WORKING paper

BANQUE DE FRANCE
EUROSYSTÈME

Estimating Non-Linear DSGEs with the Approximate Bayesian Computation: an application to the Zero Lower Bound

Valerio Scalone[1]

August 2018, WP #688

ABSTRACT

Abstract: Estimation of non-linear DSGE models is still very limited due to high computational costs and identification issues arising from the non-linear solution of the models. Besides, the use of small sample amplifies those issues. This paper advocates for the use of Approximate Bayesian Computation (ABC), a set of Bayesian techniques based on moments matching. First, through Monte Carlo exercises, I assess the small sample performance of ABC estimators and run a comparison with the Limited Information Method (Kim, 2002), the state-of-the-art Bayesian method of moments used in DSGE literature. I find that ABC has a better small sample performance, due to the more efficient way through which the information provided by the moments is used to update the prior distribution. Second, ABC is tested on the estimation of a new-Keynesian model with a zero lower bound, a real life application where the occasionally binding constraint complicates the use of traditional method of moments.

Keywords: Monte Carlo analysis; Method of moments, Bayesian, Zero Lower Bound, DSGE estimation

JEL classification: C15, C11, E2.

[1]Banque de France, Macro-prudential policy Division - Directorate of Financial Stability. Email: valerio.scalone@banque-france.fr. I would like to thank Salvatore Nistico' and Giuseppe Ragusa for their support. I also thank Stephane Adjemian, Pierpaolo Benigno, Nicola Borri, Nuno Coimbra, Sergio de Ferra, Giorgio Di Giorgio, Jesus Fernandez-Villaverde, Maddalena Galardo, Raffaella Giacomini, Anastasios Karantounias, Lena Koerber, Dennis Kristensen, Riccardo Masolo, Marco Lippi, Lucrezia Reichlin, Pietro Reichlin, Federica Romei, Frank Schorfheide, Paolo Surico for useful comments. I am also grateful to seminar participants at LUISS Guido Carli, the London Business School, UCL, University of Pennsylvania, Banque de France and EABCN-EUI Conference.

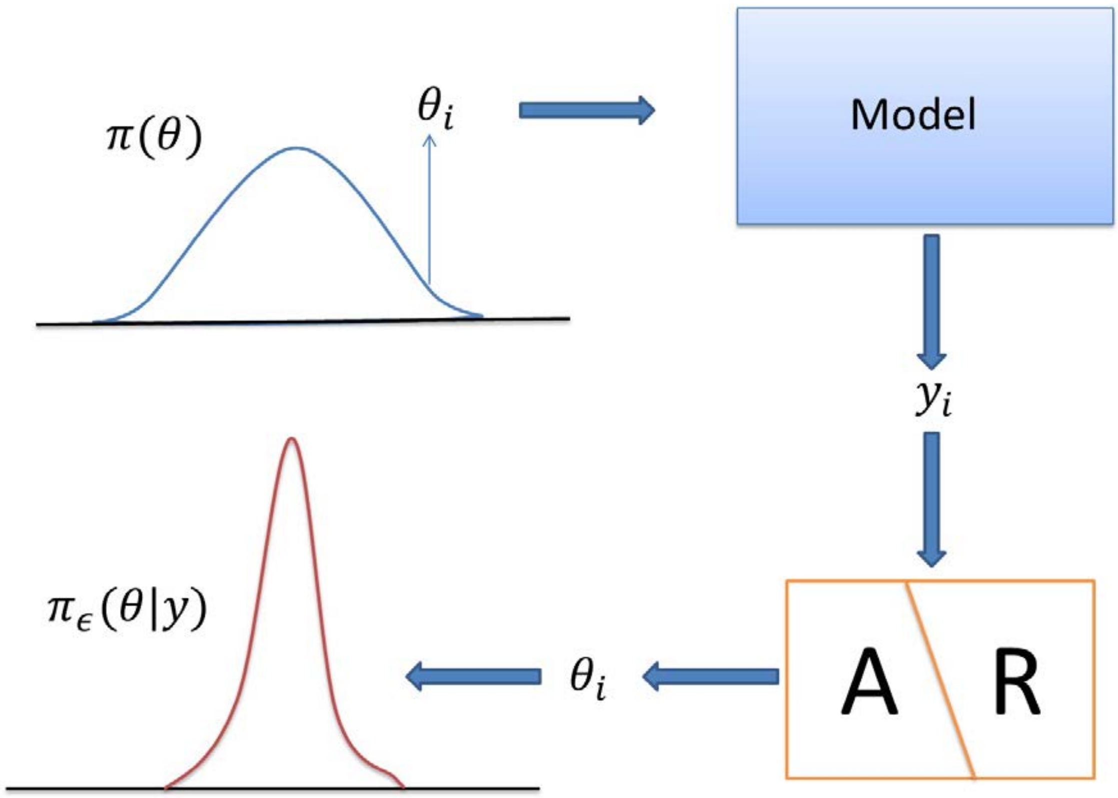
NON-TECHNICAL SUMMARY

Despite the growing importance of non-linear DSGE models, their estimation is still very limited mainly due to the high computational burden associated to the non-linear estimation techniques. Beside, some identification issues can arise due to non-regularities deriving from the non-linear solution.

