ρ_c,ρ_n,ρ_i,ρ_ν,$\rho_{\pi e}$,ρ_{ye}), a Gamma distribution for parameters that should be positive (τ,ϕ_{iy},ϕ_{ir}) and a Uniform distribution for parameters for which we perceive, a priori, any realization within a given range as equally likely (ϕ_{gy},$\phi_{r\pi}$,ϕ_{ry},$\phi_{\nu l}$). Most of the prior means have been chosen loosely following calibration choices in the related PK literature as well as previous empirical contributions. Yet, we assign low confidence to these prior choices by setting the standard deviations rather large. The mean for the price adjustment parameter is set to $\tau = 50$ which is in the range of the findings by Schoder (2015a) estimating a Dynamic Stochastic Disequilibrium (DSDE) model as well as a DSGE model for the Euro Area. For the income elasticity of rentier's consumption and the expected sales elasticity of investment as well as the negative interest elasticity of investment we choose $\phi_c = 0.5$, $\phi_{iy} = 1$ and $\phi_{ir} = 1$, respectively, which is around the findings in the PK literature which typically estimates marginal effects (cf. Hein and Schoder 2011). We attach very low confidence to these beliefs. The priors attached to the fiscal and monetary policy responses as well as the unemployment elasticity of wage inflation are uninformative zero mean uniform distributions. The means of the autoregressive coefficients are set to 0.5 and of the standard deviations of the shocks to 0.01 with standard deviations of 0.2 and 0.05, respectively.

For each specification, the posterior mode and its standard deviation, Table 1 also reports the posterior mean as well as its 90% confidence intervals. The posterior mode obtained by a MCMC optimization routine serves as the initialization of the burn-in stage of the MH-MCMC posterior simulator. The fact that posterior mode and mean do not diverge too much indicates that our likelihood or our priors are strong enough to prevent the posterior from spiking at different locations in the parameter space. A strong likelihood is also confirmed by the observation that, for most parameters, the posterior confidence intervals are smaller than implied by the prior alone.

The posterior mean for the price adjustment cost scaling parameter is $\tau = 44.84$ in the third specification. This result is robust to excluding endogenous wage formation and, additionally, endogenous policy. It is also broadly consistent with the findings by Schoder (2015c) based on a DSDE and a DSGE model.

The mean of the income elasticity of rentier's consumption is $\phi_c = 0.18$ with a 90% probability that it lies between 0.03 and 0.32 which is a rather wide range. This result

is robust to excluding the endogenous wage formation. Without endogenous policy, the parameter mean is found to be 0.05. The corresponding marginal propensities to consume out of rentier's income $\phi_c C_r^{ss}/Z_r^{ss}$ are 0.09 for the second and third specification and 0.02 for the first specification. The marginal propensities to consume out of (steady-state) wealth $(1-\phi_c) C_r^{ss}/B_r^{ss}$ are 0.0028 and 0.0033, respectively. These values are broadly consistent with previous empirical analyses (cf. Naastepad and Storm 2006-07, Hein and Vogel 2008, Hein and Schoder 2011) and with calibration choices in the literature on stock-flow consistent modeling (cf. Godley and Lavoie 2012, Schoder 2014b).

The response of investment to expected sales is around $\phi_{iy} = 1.82$ but becomes much less pronounced ($\phi_{iy} = 0.49$) if we exclude endogenous fiscal and monetary policy. The interest rate elasticity of investment is about $\phi_{ir} = 0.79$ in the second and third specifications and lower (0.44) in the first specification. Transforming these values into marginal effects on investment, $\phi_{iy} I^{ss}/Y^{ss}$ and $\phi_{ir} I^{ss}/R^{ss}$, we obtain 0.53 and 0.23, respectively, for the endogenous policy specifications and 0.14 and 0.13, respectively, for the exogenous policy specification. The former results are more consistent with the findings in the related literature (cf. Hein and Schoder 2011, Stockhammer et al. 2011).

The output elasticity of fiscal policy $\phi_{gy} = -3.74$ is strong, exhibits the expected sign and is robust. The marginal effect of output on government spending is $\phi_{gy} G^{ss}/Y^{ss} = -0.75$. The monetary policy responses to inflation and output, $\phi_{r\pi} = 0.12$ and $\phi_{ry} = 0.10$, respectively, are weak compared to Schoder's (2015a) findings obtained from models with household and firm behavior derived from standard micro-foundations.

One striking finding is that the Euro Area data does not seem to support a strong or significant effect of employment on the rate of nominal wage inflation. The posterior mean 0.08 is positive but the 90% interval ranges from -0.14 to 0.29. Further research may want to investigate these failure to find a Goodwin (1967) labor market feedback mechanism that is at the core of prominent PK business cycle theories (cf. Taylor 2004, Flaschel 2009, Schoder 2014b). One reason may be the fact that the Euro Area is a highly aggregated economy which may blur a possible causal relationship from employment to wages on the level of the nation state.

Regarding the persistence of shocks, the Bayesian Maximum Likelihood estimation yields high autoregressive coefficients for the productivity shock, consumption shock, labor supply shock and investment shock. Note that the bargaining power shock process exhibits almost no persistence with $\rho_\nu = 0.11$. Finally, an autoregressive parameter for the adjustment of inflation expectations of $\rho_{\pi e} = 0.02$ implies that a disappointment in the expectation formation causes a strong revision of expectations.

In general, note that the 90% confidence interval is rather narrow despite weak priors. This suggests that the likelihood is strong and drives the results.

To compare the overall performance of the three specifications, we assess the marginal likelihood, i.e. the unconditional probability of the observed data given only the economic model (and λ), for different specifications: (a) an ET-VAR with restrictions as loose as

Table 2: Marginal likelihoods of ET-VAR(λ) for different λ and of the economic model.

	(a) ET-VAR(0.2537)	(b) ET-VAR($\hat{\lambda}$)	(c) ET-VAR(∞)	(d) Econ. model (full smpl.)	(e) Econ. model (full smpl. minus training smpl.)
Spec. 1	-385.02	-355.50 ($\hat{\lambda} = 0.46$)	-487.12	-471.47	-307.07
Spec. 2	-349.52	-313.78 ($\hat{\lambda} = 0.50$)	-418.49	-424.13	-263.49
Spec. 3	-349.21	-314.51 ($\hat{\lambda} = 0.51$)	-439.51	-425.62	-265.10
Smets and Wouters (2003)	-234.92	-171.06 ($\hat{\lambda} = 0.66$)	-269.42	-272.97	-138.04

possible, i.e. choosing the smallest hyperparameter possible ($\lambda = 0.2537$) for the prior distributions not to be degenerate (cf. Del Negro and Schorfheide 2004, Del Negro et al. 2007); (b) an ET-VAR with the optimal $\lambda = \hat{\lambda}$ that maximizes the marginal likelihood $\mathcal{P}(\mathcal{M}_\mathcal{T}|\lambda)$ estimated; (c) an ET-VAR that corresponds to a fully restricted approximation of the economic model ($\lambda = \infty$). We also consider the marginal likelihood of the (d) log-linearized economic model. To reduce the sensitivity of the results to the choice of priors and number of estimated parameters, we finally report the marginal likelihood of a (e) specification that uses the first 40 observations as a training sample (Sims 2003). Regarding economic models, we consider the three specifications outlined above and, as a reference, the Euro Area DSGE model proposed by Smets and Wouters (2003) which we have fitted to our data set.[16]

For each specification including Smets and Wouters (2003), Table 2 reports the marginal likelihoods $\mathcal{P}(\mathcal{M}_\mathcal{T}|\lambda)$ of the ET-VARs with different λ's as well as the economic model for the full sample and the training-corrected sample. The following observations are worth to note:

First, the marginal likelihoods for ET-VAR(∞) and the economic model differ even

---

[16]Note that the Smets and Wouters (2003)-model does not feature unemployment. Since we would like to fit the model to our data set which includes a series for unemployment, we add a measurement equation which relates unemployment to the ratio of employed labor and the labor force, which obviously introduces some degree of misspecification to the model.

though the former is an approximation of the latter. This is basically an approximation error also present in Del Negro et al. (2007).

Second, the optimal λ's lie somewhere between the fully unrestricted and the fully restricted models. Hence, starting from the unrestricted ET-VAR, adding information contained in the economic model by tightening up the priors of the ET-VAR improves the fit. Yet, if the confidence put upon the prior becomes too strong, i.e. if the restrictions implied by the economic model get too much emphasis, the fit will deteriorate again. This pattern is also observed by Del Negro et al. (2007).

Third, for the PK model specifications, the estimated $\hat{\lambda}$'s are rather low indicating that the theory-imposed restrictions should not be too strong in order for the ET-VAR(λ) to describe the data well. This is first evidence that the PK model outlined above may not perform too well empirically. The $\hat{\lambda}$ of the DSGE model is also low but considerable larger than those of the PK models. Hence, the DSGE model seems to outperform the PK models on our data set.

Fourth, the marginal likelihoods are considerably larger for the DSGE model than for the PK models suggesting, again, that the former performs relatively better.

Fifth, including endogenous policy in the PK model improves its empirical performance considerably. Even $\hat{\lambda}$ increases from the first to the second specification. Yet, introducing labor market feedback on nominal wage formation does not improve the fit. Again, this is a puzzling result given the emphasis on Goodwin labor market dynamics in the PK literature.

3.3   Impulse-response analysis

To further evaluate the economic model and spot potential sources of misspecification, this section compares the macroeconomic responses to structural shocks as predicted by the estimated ET-VAR(∞) approximation of the economic model and the ET-VAR($\hat{\lambda}$). The latter is the benchmark against which the performance of the economic model is assessed.[17] In particular we analyze the Bayesian impulse-response functions (IRFs) for a government spending shock, a monetary policy shock, a shock to the worker's bargaining power and a productivity shock. Note further that this analysis is conducted only for the third specification of our model including both endogenous policy and wage formation.

The Bayesian IRFs are obtained from the NH-MCMC simulations. Each iteration of this algorithm generates and stores not only a set of parameters but also a set of impulse-response functions based on the policy functions, or decision rules, implied by the model. A response $y_{xz,l,h}$ for a specific variable x to a specific impulse z in a specific iteration l illustrates how, in

---

[17]Note that the ET-VAR(∞) approximation should yield predictions very similar to the predictions of the economic model itself. Yet, since estimating the model and its approximation yields slightly different results, the Bayesian impulse-response functions are not identical. Moreover, the impulse-response functions will not be comparable as the identified impulses will differ. Hence, we limit our analysis to comparing the impulse-responses of the ET-VAR(∞) and the ET-VAR($\hat{\lambda}$).
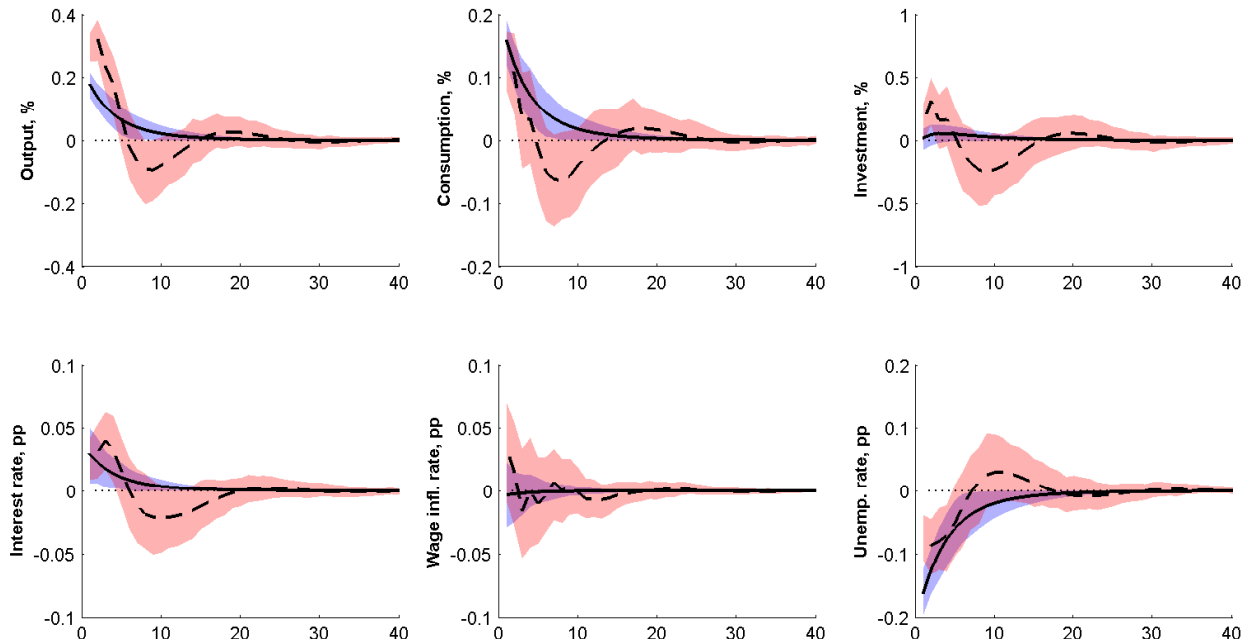
Figure 1: Macroeconomic responses to a fiscal policy shock, ε_g, as predicted by the ET-VAR(∞) (solid line, blue band) and by the ET-VAR($\hat{\lambda}$) (dashed line, red band) and measured as deviations from the steady state.

a given iteration l, x evolves over time h after a one-time rise in z by one standard deviation. The Bayesian IRF plots for each $h = 1, \ldots, 40$, the median of $y_{xz,l,h}$ over all iterations l as well as the 5% and 95% percentiles which constitute the confidence band.

Government spending shock. We consider a one-standard-deviation one-time shock to government expenditures, ε_g. The Bayesian IRFs for the economic model, for its ET-VAR(∞) approximation and for the ET-VAR(λ) are plotted in Figure 1.

Let us first focus on the response as predicted by the ET-VAR(∞) approximation of the model. As expected, a fiscal stimulus has expansionary effects. A rise in government consumption immediately translates into higher aggregate demand causing output to rise and unemployment to fall. Since the labor market feedback on wage formation is weak and insignificant, nominal wage inflation is not predicted to change considerably. With unchanged unit variable costs, inflation remains basically unaffected. Hence, the real wage is not found to respond either. We do not find evidence for the real wage to move pro-cyclically with a lag as observed by Barbosa-Filho and Taylor (2006), Zipperer and Skott (2010) and Flaschel (2009). Expanding employment increases the wage bill and stimulates consumption. With output going up, expected sales increase slowly causing firms to increase investment. Hence, investment responds with a lag. Overall the multiplier effects through consumption and

23

investment as well as the persistence in the fiscal stimulus keep output above steady state for about 15 to 20 quarters. Note that government spending is suppressed by this expansion due to automatic stabilizers built into the fiscal rule. Also monetary policy responds by increasing the interest rate.

What are the quantitative effects? Estimation of the economic model has identified the initial expansion of government consumption to be about 0.5% of steady-state government consumption which has been calibrated to be 20% of GDP which, in turn, has been normalized to unity. Hence, $0.005 \cdot 0.2 = 0.001$. The rise in output is about 0.2% of steady-state GDP, i.e. 0.002. Hence, the impact multiplier is about $0.002/0.001 = 2$. Note that this multiplier effect holds when the economy is shocked in the steady state and the shock does not push the economy away from it too far. The size of the fiscal multiplier is in line with the DSDE model estimated in Schoder (2015a) but large compared to the typical findings in the DSGE literature. Smets and Wouters (2003) estimate a multiplier of about one due to crowding out of consumption and investment. In contrast to that, our PK model exhibits a strong multiplier effect on consumption increasing by 0.15% which is consistent with empirical observation (cf. Fatás and Mihov 2001 and Blanchard and Perotti (2002)). Investment increases by up to 0.1%. The output expansion is associated with a rise of the interest rate by up to 0.05%-points and a reduction of unemployment by 0.5%-points.

Contrasting the IRFs of the ET-VAR(∞) to the IRFs of the ET-VAR($\hat{\lambda}$) indicates that one core reason for the relatively poor performance of our estimated PK model in terms of marginal likelihood is its failure to predict the cyclical adjustment back to the steady state after a fiscal expansion as suggested by the data. These cycles do not seem to be driven by the interaction of distribution and economic activity as the ET-VAR($\hat{\lambda}$) does not suggest a cyclical movement of the wage inflation rate.

Monetary policy shock. Figure 2 plots the Bayesian IRFs for a contractionary monetary policy shock amounting to an increase of the interest rate by 12%-points. The macroeconomic effects of monetary policy are predicted to be small in our PK model as the interest rate enters the model only through the investment behavior of the firm. In DSDE and DSGE models based on mainstream micro-foundations, the interest rate affects the output additionally through inter-temporal substitution of consumption and investment. Overall the comparison of the two sets of IRFs suggest three issues of the PK model outlined above regarding monetary policy. First, it is unable to replicate the quantitative effect that monetary policy has on core macroeconomic variables. Second, in contrast to the economic model, the ET-VAR($\hat{\lambda}$) suggests the interest rate to affect the economy with a considerable lag with the peak after 8 quarters. Third, again the estimated economic model does not generate a cyclical adjustment to the steady state.

The lack of our model to account for the monetary transmission mechanism observed empirically as also documented, for instance, by Peersman and Smets (2001) indicates that PK theory may have to be revised to better suit the purpose of monetary policy analysis (as suggested by Lavoie 1995 and Hein 2007).
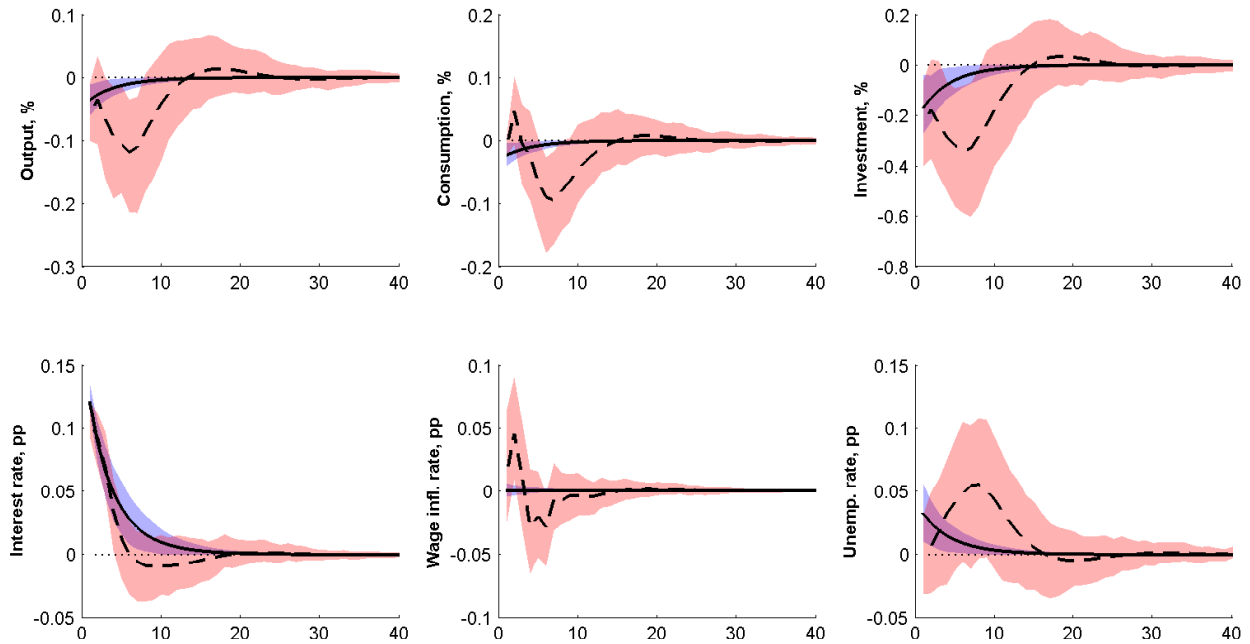
Figure 2: Macroeconomic responses to a monetary policy shock, ε_r, as predicted by the ET-VAR(∞) (solid line, blue band) and by the ET-VAR($\hat{\lambda}$) (dashed line, red band) and measured as deviations from the steady state.

Wage bargaining power shock. Figure 3 presents the Bayesian IRFs for a shock to the relative bargaining power of workers. A strengthening of the workers' bargaining power immediately increases the real wage even though nominal inflation is increasing, too. The overall effect on output and employment is ambiguous. On the one hand, a higher real wage stimulates worker consumption. On the other, monetary policy responds to inflationary pressures with a higher interest rate cutting into investment. Due to the low persistence of the bargaining power shock, wage inflation returns quickly back to the steady state. Price inflation adjusts slower causing the real wage to decrease below steady state after the first few quarters. Hence, our estimated PK model does not give a clear answer on the question of wage-led or profit-led demand.

Attributing only little weight to the restrictions imposed by economic theory, the ET-VAR($\hat{\lambda}$) does not identify a straightforward effect of the worker's bargaining power on economic activity.

Productivity shock. Finally, the Bayesian IRFs for a productivity shock are plotted in Figure 4. An increase in productivity raises unemployment and reduces the price inflation rate slightly. Due to the lack of a labor market feedback on the wage formation, the PK model does not predict productivity to affect the economy much. Similarly, the ET-VAR($\hat{\lambda}$) does not indicate an unambiguous response to a productivity shock.
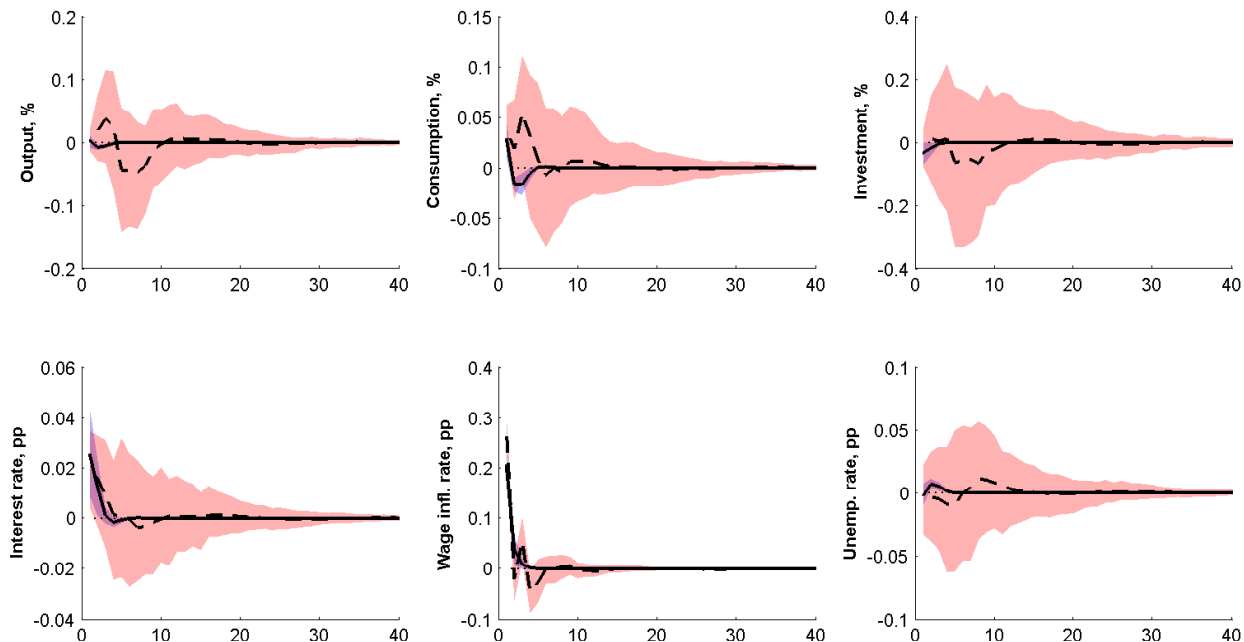
Figure 3: Macroeconomic responses to a wage bargaining power shock, ε_ν, as predicted by the ET-VAR(∞) (solid line, blue band) and by the ET-VAR($\hat{\lambda}$) (dashed line, red band) and measured as deviations from the steady state.