Are there estimation methods which overcome those issues and make non-linear DSGE estimation easier to apply? The Bayesian versions of the methods of moments can provide a useful response in this sense. In particular, this paper advocates for the use of Approximate Bayesian Computation, a set of techniques developed in epidemiology and population genetics (Pritchard et al. [2000]).

ABC is a set of Bayesian methods through which the prior distribution of the structural parameters of the model is updated by the information provided by the moments (i.e. variances, covariances, etc.). In a first step, a large number of vectors of parameters is drawn from the prior distribution. In a second step, for each vector o parameter, the model is simulated. For each simulation, the distance between the vector of simulated moments and the vector of the observed moments is computed. In a third step, a selection step is made for each simulation: simulations are accepted only if the Euclidean distance is smaller than a fixed threshold, otherwise they are rejected. Finally, the parameters associated to the accepted simulations provide an approximation to the posterior distribution.

Schematic representation of the basic algorithm of Approximate Bayesian Computation



Through Monte Carlo exercises, I show that ABC has a better small sample performance compared to the state-of-the-art Bayesian method of moments used in DSGE estimation, the Bayesian Limited Information (Kim [2002], Christiano et al. [2015]). This result hinges on the way

through which the distribution of the moments is obtained and used to update the prior distribution.

Furthermore, I estimate a New-Keynesian DSGE model with an occasionally binding ZLB by using simple ABC-rejection. I estimate the model on US data by using ABC and I find three results. First, including the period of the ZLB in the sample is crucial to correctly estimate the probability of hitting the Zero Lower bound. Second, the use of conditional moments in the estimation (i.e. moments conditional on the state of the economy) can ease the identification issues generated by the presence of the Zero Lower Bound. Third, estimation helps the model to replicate some of the main features observed before and after the Great Recession and to ease the over-reaction of macroeconomic variables predicted by the DSGE with ZLB, as highlighted by Fernandez-Villaverde [2015].

Estimation des modèles DSGE non-linéaires par *Approximate Bayesian Computation* : une application au *Zero Lower Bound*

L'estimation des modèles DSGE non-linéaires est encore très limitée, vu les coûts de calcul importants et les problèmes d'identification qui peuvent découler de la forme non-linéaire de la solution du modèle. Ce papier préconise l'utilisation de la *Approximate Bayesian Computation* (ABC), une méthode des moments bayésienne. À travers des exercices Monte-Carlo, la performance de l'ABC pour de petits échantillons est comparée à celle de la méthode de l'Information limitée (BLI- Kim, 2002), l'état de l'art dans l'estimation des modèles DSGE par méthodes des moments. On observe une meilleure performance de l'ABC par rapport au BLI, en particulier quand le nombre d'observations est réduit. A titre d'exemple, l'ABC est testée sur l'estimation d'un modèle néo-keynésien avec une politique monétaire au zero lower bound.

Mots-clés : Analyse Monte Carlo; Méthode des moments, Bayésien, Zero Lower Bound, estimation des modèles DSGE

1 Introduction

Despite the growing importance of non-linear DSGE models, their estimation is still very limited. First, considering the likelihood approach (Fernández-Villaverde and Rubio-Ramírez [2007]), this fact can be explained by the computational burden associated to the Particle filter. Second, standard methods of moments (GMM, SMM) can present important small sample bias and encounter identification issues, related to the non-linear structure of the model.

Are there estimation methods which overcome those issues and make non-linear DSGE estimation easier to apply? The Bayesian versions of the methods of moments can provide a useful response in this sense. In particular, this paper advocates for the use of Approximate Bayesian Computation, a set of techniques developed in epidemiology and population genetics (Pritchard et al. [2000], Blum [2010]). In ABC, the information provided by the data is conveyed in moments, which are used to update the prior distribution on the parameters. In the paper, I highlight the positive aspects related to the use of ABC in estimation of non-linear DSGEs. Through Monte Carlo exercises, I show that ABC has a better small sample performance compared to the state-of-the-art Bayesian method of moments used in DSGE estimation, the Bayesian Limited Information (BLI, Kim [2002], Christiano et al. [2010b, 2015]). This result hinges on the way through which the distribution of the moments is obtained and used to update the prior distribution. Furthermore, by applying ABC to a New-Keynesian DSGE model with an occasionally binding ZLB, I show that the use of moments limits the computational burden associated to the likelihood approach, making the estimation of medium-scale DSGE models feasible. Before running the estimation, I highlight the computational and identification issues related to the use of moments in non-linear models, generated by the non-regular mapping functions relating the structural parameters and the moments used in the estimation. To this extent, I show how ABC -and Bayesian method of moments- can ease those problems. First, the Bayesian structure of ABC allows to fix the computational issues related to non-regularity in objective functions, generated by the use of moments in a non-linear framework. Second, the good small sample properties of ABC simplify the use of conditional moments (i.e. moments computed conditional on the state of the economy), easing identification issues.