3.4 Forecast error variance decomposition

We may want to ask which shocks drive the fluctuations in the data conditional on our economic model and estimated parameters. The forecast error variance decomposition reported in Table 3 addresses this question for the specification including endogenous policy and wage formation. It indicates the shares of each shock in contributing to the forecast error variance in each of the observed variables. We consider the forecast error for two time horizons: one-period and four-periods ahead. The results do not differ considerably.

The government spending shock is found to contribute tremendously to the variation in the data. Almost 80% of the variation in forecast error of output is traced back to this shock. Also big junks of the variations in consumption, investment, and the unemployment rate are found be related to government spending shocks. Yet, they hardly drive prices and wages. Wage inflation is almost completely driven by shocks to the bargaining power of workers which is not surprising given the low feedback of the labor market. Through unit variable costs in the Kaleckian mark-up price equation, bargaining power shocks as well as productivity shocks contribute to the variation of price inflation. Note that because of only weak feedback effects of the rest of the economy on price and wage formation, price and wage inflation are primarily driven by exogenous processes. Demand shocks do not induce firms to increase prices because of more or less stable unit variable costs. Due to the missing labor
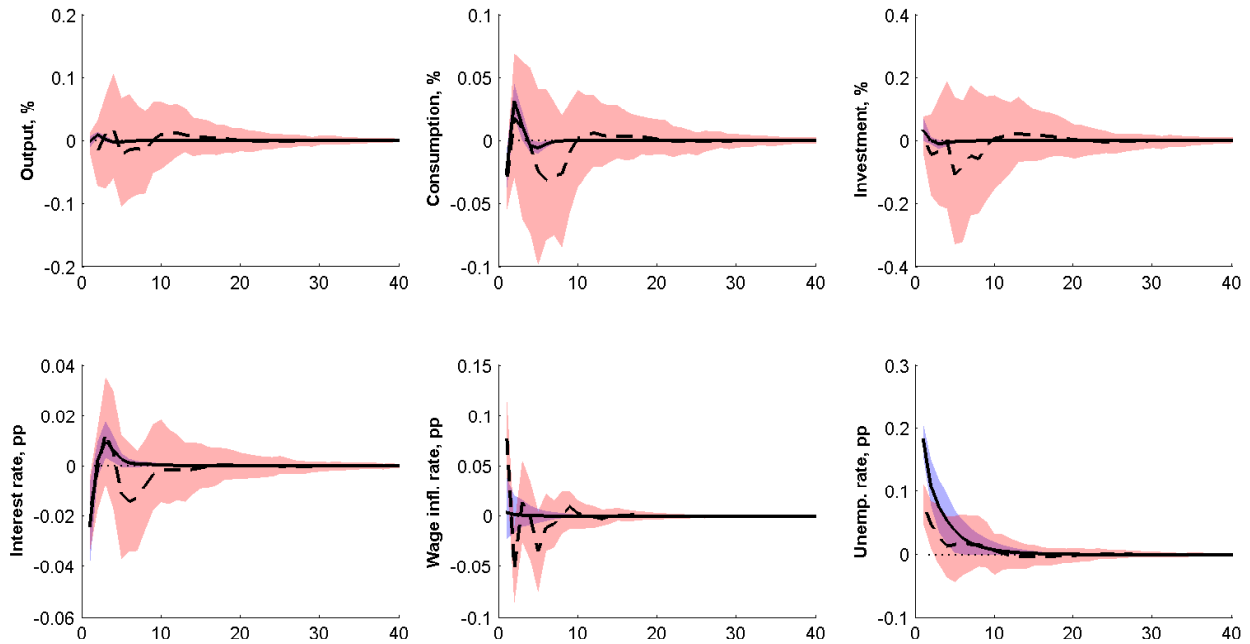
Figure 4: Macroeconomic responses to a productivity shock, ε_a, as predicted by the ET-VAR(∞) (solid line, blue band) and by the ET-VAR($\hat{\lambda}$) (dashed line, red band) and measured as deviations from the steady state.

market feedback on wage formation, labor supply shocks only affect unemployment but not any of the other variables.

Figure 5 addresses the question what shocks drove the variation in the variables form a historical perspective. It illustrates how the Bayesian estimation procedure decomposes the variation of the data into variations of different shocks over time. We conduct this exercise for the historical trend deviation of unemployment in percent, which is plotted in each panel over time as the bold black line. The blue line is the hypothetical trend deviation over time if the model was only hit by series of a specific shock as estimated by the Kalman filter during the MCMC simulation. For instance, the blue line in the first panel is the hypothetical trend deviation of unemployment if the economy was only hit by the series of government spending shocks that the Kalman filter extracted from the data. The following observations are worth to note: First, shocks to government consumption and investment seem to closely track the unemployment rate, the former especially since the introduction of the Euro. Second, the bargaining power shock does not create much variation in the unemployment rate. Finally, shocks to productivity, consumption and labor supply do generate variation in the unemployment rate but do not track closely the observed time series.

Table 3: Conditional forecast error variance decomposition (in percent)

Period 1	ε_g	ε_a	ε_r	ε_c	ε_n	ε_i	ε_ν
$d_{y,t}$	78.87	0.16	0.12	8.31	0.00	12.38	0.16
$d_{c,t}$	40.63	1.41	0.04	51.89	0.00	4.63	1.39
$d_{i,t}$	24.25	0.00	0.69	2.55	0.00	72.51	0.00
$d_{r,t}$	11.46	6.30	73.01	1.21	0.00	1.80	6.22
$d_{\pi_p,t}$	0.05	50.31	0.00	0.01	0.00	0.01	49.62
$d_{\pi_w,t}$	0.11	0.06	0.00	0.01	0.00	0.02	99.81
$d_{u,t}$	38.74	20.95	0.06	4.08	30.01	6.08	0.08

Period 4	ε_g	ε_a	ε_r	ε_c	ε_n	ε_i	ε_ν
$d_{y,t}$	79.37	0.13	0.13	7.26	0.00	13.02	0.10
$d_{c,t}$	46.36	1.11	0.06	45.90	0.00	5.68	0.89
$d_{i,t}$	35.89	0.01	0.60	3.26	0.00	60.22	0.02
$d_{r,t}$	13.63	3.11	76.64	1.25	0.00	2.24	3.13
$d_{\pi_p,t}$	0.21	48.89	0.00	0.02	0.00	0.03	50.85
$d_{\pi_w,t}$	0.30	0.10	0.00	0.03	0.00	0.05	99.52
$d_{u,t}$	43.99	15.41	0.07	4.02	29.23	7.21	0.06

Notes: $d_{y,t}$, $d_{c,t}$, $d_{i,t}$, $d_{r,t}$, $d_{\pi_p,t}$, $d_{\pi_w,t}$, and $d_{u,t}$ are the steady-state deviations of output, consumption, investment, the interest rate, the price inflation rate, the wage inflation rate, and the unemployment rate, respectively. ε_g, ε_a, ε_r, ε_c, ε_n, ε_i, and ε_ν are disturbances to government consumption, productivity, the interest rate, consumption, labor supply, investment, and the worker's bargaining power, respectively.

4   Concluding remarks

The paper's main objective has been to introduce the method of Bayesian Maximum Likelihood to the empirical post-Keynesian (PK) literature. By showing its virtues in combining theory and empirics in a transparent and rigorous manner, we conclude, on a methodological level, that Bayesian Maximum Likelihood should be established as one standard tool for PK model estimation and evaluation. It may help considerably to advance PK theorizing by strengthening the ties between theory and empirics.

In particular, we have outlined a simple neo-Kaleckian type of model which is sufficiently rich in dynamics, variables and shocks to be fitted to the data. The model has then been estimated for the Euro Area using Bayesian Maximum Likelihood. Special attention has been paid to providing a thorough understanding of the intuition of this estimation strategy.
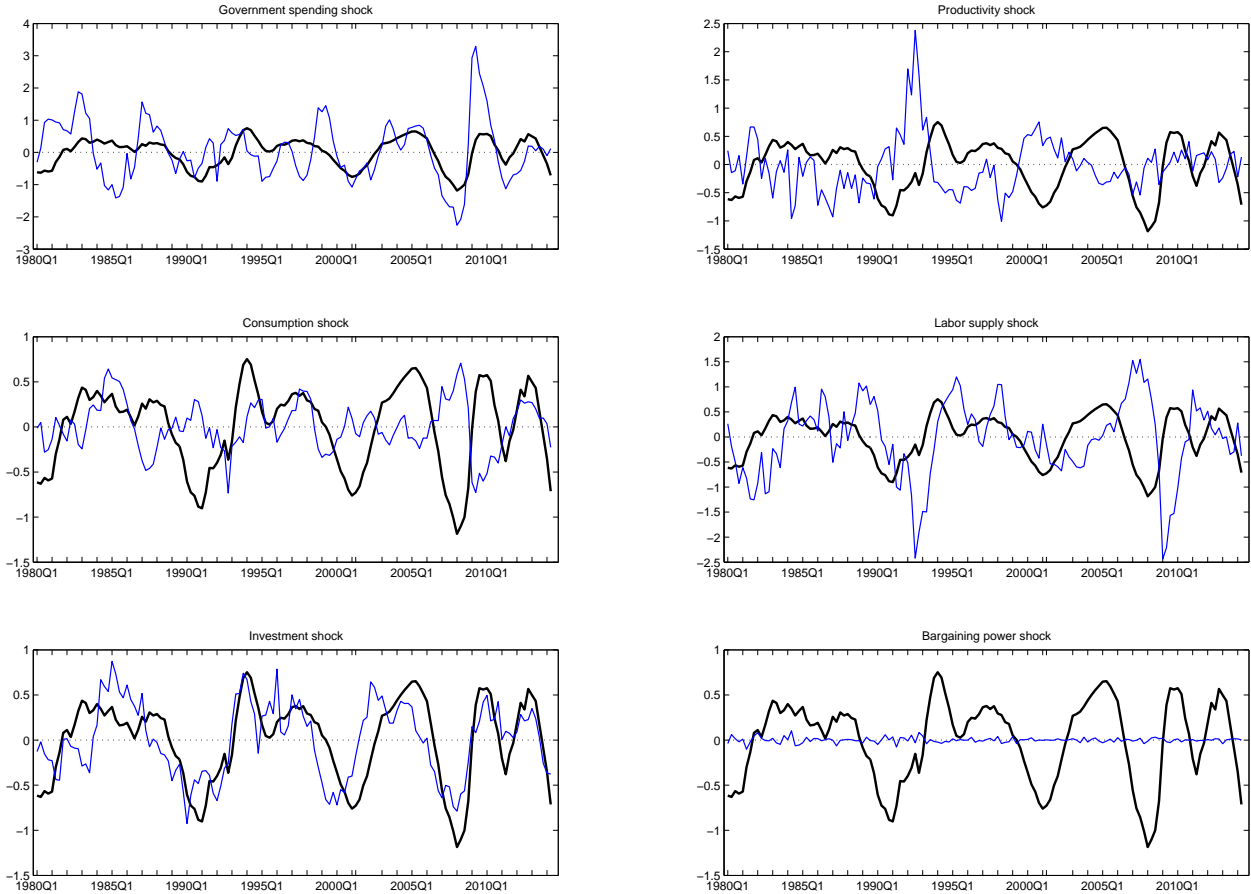
Figure 5: Historical decomposition for the trend deviation of unemployment in percent.

We have shown how this method can be used to evaluate the model fit in terms of marginal likelihood, analyze the dynamics of the system in terms of Bayesian impulse-response functions, and to decompose the variation observed in the data into contributions of different structural shocks.

The results have been obtained for a highly stylized PK model of business cycle fluctuations. They have to be interpreted with caution and other very legitimate specifications may yield very different results. Nevertheless, the following conclusions may be drawn from our analysis and may be seen as a starting point for future research seeking to shed further light on the issues identified.

First, the posterior means of the estimated parameters are more or less of the expected size despite very loose or even uninformative priors. The results obtained for the full model including endogenous policy and wage formation are robust to excluding these features. The only surprising result is the insignificant employment elasticity of the bargaining power which suggests that there is not much evidence for Goodwin wage dynamics in the data. Since this mechanism is the main channel through which the supply side affects the demand side in

our economy, we do not find productivity and bargaining power shocks to affect aggregate demand and its components much.

Second, despite the robust estimation results and good diagnostics of the estimation, the PK model's empirical performance is rather poor. The marginal likelihood of the data increases considerably if we loosen the theory-implied restrictions on the prior of a Bayesian Vector Auto-Regression (VAR) model approximating the PK model which indicates misspecification. Moreover, the benchmark Dynamic Stochastic General Equilibrium model proposed by Smets and Wouters (2003) performs considerably better.

Third, comparing the impulse-response functions of two Bayesian VARs, which differ only by the tightness of the theory-implied cross-coefficient restrictions on the prior means, allows us to narrow down the sources of misspecification. The estimated PK model outlined in this paper does not replicate the cyclical adjustment to the steady state after demand shocks and is unable to capture the fiscal and monetary transmission mechanisms satisfyingly.

Fourth, a change in income distribution triggered by a bargaining power shock does not unambiguously affect economic activity. Given the discordance in the empirical literature on wage-led vs. profit-led demand, this may not be that surprising after all.

Rather than finding the best-performing PK model, the main objective of the present paper has been to introduce an appealing method of model estimation and evaluation to the PK literature. Future research may take the current model as a starting point, modify or extend it to capture competing economic arguments. Advancing economic theory with a steady glance on the strengths and weaknesses of a given model in explaining the macroeconomic patterns observed may prove a fruitful endeavor.

References

Allain, O. (2015). Tackling the instability of growth: a kaleckian-harrodian model with an autonomous expenditure component. *Cambridge Journal of Economics*, 39(5):1351–1371.

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews*, 26(2-4):113–172.

Barbosa-Filho, N. H. and Taylor, L. (2006). Distributive and demand cycles in the us economy-a structuralist goodwin model. *Metroeconomica*, 57(3):389–411.

Bhaduri, A. and Marglin, S. (1990). Unemployment and the real wage: the economic basis for contesting political ideologies. *Cambridge Journal of Economics*, 14:375–93.

Blanchard, O. and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, 117(4):pp. 1329–1368.

Blecker, R. A. (1989). International competition, income distribution and economic growth. *Cambridge Journal of Economics*, 13:395–412.

Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.

Carvalho, L. and Rezai, A. (2014). Personal Income Inequality and Aggregate Demand. Working Papers, Department of Economics 2014-23, University of São Paulo (FEA-USP).

Christiano, L. J., Trabandt, M., and Walentin, K. (2010). DSGE Models for Monetary Policy Analysis. In Friedman, B. M. and Woodford, M., editors, *Handbook of Monetary Economics*, volume 3 of *Handbook of Monetary Economics*, chapter 7, pages 285–367. Elsevier.

Christoffel, K., Coenen, G., and Warne, A. (2008). The new area-wide model of the euro area: a micro-founded open-economy model for forecasting and policy analysis. Working Paper Series 0944, European Central Bank.

Del Negro, M. and Schorfheide, F. (2004). Priors from General Equilibrium Models for VARS. *International Economic Review*, 45(2):643–673.

Del Negro, M., Schorfheide, F., Smets, F., and Wouters, R. (2007). On the fit of new keynesian models. *Journal of Business & Economic Statistics*, 25(2):123–143.

Duménil, G. and Lévy, D. (1999). Being keynesian in the short term and classical in the long term: The traverse to classical long-term equilibrium. *The Manchester School*, 67(6):684–716.

Edge, R. M. and Gurkaynak, R. S. (2010). How Useful Are Estimated DSGE Model Forecasts for Central Bankers? *Brookings Papers on Economic Activity*, 41(2 (Fall)):209–259.

Fatás, A. and Mihov, I. (2001). The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. CEPR Discussion Papers 2760, C.E.P.R. Discussion Papers.

Fernández-Villaverde, J. (2009). The Econometrics of DSGE Models. NBER Working Papers 14677, National Bureau of Economic Research, Inc.

Fernández-Villaverde, J., Guerrón-Quintana, P., and Rubio-Ramŕez, J. F. (2010). Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data. NBER Working Papers 15928, National Bureau of Economic Research, Inc.

Fernández-Villaverde, J., Guerrón-Quintana, P. A., Kuester, K., and Rubio-Ramírez, J. (2011). Fiscal Volatility Shocks and Economic Activity. NBER Working Papers 17317, National Bureau of Economic Research, Inc.

Flaschel, P. (2009). *The Macrodynamics of Capitalism: Elements for a Synthesis of Marx, Keynes, and Schumpeter*. Springer, Berlin.

Foley, D. K. and Michl, T. R. (1999). *Growth and Distribution*. Harvard University Press, Cambridge, London.

Godley, W. and Lavoie, M. (2012). *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*. Palgrave Macmillan, second edition edition.

Goodwin, R. M. (1967). A growth cycle. In Feinstein, C. H., editor, *Capitalism and Economic Growth*, pages 54–8. Cambridge University Press, Cambridge.

Guerron-Quintana, P. A. and Nason, J. M. (2012). Bayesian estimation of DSGE models. Working Papers 12-4, Federal Reserve Bank of Philadelphia.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica*, 12(Supplement):1–118.

Hein, E. (2007). Interest rate, debt, distribution and capital accumulation in a post-kaleckian model. *Metroeconomica*, 58(2):310–39.

Hein, E. and Schoder, C. (2011). Interest rates, distribution and capital accumulation – a post-kaleckian perspective on the us and germany. *International Review of Applied Economics*, 25(6):693–723.

Hein, E. and Stockhammer, E. (2010). Macroeconomic policy mix, employment and inflation in a post-keynesian alternative to the new consensus model. *Review of Political Economy*, 22(3):317–354.

Hein, E. and Vogel, L. (2008). Distribution and growth reconsidered - empirical results for six oecd countries. *Cambridge Journal of Economics*, 32(3):479–511.

Jean, S. and Nicoletti, G. (2002). Product Market Regulation and Wage Premia in Europe and North America: An Empirical Investigation. OECD Economics Department Working Papers 318, OECD Publishing.

Kalecki, M. (1971). *Selected Essays on the Dynamics of the Capitalist Economy*. Cambridge University Press, Cambridge.

Lavoie, M. (1995). Interest rates in post-keynesian models of growth and distribution. *Metroeconomica*, 46:146–77.

Lavoie, M. (1996). Traverse, hysteresis, and normal rates of capacity utilization in kaleckian models of grwoth and distribution. *Review of Radical Political Economics*, 28(4):113–47.

Lavoie, M., Rodríguez, G., and Seccareccia, M. (2004). Similitudes and discrepancies in post-keynesian and marxist theories of investment: A theoretical and empirical investigation. *International Review of Applied Economics*, 18(2):127–49.

Lima, G. T. and Meirelles, A. J. A. (2007). Macrodynamics of debt regimes, financial instability and growth. *Cambridge Journal of Economics*, 31(4):563–580.

Martins, J. O., Scarpetta, S., and Pilat, D. (1996). Mark-Up Ratios in Manufacturing Industries: Estimates for 14 OECD Countries. OECD Economics Department Working Papers 162, OECD Publishing.

Meirelles, A. J. and Lima, G. T. (2006). Debt, financial fragility, and economic growth: a Post Keynesian macromodel. *Journal of Post Keynesian Economics*, 29(1):93–115.

Minsky, H. P. (1976). *John Maynard Keynes*. Macmillan, London.

Minsky, H. P. (1985). The financial instability hypothesis: A restatement. In Arestis, P. and Skouras, T., editors, *Post Keynesian Economic Theory. A Challenge to Neoclassical Economics*, pages 24–54. M. E. Sharpe, New York.

Modigliani, F. (1986). Life Cycle, Individual Thrift, and the Wealth of Nations. *American Economic Review*, 76(3):297–313.

Naastepad, C. W. M. and Storm, S. (2006-07). Oecd demand regimes (1960-2000). *Journal of Post Keynesian Economics*, 29(2):211–46.

Papadimitriou, D. B., Zezza, G., and Nikiforos, M. (2013). A levy institute model for greece. Technical paper, Levy Institute.

Peersman, G. and Smets, F. (2001). The monetary transmission mechanism in the euro area: more evidence from VAR analysis. Working Paper Series 0091, European Central Bank.

Ratto, M., Roeger, W., and Veld, J. i. t. (2009). Quest iii: An estimated open-economy dsge model of the euro area with fiscal and monetary policy. *Economic Modelling*, 26(1):222–233.

Schoder, C. (2014a). Effective demand, exogenous normal utilization and endogenous capacity in the long run: Evidence from a cointegrated vector autoregression analysis for the usa. *Metroeconomica*, 65(2):298–320.

Schoder, C. (2014b). Instability, stationary utilization and effective demand: A structuralist model of endogenous cycles. *Structural Change and Economic Dynamics*, 30(0):10 – 29.

Schoder, C. (2015a). An estimated dynamic stochastic labor market disequilibrium model for explaining unemployment in the euro area.

Schoder, C. (2015b). A keynesian dynamic stochastic labor market disequilibrium model for business cycle analysis. Working Paper 157, Macroeconomic Policy Institute (IMK), Düsseldorf, Germany.

Schoder, C. (2015c). Methodological, internal and ontological inconsistencies in the conventional micro-foundation of post-keynesian theory.

Shaikh, A. (2009). Economic policy in a growth context: A classical synthesis of keynes and harrod. *Metroeconomica*, 60(3):455–94.

Sims, C. A. (2003). Probability models for monetary policy decisions. Princeton University.

Skott, P. (1989). Effective demand, class struggle and cyclical growth. *International Economic Review*, 30(1):231–47.

Skott, P. and Ryoo, S. (2008). Macroeconomic implications of financialisation. *Cambridge Journal of Economics*, 32(6):827–862.

Skott, P. and Zipperer, B. (2012). An empirical evaluation of three post-Keynesian models. *European Journal of Economics and Economic Policies: Intervention*, 9(2):277–307.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review*, 97(3):pp. 586–606.

Steindl, J. (1952). *Maturity and Stagnation in American Capitalism*. Basil Blackwell, Oxford.

Stockhammer, E. (2005-06). Shareholder value orientation and the investment-profit puzzle. *Cambridge Journal of Economics*, 28:193–215.

Stockhammer, E., Hein, E., and Grafl, L. (2011). Globalization and the effects of changes in functional income distribution on aggregate demand in Germany. *International Review of Applied Economics*, 25(1):1–23.

Stockhammer, E. and Onaran, Ö. (2004). Accumulation, distribution and employment: a structural var approach to a post-keynesian macro model. *Structural Change and Economic Dynamics*, 15:421–47.

Stockhammer, E., Onaran, Ö., and Ederer, S. (2009). Functional income distribution and aggregate demand in the euro area. *Cambridge Journal of Economics*, 33(1):19–40.

Taylor, L. (1985). A stagnationist model of economic growth. *Cambridge Journal of Economics*, 9(4):381–403.

Taylor, L. (2004). *Reconstructing Macroeconomics: Structuralist Proposals and Critiques of the Mainstream.* Harvard University Press, Cambridge.

Taylor, L. (2012). Growth, cycles, asset prices and finance. *Metroeconomica*, 63(1):40–63.

Taylor, L., Proao, C. R., de Carvalho, L., and Barbosa, N. (2012). Fiscal deficits, economic growth and government debt in the USA. *Cambridge Journal of Economics*, 36(1):189–204.

Zipperer, B. and Skott, P. (2010). Cyclical patterns of employment, utilization and profitability. Working Papers 2010-02, University of Massachusetts Amherst, Department of Economics.

A   Computing the steady state

The steady state can be computed from the parameters recursively. It is easy to see from (22)-(28) that, at the steady state, the shock processes are

$$V_g^{ss} = 1$$
$$V_a^{ss} = 1$$
$$V_r^{ss} = 1$$
$$V_c^{ss} = V_c$$
$$V_n^{ss} = 1$$
$$V_i^{ss} = 1$$
$$V_\nu^{ss} = 1$$

and from (29)-(35) that the observed variables are

$$g_y^{ss} = 1$$
$$g_c^{ss} = 1$$
$$g_i^{ss} = 1$$
$$d_r^{ss} = 0$$
$$d_{\pi_p}^{ss} = 0$$
$$d_{\pi_w}^{ss} = 0$$
$$d_u^{ss} = 0.$$

Then, from (4),

$$N_w^{ss} = N_w.$$

Because of (20), price adjustment costs are zero in the steady state, and (7) and (6) imply

$$\varphi^{ss} = \frac{1}{1 + \epsilon}$$

and

$$w^{ss} = \varphi^{ss}(1 - \alpha)\psi^{-\frac{\alpha}{1-\alpha}} = \frac{1}{1 + \epsilon}(1 - \alpha)\psi^{-\frac{\alpha}{1-\alpha}},$$

respectively. To proceed, let us assume for the moment that $Y^{ss} = Y$. In the end, we will show how V_c can be restricted for this condition to hold. With $Y^{ss} = Y$, (5) implies

$$L^{ss} = \psi^{\frac{\alpha}{1-\alpha}}.$$

Then, (3) pins down workers' consumption as

$$C_w^{ss} = \frac{1 - t_w}{1 + t_c}w^{ss}L^{ss}.$$

36

Government spending is implied by (11) as

$$G^{ss} = G.$$

Since N_w is calibrated such that $1 - u = L^{ss}/N^{ss}$, (15) and (18) imply

$$\nu^{ss} = \nu$$

and

$$u^{ss} = u,$$

respectively. Also, (21) implies

$$Y^{e,ss} = Y^{ss}.$$

Because $\frac{I^{ss}}{K^{ss}} = (\Gamma - (1 - \delta))$ due to (9), (8) then requires

$$R^{ss} = R.$$

Eqs. (13), (19), and (20) require

$$\Pi_p^{ss} = \Pi_p,$$
$$\Pi_w^{ss} = \Pi_p$$

and

$$\Pi_p^{e,ss} = \Pi_p,$$

respectively. Note that we require the parameter restriction that $\Pi_p = \Pi_w$. With $Y^{ss} = Y$, and the assumption that firms operate with a fixed output-capital ratio ψ in the long run, we know

$$K^{ss} = \frac{Y^{ss}}{\psi}.$$

Eq. (9) then implies for investment

$$I^{ss} = (\Gamma - (1 - \delta)) K^{ss}.$$

which combined with the steady state values already known allow us to obtain the rentier's income through (17)

$$Z_r^{ss} = Y^{ss} - \omega^{ss} L^{ss} - (1 - \xi^{ss}) I^{ss}.$$

Rentier's consumption can now be obtained from the macroeconomic balance condition in (16) as

$$C_r^{ss} = Y^{ss} - C_w^{ss} - I^{ss} - G^{ss}.$$

Eq. (2) implies

$$B_r^{ss} = \frac{\frac{R^{ss}}{\Pi_p^{ss}} \frac{1}{\Gamma}}{1 - \frac{R^{ss}}{\Pi_p^{ss}} \frac{1}{\Gamma}} \left((1 - t_z)Z_r^{ss} - C_r^{ss}\right).$$

Taxes are determined by (10) as

$$T^{ss} = t_w \omega^{ss} L^{ss} + t_z Z_r^{ss} + t_c (C_r^{ss} + C_w^{ss}).$$

and the steady-state government debt by (12) as

$$D_g^{ss} = \frac{\frac{R^{ss}}{\Pi_p^{ss}} \frac{1}{\Gamma}}{1 - \frac{R^{ss}}{\Pi_p^{ss}} \frac{1}{\Gamma}} \left(G^{ss} - T^{ss}\right).$$

Note that we calibrate t_w such that $D_g^{ss} = 2.4$. Finally, note that we have assumed $Y^{ss} = Y$ which is achieved by restricting V_c such that the macroeconomic balance condition (16) after substituting in the consumption function (1) as well as all steady state values implies that $Y = ((1 - t_z)Z_r^{ss})^{\phi_c} (B_r^{ss})^{1 - \phi_c} V_c + C_w^{ss} + I^{ss} + G^{ss}$.

B   Log-linear approximations around the steady state

A log-linear approximation of the non-linear model is performed by *Dynare*. Here, we provide a manual derivation. For each variable X_t, we define $\hat{x}_t = \log X_t - \log X^{ss}$. Then we can write $X_t = X^{ss}\exp\{\hat{x}_t\} \approx X^{ss}(1 + \hat{x}_t)$.

- Rentier's consumption function (1):

$$\frac{C_{r,t}}{B_r} = \left((1 - t_z)\frac{Z_{r,t}}{B_r}\right)^{\phi_c} V_{c,t}$$

$$\frac{C_r^{ss}}{B_r}\exp\{\hat{c}_{r,t}\} = \left((1 - t_z)\frac{Z_r}{B_r}\right)^{\phi_c} V^{ss}\exp\{\phi_c(1 - t_z)\hat{z}_{r,t} + \hat{v}_t\}$$

$$\frac{C_r^{ss}}{B_r}(1 + \hat{c}_{r,t}) \approx \frac{C_r^{ss}}{B_r}\left(1 + \phi_c(1 - t_z)\hat{z}_{r,t} + \hat{v}_t\right)$$

$$\hat{c}_{r,t} \approx \phi_c(1 - t_z)\hat{z}_{c,t} + \hat{v}_t$$

- Rentier's budget constraint (2):

$$B_{r,t} = \frac{R_t}{\Pi_{p,t+1}}\frac{1}{\Gamma}\left(B_{r,t-1} + (1 - t_z)Z_{r,t} - (1 + t_c)C_{r,t}\right)$$

$$\Gamma\frac{\Pi_p^{ss}}{R^{ss}}B_r^{ss}\exp\{\hat{\pi}_{p,t+1} - \hat{r}_t + \hat{b}_t\} = B_r^{ss}\exp\{\hat{b}_{r,t-1}\} + (1 - t_z)Z_r^{ss}\exp\{\hat{z}_{r,t}\} -$$
$$- (1 + t_c)C_r^{ss}\exp\{\hat{c}_{r,t}\}$$

$$\Gamma\frac{\Pi_p^{ss}}{R^{ss}}B_r^{ss}(1 + \hat{\pi}_{p,t+1} - \hat{r}_t + \hat{b}_{r,t}) \approx B_r^{ss}(1 + \hat{b}_{r,t-1}) + (1 - t_z)Z_r^{ss}(1 + \hat{z}_{r,t}) -$$
$$- (1 + t_c)C_r^{ss}(1 + \hat{c}_{r,t})$$

$$\Gamma\frac{\Pi_p^{ss}}{R^{ss}}B_r^{ss}(\hat{\pi}_{p,t+1} - \hat{r}_t + \hat{b}_{r,t}) \approx B_r^{ss}\hat{b}_{r,t-1} + (1 - t_z)Z_r^{ss}\hat{z}_{r,t} - (1 + t_c)C_r^{ss}\hat{c}_{r,t}$$

$$\Gamma\frac{\Pi_p^{ss}}{R^{ss}}(\hat{\pi}_{p,t+1} - \hat{r}_t + \hat{b}_{r,t}) \approx \hat{b}_{r,t-1} + (1 - t_z)\frac{Z_r^{ss}}{B_r^{ss}}\hat{z}_{r,t} - (1 + t_c)\frac{C_r^{ss}}{B_r^{ss}}\hat{c}_{r,t}$$

$$\hat{b}_t \approx \hat{r}_t - \hat{\pi}_{p,t+1} +$$
$$+ \frac{R^{ss}}{\Pi_p^{ss}}\frac{1}{\Gamma}\left(\hat{b}_{r,t-1} + (1 - t_z)\frac{Z_r^{ss}}{B_r^{ss}}\hat{z}_{r,t} - (1 + t_c)\frac{C_r^{ss}}{B_r^{ss}}\hat{c}_{r,t}\right)$$

- Worker's consumption (3):

$$(1 + t_c)C_{w,t} = (1 - t_w)\omega_t L_t$$
$$(1 + t_c)C_w^{ss}\exp\{\hat{c}_{w,t}\} = (1 - t_w)\omega^{ss}L^{ss}\exp\{\hat{\omega}_t + \hat{l}_t\}$$
$$(1 + t_c)C_w^{ss}(1 + \hat{c}_{w,t}) \approx (1 - t_w)\omega^{ss}L^{ss}(1 + \hat{\omega}_t + \hat{l}_t)$$
$$\hat{c}_{w,t} \approx \hat{\omega}_t + \hat{l}_t$$

- Labor supply (4):

$$\frac{N_{w,t}}{N_w} = V_{n,t}$$

$$\frac{1}{N_w} N_w^{ss} \exp\{\hat{n}_{w,t}\} = V_n^{ss} \exp\{\hat{v}_{n,t}\}$$

$$\frac{1}{N_w} N_w^{ss}(1 + \hat{n}_{w,t}) \approx V_n^{ss}(1 + \hat{v}_{n,t})$$

$$\hat{n}_{w,t} \approx \hat{v}_{n,t}$$

- Labor demand (5):

$$L_t = \frac{Y_t}{V_{a,t}} \psi^{\frac{\alpha}{1-\alpha}}$$

$$L^{ss} \exp\{\hat{l}_t\} = \psi^{\frac{\alpha}{1-\alpha}} \frac{Y^{ss}}{V_a^{ss}} \exp\{\hat{y}_t - \hat{v}_{a,t}\}$$

$$L^{ss}(1 + \hat{l}_t) \approx \psi^{\frac{\alpha}{1-\alpha}} \frac{Y^{ss}}{V_a^{ss}}(1 + \hat{y}_t - \hat{v}_{a,t})$$

$$\hat{l}_t \approx \hat{y}_t - \hat{v}_{a,t}$$

- Unit costs (6):

$$\varphi_t = \frac{\omega_t}{V_{a,t}} \frac{1}{1-\alpha} \left(\frac{\psi}{V_{a,t}}\right)^{\frac{\alpha}{1-\alpha}}$$

$$\varphi_t = \frac{1}{1-\alpha} \psi^{\frac{\alpha}{1-\alpha}} \omega_t (V_{a,t})^{\frac{-1}{1-\alpha}}$$

$$\varphi^{ss} \exp\{\hat{\varphi}_t\} = \frac{1}{1-\alpha} \psi^{\frac{\alpha}{1-\alpha}} \omega^{ss} (V_a^{ss})^{\frac{-1}{1-\alpha}} \exp\left\{\hat{\omega}_t - \frac{1}{1-\alpha}\hat{v}_{a,t}\right\}$$

$$\varphi^{ss}(1 + \hat{\varphi}_t) \approx \frac{1}{1-\alpha} \psi^{\frac{\alpha}{1-\alpha}} \omega^{ss} (V_a^{ss})^{\frac{-1}{1-\alpha}} \left(\hat{\omega}_t - \frac{1}{1-\alpha}\hat{v}_{a,t}\right)$$

$$\hat{\varphi}_t \approx \hat{\omega}_t - \frac{1}{1-\alpha}\hat{v}_{a,t}$$

- Price formation (7):

$$(1 - \varphi_t) \frac{1}{\epsilon} \exp\{\sigma_\epsilon \varepsilon_{\epsilon,t}\} - \varphi_t = \tau \left(\Pi^e_{p,t} - \Pi_{p,t} \right)$$

$$\frac{1}{\epsilon} \exp\{\sigma_\epsilon \varepsilon_{\epsilon,t}\} - \frac{1}{\epsilon} \exp\{\sigma_\epsilon \varepsilon_{\epsilon,t}\} \varphi_t - \varphi_t = \tau \left(\Pi^e_{p,t} - \Pi_{p,t} \right)$$

$$\left[\begin{matrix} \frac{1}{\epsilon} \exp\{\sigma_\epsilon \varepsilon_{\epsilon,t}\} - \frac{1}{\epsilon} \exp\{\sigma_\epsilon \varepsilon_{\epsilon,t}\} \varphi^{ss} \exp\{\hat{\varphi}_t\} - \\ -\varphi^{ss} \exp\{\hat{\varphi}_t\} \end{matrix} \right] = \tau \left(\Pi^{e,ss}_p \exp\{\hat{\pi}^e_{p,t}\} - \Pi^{ss}_p \exp\{\hat{\pi}_{p,t}\} \right)$$

$$\left[\begin{matrix} \frac{1}{\epsilon}(1 + \sigma_\epsilon \varepsilon_{\epsilon,t}) - \frac{1}{\epsilon}(1 + \sigma_\epsilon \varepsilon_{\epsilon,t}) \varphi^{ss}(1 + \hat{\varphi}_t) - \\ -\varphi^{ss}(1 + \hat{\varphi}_t) \end{matrix} \right] \approx \tau \left(\Pi^{e,ss}_p(1 + \hat{\pi}^e_{p,t}) - \Pi^{ss}_p(1 + \hat{\pi}_{p,t}) \right)$$

$$\left[\begin{matrix} 1\epsilon + \frac{1}{\epsilon}\sigma_\epsilon \varepsilon_{\epsilon,t} - \frac{1}{\epsilon}\varphi^{ss} + \frac{1}{\epsilon}\varphi^{ss}\hat{\varphi}_t - \frac{1}{\epsilon}\sigma_\epsilon \varepsilon_{\epsilon,t}\varphi^{ss} + \\ +\frac{1}{\epsilon}\sigma_\epsilon \varepsilon_{\epsilon,t}\varphi^{ss}\hat{\varphi}_t - \varphi^{ss} - \varphi^{ss}\hat{\varphi}_t \end{matrix} \right] \approx \tau \Pi^{ss}_p \left(\hat{\pi}^e_{p,t} - \hat{\pi}_{p,t} \right)$$

$$\frac{1}{\epsilon}\sigma_\epsilon \varepsilon_{\epsilon,t} + \frac{1}{\epsilon}\varphi^{ss}\hat{\varphi}_t - \frac{1}{\epsilon}\sigma_\epsilon \varepsilon_{\epsilon,t}\varphi^{ss} - \varphi^{ss}\hat{\varphi}_t \approx \tau \Pi^{ss}_p \left(\hat{\pi}^e_{p,t} - \hat{\pi}_{p,t} \right)$$

- Investment (8):

$$\frac{I_t}{K} = (\Gamma - (1 - \delta)) \left(\frac{Y^e_t}{Y} \right)^{\phi_{iy}} \left(\frac{R_t}{R} \right)^{-\phi_{ir}} V_{i,t}$$

$$\frac{I^{ss}}{K} \exp\{\hat{i}_t\} = (\Gamma - (1 - \delta)) \left(\frac{Y^{e,ss}}{Y} \right)^{\phi_{iy}} \left(\frac{R^{ss}}{R} \right)^{-\phi_{ir}} V^{ss}_i \exp\{\phi_{iy}\hat{y}^e_t - \phi_{ir}\hat{r}_t + \hat{v}_{i,t}\}$$

$$\frac{I^{ss}}{K}(1 + \hat{i}_t) \approx (\Gamma - (1 - \delta)) \left(\frac{Y^{e,ss}}{Y} \right)^{\phi_{iy}} \left(\frac{R^{ss}}{R} \right)^{-\phi_{ir}} V^{ss}_i(1 + \phi_{iy}\hat{y}^e_t - \phi_{ir}\hat{r}_t + \hat{v}_{i,t})$$

$$\hat{i}_t \approx \phi_{iy}\hat{y}^e_t - \phi_{ir}\hat{r}_t + \hat{v}_{i,t}$$

- Law of motion of capital (9):

$$\Gamma K_t = I_t + (1 - \delta) K_{t-1}$$

$$\Gamma K^{ss} \exp\{\hat{k}_t\} = I^{ss} \exp\{\hat{i}_t\} + (1 - \delta) K^{ss} \exp\{\hat{k}_{t-1}\}$$

$$\Gamma K^{ss}(1 + \hat{k}_t) \approx I^{ss}(1 + \hat{i}_t) + (1 - \delta) K^{ss}(1 + \hat{k}_{t-1})$$

$$\Gamma \hat{k}_t \approx \frac{I^{ss}}{K^{ss}}\hat{i}_t + (1 - \delta) \hat{k}_{t-1}$$

- Taxes (10):

$$T_t = t_w \omega_t L_t + t_z Z_{r,t} + t_c(C_{r,t} + C_{w,t})$$

$$T^{ss}\hat{t}_t \approx t_w \omega^{ss} L^{ss}(\hat{\omega}_t + \hat{l}_t) + t_z Z^{ss}_r \hat{z}_t + t_c(C^{ss}_r \hat{c}_{r,t} + C^{ss}_w \hat{c}_{w,t})$$

- Government spending (11):

$$\frac{G_t}{G} = \left(\frac{Y_t}{Y} \right)^{\phi_{gy}} V_{g,t}$$

$$\hat{g}_t \approx \phi_{gy}\hat{y}_t + \hat{v}_{g,t}$$

- Law of motion of government debt (12):

$$D_{g,t} = \frac{R_t}{\Pi_{p,t+1}} \frac{1}{\Gamma} (D_{g,t-1} + G_t - T_t)$$

$$\Gamma \frac{\Pi_p^{ss}}{R^{ss}} D_g^{ss}(1 + \hat{\pi}_{p,t+1} - \hat{r}_t + \hat{d}_{g,t}) \approx D_g^{ss}(1 + \hat{d}_{g,t-1}) + G^{ss}(1 + \hat{g}_t) - $$
$$ - T^{ss}(1 + \hat{t}_t)$$

$$\Gamma \frac{\Pi_p^{ss}}{R^{ss}} D_g^{ss}(\hat{\pi}_{p,t+1} - \hat{r}_t + \hat{d}_{g,t}) \approx D_g^{ss} \hat{d}_{g,t-1} + G^{ss}\hat{g}_t - T^{ss}\hat{t}_t$$

$$\Gamma \frac{\Pi_p^{ss}}{R^{ss}} (\hat{\pi}_{p,t+1} - \hat{r}_t + \hat{d}_{g,t}) \approx \hat{d}_{g,t-1} + \frac{G^{ss}}{D_g^{ss}}\hat{g}_t - \frac{T^{ss}}{D_g^{ss}}\hat{t}_t$$

$$\hat{d}_{g,t} \approx \hat{r}_t - \hat{\pi}_{p,t+1} + $$
$$ + \frac{R^{ss}}{\Pi_p^{ss}} \frac{1}{\Gamma} \left(\hat{d}_{g,t-1} + \frac{G^{ss}}{D_g^{ss}}\hat{g}_t - \frac{T^{ss}}{D_g^{ss}}\hat{t}_t \right)$$

- Interest rate rule (13):

$$\frac{R_t}{R} = \left(\frac{\Pi_{p,t}}{\Pi_p} \right)^{\phi_{r\pi}} \left(\frac{Y_t}{Y} \right)^{\phi_{ry}} V_{r,t}$$

$$\hat{r}_t \approx \phi_{r\pi}\hat{\pi}_{p,t} + \phi_{ry}\hat{y}_t + \hat{v}_{r,t}$$

- Wage bargaining solution (14):

$$1 = \left(1 - \frac{1}{\nu_t} \right) \frac{\omega(\Pi_{w,t})}{r(\Pi_{w,t})} \frac{r'(\Pi_{w,t})}{\omega'(\Pi_{w,t})}$$

where $\omega(\cdot)$ and $r(\cdot)$ are of the from $\omega(\cdot) = A_\omega + B_\omega \Pi_{w,t}$ and $r(\cdot) = A_r + B_r\Pi_{w,t}$, respectively, with A_ω, B_ω, A_r, and B_r being parameter convolutes. $\omega'(\cdot)$ and $r'(\cdot)$ are constant. Then, log-linearization yields

$$A_r + B_r\Pi_{w,t} = \frac{r'(\cdot)}{\omega'(\cdot)}A_\omega + \frac{r'(\cdot)}{\omega'(\cdot)}B_\omega\Pi_{w,t} - \frac{r'(\cdot)}{\omega'(\cdot)}A_\omega\frac{1}{\nu_t} + \frac{r'(\cdot)}{\omega'(\cdot)}B_\omega\frac{1}{\nu_t}\Pi_{w,t}$$

$$A_r + B_r\Pi_w^{ss}(1 + \hat{\pi}_{w,t}) \approx \frac{r'(\cdot)}{\omega'(\cdot)}A_\omega + \frac{r'(\cdot)}{\omega'(\cdot)}B_\omega\Pi_w^{ss}(1 + \hat{\pi}_{w,t}) - \frac{r'(\cdot)}{\omega'(\cdot)}A_\omega\frac{1}{\nu^{ss}}(1 + \hat{\nu}_t) + $$
$$ + \frac{r'(\cdot)}{\omega'(\cdot)}B_\omega\frac{1}{\nu^{ss}}\Pi_w^{ss}(1 - \hat{\nu}_t + \hat{\pi}_{w,t})$$

$$B_r\Pi_w^{ss}\hat{\pi}_{w,t} \approx \frac{r'(\cdot)}{\omega'(\cdot)} \left(B_\omega\Pi_w^{ss}\hat{\pi}_{w,t} - A_\omega\frac{1}{\nu^{ss}}\hat{\nu}_t - B_\omega\frac{1}{\nu^{ss}}\Pi_w^{ss}(\hat{\nu}_t - \hat{\pi}_{w,t}) \right)$$

- Bargaining power (15):

$$\frac{\nu_t}{\nu} = \left(\frac{L_t}{L} \right)^{\phi_{\nu l}} V_{\nu,t}$$

$$\hat{\nu}_t \approx \phi_{\nu l}\hat{l}_t + \hat{v}_{\nu,t}$$

- Macroeconomic balance condition (16):

$$Y_t = C_{r,t} + C_{w,t} + I_t + G_t + \frac{\tau}{2}\left(\Pi_{p,t} - \Pi_p\right)^2$$

$$Y^{ss}(1 + \hat{y}_t) \approx C_r^{ss}(1 + \hat{c}_{r,t}) + C_w^{ss}(1 + \hat{c}_{w,t}) + I^{ss}(1 + \hat{i}_t) + G^{ss}(1 + \hat{g}_t) +$$
$$+ \frac{\tau}{2}\left((\Pi_p^{ss})^2(1 + 2\hat{\pi}_{p,t}) - 2\Pi_p\Pi_p^{ss}(1 + \hat{\pi}_{p,t}) + (\Pi_p)^2\right)$$

$$Y^{ss}(1 + \hat{y}_t) \approx C_r^{ss}(1 + \hat{c}_{r,t}) + C_w^{ss}(1 + \hat{c}_{w,t}) + I^{ss}(1 + \hat{i}_t) + G^{ss}(1 + \hat{g}_t)$$

$$\hat{y}_t \approx \frac{C_r^{ss}}{Y^{ss}}\hat{c}_{r,t} + \frac{C_w^{ss}}{Y^{ss}}\hat{c}_{w,t} + \frac{I^{ss}}{Y^{ss}}\hat{i}_t + \frac{G^{ss}}{Y^{ss}}\hat{g}_t$$

- Rentier's income (17):

$$Z_{r,t} = Y_t - \omega_t L_t - (1 - \xi_t) I_t - \frac{\tau}{2}\left(\Pi_{p,t} - \Pi_p\right)^2$$

$$Z_{r,t} = Y_t - \omega_t L_t - I_t + \xi_t I_t - \frac{\tau}{2}\left((\Pi_{p,t})^2 - 2\Pi_{p,t}\Pi_p + (\Pi_p)^2\right)$$

$$Z_r^{ss}(1 + \hat{z}_{r,t}) \approx Y^{ss}(1 + \hat{y}_t) - \omega^{ss}L^{ss}(1 + \hat{\omega}_t + \hat{l}_t) - I^{ss}(1 + \hat{i}_t) + \xi^{ss}I^{ss}(1 + \hat{\xi}_t + \hat{i}_t) -$$
$$- \frac{\tau}{2}\left((\Pi_p^{ss})^2(1 + 2\hat{\pi}_{p,t}) - 2\Pi_p\Pi_p^{ss}(1 + \hat{\pi}_{p,t}) + (\Pi_p)^2\right)$$

$$Z_r^{ss}\hat{z}_{r,t} \approx Y^{ss}\hat{y}_t - \omega^{ss}L^{ss}(\hat{\omega}_t + \hat{l}_t) - I^{ss}\hat{i}_t + \xi^{ss}I^{ss}(\hat{\xi}_t + \hat{i}_t)$$

$$\hat{z}_{r,t} \approx \frac{Y^{ss}}{Z_r^{ss}}\hat{y}_t - \frac{\omega^{ss}L^{ss}}{Z_r^{ss}}(\hat{\omega}_t + \hat{l}_t) - \frac{I^{ss}}{Z_r^{ss}}\hat{i}_t + \frac{\xi^{ss}I^{ss}}{Z_r^{ss}}(\hat{\xi}_t + \hat{i}_t)$$

- Unemployment rate (18):

$$u_t = 1 - \frac{L_t}{N_{w,t}}$$

$$u^{ss}(1 + \hat{u}_t) \approx 1 - \frac{L^{ss}}{N^{ss}}(1 + \hat{l}_t - \hat{n}_t)$$

$$u^{ss}\hat{u}_t \approx (1 - u^{ss})(\hat{l}_t - \hat{n}_t)$$

$$\hat{u}_t \approx \frac{1 - u^{ss}}{u^{ss}}(\hat{l}_t - \hat{n}_t)$$

- Law of motion of the real wage (19):

$$\frac{\omega_t}{\omega_{t-1}} - 1 = \Pi_{w,t} - \Pi_{p,t}$$

$$\frac{\omega^{ss}}{\omega^{ss}}(1 + \hat{\omega}_t - \hat{\omega}_{t-1}) - 1 \approx \Pi_p^{ss}(\hat{\pi}_{w,t} - \hat{\pi}_{p,t})$$

$$\hat{\omega}_t - \hat{\omega}_{t-1} \approx \Pi_p^{ss}(\hat{\pi}_{w,t} - \hat{\pi}_{p,t})$$

- Price expectations (20):

$$\frac{\Pi_{p,t}^e}{\Pi_{p,t}} = \left(\frac{\Pi_{p,t-1}^e}{\Pi_{p,t}}\right)^{\rho_{\pi e}}$$

$$\hat{\pi}_t^e \approx \rho_{\pi e}\hat{\pi}_{t-1}^e + (1 - \rho_{\pi e})\hat{\pi}_t$$

- Output expectations (21):

$$\frac{Y_t^e}{Y_t} = \left(\frac{Y_{t-1}^e}{Y_t}\right)^{\rho_{ye}}$$

$$\hat{y}_t^e \approx \rho_{ye}\hat{y}_{t-1}^e + (1 - \rho_{ye})\hat{y}_t$$

- Shock processes (22)-(28):

$$\hat{v}_{g,t} \approx \rho_g\hat{v}_{g,t-1} + \sigma_g\varepsilon_{g,t},$$

$$\hat{v}_{a,t} \approx \rho_a\hat{v}_{a,t-1} + \sigma_a\varepsilon_{a,t},$$

$$\hat{v}_{r,t} \approx \rho_r\hat{v}_{r,t-1} + \sigma_r\varepsilon_{r,t},$$

$$\hat{v}_{c,t} \approx \rho_c\hat{v}_{c,t-1} + \sigma_c\varepsilon_{c,t},$$

$$\hat{v}_{n,t} \approx \rho_n\hat{v}_{n,t-1} + \sigma_n\varepsilon_{n,t},$$

$$\hat{v}_{i,t} \approx \rho_i\hat{v}_{i,t-1} + \sigma_i\varepsilon_{i,t},$$

$$\hat{v}_{\nu,t} \approx \rho_\nu\hat{v}_{\nu,t-1} + \sigma_\nu\varepsilon_{\nu,t},$$

- Measurement equations (29)-(35): Note that the observed variables have been specified as steady-state deviations in percent in the first place. Hence,

$$d_{y,t} = 100\hat{y}_t$$

$$d_{c,t} = 100\left(\frac{C_r^{ss}}{C_r^{ss} + C_w^{ss}}\hat{c}_{r,t} + \frac{C_w^{ss}}{C_r^{ss} + C_w^{ss}}\hat{c}_{w,t}\right)$$

$$d_{i,t} = 100\hat{i}_t$$

$$d_{r,t} = 100\hat{r}_t$$

$$d_{\pi p,t} = 100\hat{\pi}_{p,t}$$

$$d_{\pi w,t} = 100\hat{\pi}_{w,t}$$

$$d_{u,t} = 100\hat{u}_t$$

C   Computing the likelihood using the Kalman filter

Basic principles of probability theory imply that the joint probability density of the entire data set $\mathcal{M}_T$ conditional on $\boldsymbol{\Theta}$ is simply the product of the probability densities of the observations $\mathbf{M}_t$ in each single period t conditional on all previous observations up to $t-1$, i.e. $\mathcal{M}_{t-1}$, and $\boldsymbol{\Theta}$. Hence,

$$\mathcal{L}(\mathcal{M}_T|\boldsymbol{\Theta}) = \prod_{t=1}^{T} L(\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta}) \tag{38}$$

What is the likelihood of $\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta}$? Due to the assumption of normally and independently distributed innovations with zero mean and unity variance and the fact that our model has been linearized, $\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta}$ obeys a normal distribution

$$\mathcal{L}(\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta}) = (2\pi)^{-\frac{m}{2}} \left|\boldsymbol{\Omega}_{t|t-1}\right|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)'\boldsymbol{\Omega}_{t|t-1}^{-1}\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)\right\}$$

with mean vector

$$\mathbf{M}_{t|t-1} \equiv \mathrm{E}\big[\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta}\big]$$

and covariance matrix

$$\boldsymbol{\Omega}_{t|t-1} \equiv \mathrm{E}\big[\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)'\big].$$

If all model variables were observed, then $\mathbf{M}_t|\mathcal{M}_{t-1},\boldsymbol{\Theta} \sim n.i.d.\,(\mathbf{FM}_{t-1},\mathbf{QQ}')$. This is because, in this case, $\mathbf{M}_t = \mathbf{S}_t$ and the mean vector would simply be

$$\begin{aligned}\mathbf{M}_{t|t-1} &= \mathrm{E}\big[\mathbf{S}_t|\mathcal{S}_{t-1},\boldsymbol{\Theta}\big] \\ &= \mathbf{FS}_{t-1} \\ &= \mathbf{FM}_{t-1}\end{aligned}$$

and the covariance matrix

$$\begin{aligned}\boldsymbol{\Omega}_{t|t-1} &= \mathrm{E}\big[\left(\mathbf{M}_t - \mathbf{FM}_{t-1}\right)\left(\mathbf{M}_t - \mathbf{FM}_{t-1}\right)'\big] \\ &= \mathrm{E}\big[\mathbf{Q}\varepsilon_t\varepsilon_t'\mathbf{Q}'\big] \\ &= \mathbf{QQ}'.\end{aligned}$$

The likelihood function could be computed very easily. Unfortunately, only a subset of model variables is observed. Hence the expected $\mathbf{M}_{t|t-1}$ depends on the current state which is not observed. To form the expectation $\mathbf{M}_{t|t-1}$ we first need to compute an unbiased expectation of the current state $\mathbf{S}_{t|t-1} \equiv \mathrm{E}\big[\mathbf{S}_t|\mathcal{S}_{t-1},\boldsymbol{\Theta}\big]$ as we do not observe it. Since the state has to be guessed, some uncertainty around this guess arises which is measured by the covariance matrix $\mathbf{P}_{t|t-1} \equiv \mathrm{E}\big[\left(\mathbf{S}_t - \mathbf{S}_{t|t-1}\right)\left(\mathbf{S}_t - \mathbf{S}_{t|t-1}\right)'\big]$ which has to be estimated as well. Obviously,

we want to choose a method of guessing the state which implies the uncertainty, i.e. the covariance matrix, to be as low as possible. Once the state $\mathbf{S}_{t|t-1}$ has been estimated we use the measurement equation to infer $\mathbf{M}_{t|t-1}$. The variance of this estimate will now be larger compared to the previous case of all variables observed because it needs to take into account the uncertainty around $\mathbf{S}_{t|t-1}$ as well.

The Kalman filter takes up the job of generating recursively expectations of $\mathbf{S}_{t|t-1}$ and $\mathbf{M}_{t|t-1}$ required to eventually generate the likelihood function of the entire data set (which the Maximum Likelihood estimator then maximizes by choosing the optimal $\mathbf{\Theta}_e$ given $\mathbf{\Theta}_c$). These expectations are optimal within the class of linear models if innovations are Gaussian as assumed in our model.

The Kalman filter algorithm proceeds as follows: Initialize the expectations of the first state as

$$\mathbf{S}_{1|0} = 0.$$

Substituting this into the definition of the covariance matrix and using the state equation yields

$$
\begin{aligned}
\mathbf{P}_{1|0} &= \mathrm{E}\big[\mathbf{S}_1 \mathbf{S}_1'\big] \\
&= \mathrm{E}\big[\big(\mathbf{F}\mathbf{S}_0 + \mathbf{Q}\varepsilon_0\big)\big(\mathbf{F}\mathbf{S}_0 + \mathbf{Q}\varepsilon_0\big)'\big] \\
&= \mathbf{F}\mathbf{P}_{0|0}\mathbf{F}' + \mathbf{Q}\mathbf{Q}'
\end{aligned}
$$

where $\mathbf{P}_{0|0}$ is the unconditional variance of $\mathbf{S}$ which can be obtained by solving $\mathbf{P}_{0|0} = \mathbf{F}\mathbf{P}_{0|0}\mathbf{F}' + \mathbf{Q}\mathbf{Q}'$. Using the measurement equation, we get for the expected observables of the first period

$$\mathbf{M}_{1|0} = \mathbf{H}\mathbf{S}_{1|0} = 0$$

and for its covariance

$$
\begin{aligned}
\mathbf{\Omega}_{1|0} &= \mathrm{E}\big[\big(\mathbf{M}_1 - \mathbf{M}_{1|0}\big)\big(\mathbf{M}_1 - \mathbf{M}_{1|0}\big)'\big] \\
&= \mathbf{H}'\mathbf{P}_{1|0}\mathbf{H}
\end{aligned}
$$

We now have all the ingredients to compute the likelihood function for the first period as

$$
\begin{aligned}
\mathcal{L}(\mathbf{M}_1|\mathbf{\Theta}) &= (2\pi)^{-\frac{m}{2}} \left|\mathbf{\Omega}_{1|0}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\big(\mathbf{M}_1 - \mathbf{M}_{1|0}\big)'\mathbf{\Omega}_{1|0}^{-1}\big(\mathbf{M}_1 - \mathbf{M}_{1|0}\big)\right\} \\
&= (2\pi)^{-\frac{m}{2}} \left|\mathbf{\Omega}_{1|0}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{M}_1'\mathbf{\Omega}_{1|0}^{-1}\mathbf{M}_1\right\}
\end{aligned}
$$

Before entering the next step, the Kalman filter updates the expectations $\mathbf{S}_{1|0}$ made in period 0 and the covariance $\mathbf{P}_{1|0}$ of these expectations to expectations $\mathbf{S}_{1|1}$ made in period 1 and their

covariance $\mathbf{P}_{1|1}$ by using the information contained in the set of observations $\mathbf{M}_1$ of period 1. The expectation regarding the state $\mathbf{S}_{1|0}$ and its covariance $\mathbf{P}_{1|0}$ are updated in order to take into account a possible deviation of the observation $\mathbf{M}_1$ from its expectation (or forecast) $\mathbf{M}_{1|0}$, i.e. a possible forecast error, which, given the dependence of the expectation of the observation $\mathbf{M}_{1|0}$ on the expected state $\mathbf{S}_{1|0}$ through the measurement equation, indicates that the previous expectation regarding the state $\mathbf{S}_{1|0}$ was wrong. In particular,

$$\mathbf{S}_{1|1} = \mathbf{S}_{1|0} + \mathbf{P}_{1|0}\mathbf{H}\mathbf{\Omega}_{1|0}^{-1}(\mathbf{M}_1 - \mathbf{M}_{1|0})$$

and

$$\mathbf{P}_{1|1} = \mathbf{P}_{1|0} - \mathbf{P}_{1|0}\mathbf{\Omega}_{1|0}^{-1}\mathbf{H}\mathbf{H}'\mathbf{P}_{1|0}$$

The intuitions of these updating rules are as follows: Given the proportional relationship between $\mathbf{M}_t$ and $\mathbf{S}_t$ through the measurement equation, the former equation states that a positive (negative) forecasting error of the observed variables implies that the expectation of the state was too low (high). Hence, it should be revised upwards (downwards). The extent by which we should update our expectation of the state depends proportionally on how uncertain we were about the expected state before the data became available, i.e. $\mathbf{P}_{1|0}$, and inversely on how uncertain we were about the expected observation before it was actually made, i.e. $\mathbf{\Omega}_{1|0}$. The latter is because a given forecast error should not induce us to revise the expectation of the state too much if we were very uncertain about our initial expectation of the data in the first place, as the realized state (which gave rise to the data) may well have been hit by an unfortunate idiosyncratic shock. Then, this may have been the reason for the data being rather off from what has been expected rather than a wrong expectation of the state. The latter equation states that the uncertainty $\mathbf{P}_{1|1}$ regarding the updated expectation of the state needs to become smaller than the uncertainty $\mathbf{P}_{1|0}$ of the initial expectation of the state, as the information set increased due to observing $\mathbf{M}_1$. By how much depends on the covariance of the forecast errors $\mathbf{\Omega}_{1|0}$. If the expectation of the data was rather off and, hence, the forecast errors large, then we cannot be sure that the updated expectation of the state is very accurate. The information contained in the new observation is weak and does not allow us to reduce the uncertainty regarding the expected state by much. If, however, the forecast errors were low the expected state seem to explain the data well and we can confidently reduce the uncertainty regarding the expected state.

The recursion can now start. For periods $t = 2, \ldots, T$, generate the expectations or forecasts for

$$\mathbf{S}_{t|t-1} = \mathbf{F}\mathbf{S}_{t-1|t-1},$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}\mathbf{P}_{t-1|t-1}\mathbf{F}' + \mathbf{Q}\mathbf{Q}',$$

$$\mathbf{M}_{t|t-1} = \mathbf{H}\mathbf{S}_{t|t-1},$$

$$\mathbf{\Omega}_{t|t-1} = \mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H},$$

use them to compute the likelihood function for observation $\mathbf{M}_t$ conditional on all previous observations and the model parameters,

$$\mathcal{L}(\mathbf{M}_t|\mathcal{M}_{t-1},\mathbf{\Theta}) = (2\pi)^{-\frac{m}{2}} \left|\mathbf{\Omega}_{t|t-1}\right|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)'\mathbf{\Omega}_{t|t-1}^{-1}\left(\mathbf{M}_t - \mathbf{M}_{t|t-1}\right)\right\}$$

and update the expectation of the state and its covariance matrix,

$$\mathbf{S}_{t|t} = \mathbf{S}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{H}\mathbf{\Omega}_{t|t-1}^{-1}(\mathbf{M}_t - \mathbf{M}_{t|t-1}),$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{\Omega}_{t|t-1}^{-1}\mathbf{H}\mathbf{H}'\mathbf{P}_{t|t-1}.$$

The likelihoods $\mathcal{L}(\mathbf{M}_1|\mathbf{\Theta}), \mathcal{L}(\mathbf{M}_2|\mathcal{M}_1,\mathbf{\Theta}), \ldots, \mathcal{L}(\mathbf{M}_T|\mathcal{M}_{T-1},\mathbf{\Theta})$ can then substituted into (38) to obtain the likelihood of the data $\mathcal{L}(\mathcal{M}_T|\mathbf{\Theta})$.

D   The MH-MCMC algorithm

The MH-MCMC algorithm generates a series of parameter vectors $\{\hat{\boldsymbol{\Theta}}_{e,l}\}_{l=1}^{H}$ where H is the number of iterations. The algorithm implies that the simulated series $\{\hat{\theta}_{i,e,l}\}_{l=1}^{H}$ is drawn from the marginal posterior of interest $\mathcal{P}(\hat{\theta}_{i,e}|\mathcal{M}_T, \boldsymbol{\Theta}_c)$.

The MH-MCMC simulator proceeds as follows: Choose a $\hat{\boldsymbol{\Theta}}_{e,0}$ to initialize the MH-MCMC simulator and set $c = 0$ which will be used as a counter for how often the decision rule accepts. Then, for all iterations $l = 1, \ldots, H$ do the following: Using the random walk law of motion, update the parameter vector to

$$\boldsymbol{\Theta}_{e,l} = \hat{\boldsymbol{\Theta}}_{e,l-1} + \bar{w}\mathbf{L}\xi_l, \qquad \xi_l \sim n.i.d.(\mathbf{0}_n, \mathbf{I}_n)$$

where $\bar{w}$ is the *jump* of the proposed update and $\mathbf{L}$ is the Cholesky decomposition of the covariance matrix of $\boldsymbol{\Theta}_e$. The decision rule now determines if we use the updated $\boldsymbol{\Theta}_{e,l}$ in the next iteration or keep $\hat{\boldsymbol{\Theta}}_{e,l-1}$. The decision rule compares a random variable drawn from a uniform distribution $\varsigma_l \sim U(0,1)$ to

$$\varrho_l = \min\left\{ \frac{\mathcal{L}(\mathcal{M}_T|\boldsymbol{\Theta}_{e,l}, \boldsymbol{\Theta}_c)\mathcal{P}(\boldsymbol{\Theta}_{e,l})}{\mathcal{L}(\mathcal{M}_T|\hat{\boldsymbol{\Theta}}_{e,l-1}, \boldsymbol{\Theta}_c)\mathcal{P}(\hat{\boldsymbol{\Theta}}_{e,l-1})}, 1 \right\}$$

where $\mathcal{L}(\mathcal{M}_T|\boldsymbol{\Theta}_{e,l}, \boldsymbol{\Theta}_c)$ and $\mathcal{L}(\mathcal{M}_T|\hat{\boldsymbol{\Theta}}_{e,l-1}, \boldsymbol{\Theta}_c)$ are the likelihoods evaluated by the Kalman filter for $\boldsymbol{\Theta}_{e,l}$ and $\hat{\boldsymbol{\Theta}}_{e,l-1}$, respectively, and $\mathcal{P}(\boldsymbol{\Theta}_{e,l})$ and $\mathcal{P}(\hat{\boldsymbol{\Theta}}_{e,l-1})$ are respective prior distributions. The decision rule, then, states

$$\hat{\boldsymbol{\Theta}}_{e,l} = \begin{cases} \boldsymbol{\Theta}_{e,l} & \text{if } \varsigma_l \leq \varrho_l \\ \hat{\boldsymbol{\Theta}}_{e,l-1} & \text{else} \end{cases}$$

The counter is updated as

$$c = \begin{cases} c+1 & \text{if } \varsigma_l \leq \varrho_l \\ c+0 & \text{else} \end{cases}$$

Note that the decision rule accepts whenever the prior-weighted likelihood of the proposed parameter vector $\boldsymbol{\Theta}_{e,l}$ exceeds the prior-weighted likelihood of the previous reference vector $\hat{\boldsymbol{\Theta}}_{e,l-1}$. On the other extreme, if the likelihood of the proposed vector is zero, it will never be accepted.

Delicate choices regarding the determination of the initial value $\hat{\boldsymbol{\Theta}}_{e,0}$, the number of MH chains simulated, the *jump* parameter $\bar{w}$ and the number of iterations H need some further discussion. To obtain $\hat{\boldsymbol{\Theta}}_{e,0}$, we compute the mode of the posterior distribution of the parameters which can be obtained by MCMC based optimization methods. The mode is then used to initialize a *burn-in* state of the MH-MCMC algorithm. The parameter vector $\hat{\boldsymbol{\Theta}}_e$ of the final iteration of the burn-in state gives the initial value $\hat{\boldsymbol{\Theta}}_{e,0}$ for the H MH iterations.

To further reduce the dependence of the posterior from the initial value, multiple MH chains can be simulated with each additional one starting the burn-in from the next best mode of the posterior. Note that the burn-in state is also used to estimate the covariance matrix of the parameters $\mathbf{LL}'$ needed for computing the Choleski decomposition for the random walk law of motion. The jump parameter $\bar{w}$ should be chosen such that the acceptance ratio c/H is between 0.23 and 0.30. The number of iterations H should be chosen sufficiently large to ensure convergence of the proposals $\hat{\mathbf{\Theta}}_{e,l}$ to $\mathcal{P}(\hat{\mathbf{\Theta}}_e|\mathcal{M}_T, \mathbf{\Theta}_c)$.

Hans **Böckler**
Stiftung

Fakten für eine faire Arbeitswelt.