I estimate the model on US data by using ABC and I find three results. First, including the period of the ZLB in the sample is crucial to correctly estimate the probability of hitting the Zero Lower bound. Second, the use of conditional moments in the estimation (i.e. moments conditional on the state of the economy) can ease the identification issues generated by the presence of the Zero Lower Bound.

Third, estimation helps the model to replicate some of the main features observed before and after the Great Recession and to ease the over-reaction of macroeconomic variables predicted by the DSGE with ZLB, as highlighted by Fernández-Villaverde et al. [2015].

When the economy enters in a period of Zero Lower Bound, the relations between the observable variables significantly varies with respect to normal times (for example the policy rate does not positively co-vary with income in ZLB), generating non-regular mapping functions. This phenomenon is associated to two main issues: 1) an identification issue; 2) a computational issue. First, the use of unconditional moments can cause important identification issues in small sample, as underestimating the parameters controlling the non-linearity of the model. For example, in a linear New-Keynesian model, standard deviations of demand side shocks are usually positively related to the covariance between interest rate and output. Instead in a non-linear model, large negative demand shocks can push the interest rate on its (zero) lower bound. To this extent, higher standard deviations can be associated to prolonged periods of Zero Lower bound where the covariance between interest rate and output is equal to zero. Therefore, the use of unconditional covariance (i.e. without distinguishing between periods of ZLB and Taylor rule periods) can lead to underestimate the standard deviation of demand side shocks. Instead, the use of conditional moments (moments computed conditionally on the state of the economy) can ease the identification issue, by making the use of the information provided by the data more efficient. In the case of a DSGE with ZLB, moments can be computed conditionally on being (and not being) on the Zero Lower Bound and be included in the computation of objective functions. However, the presence of conditional moments can amplify the small sample bias. To this extent, ABC proves to have a better small sample performance than the other Bayesian alternative (BLI).

Second, the presence of non-regular mapping functions generate non-regular objective functions, whose minimum can be computationally hard to find with the optimization algorithms used in frequentist methods (GMM, SMM). Instead, the Bayesian structure of ABC allows to run estimation also when the objective function is non-regular. Overall, the use of conditional moments in a Bayesian framework, as such is the case of ABC, appears crucial to help the model to replicate some of the main features observed in the macroeconomic data during the period of Zero Lower Bound.

In Economics, application of ABC is still very limited.[1] ABC is a set of Bayesian

[1] Creel and Kristensen [2011] propose the Indirect Likelihood Inference, a type of estimator whose Bayesian simulated version (Simulated Bayesian Indirect Likelihood estimator, SBIL) coincides with a variant of ABC (ABC-kernel). In their paper they also show an application to

methods through which the prior distribution $p(\theta)$ is updated by the information provided by the moments s (i.e. variances, covariances, etc.). The model is simulated a large number of times with different parameters θ_i, drawn from the prior. For each simulation, the distance between the vector of simulated moments s_i and the vector of the observed moments s is computed ($\rho_i = ||s_i - s||$). In the selection step, simulations are accepted only if the Euclidean distance is smaller than a fixed threshold ϵ. The parameters associated to the accepted simulations provide an approximation to the posterior distribution $p(\theta|s = s^*)$. Importantly, each simulation has the same sample size of the observed sample. This allows to make inference by studying the actual distribution of the moments (i.e. taking into account the sample size). This feature marks the main difference of ABC with respect to the BLI method. The BLI is often interpreted as the Bayesian version of GMM/SMM estimators, where the likelihood of the moments is used to update the prior distribution. Likelihood is obtained relaying on the asymptotic distribution of the moments. However, in small sample, the actual distribution of moments (biased with respect to the analytic value) can substantially differ with respect to the one deriving from the asymptotic distribution (normal and centered around the analytic value). This difference delivers a comparative advantage of ABC with respect to BLI under small sample. This statement is shown in two Monte Carlo exercises: on an AR(1) with different autocorrelations and on a RBC model. In order to run the comparison from a Bayesian perspective, the ABC and the BLI are assessed with respect to the full likelihood posterior distributions. The smaller the sample, the larger will be the advantage of ABC with respect to the BLI.

Literature. This paper is related to three streams of literature. First, the paper focuses on methods to estimate non-linear DSGE models. In order to compute the likelihood and estimate the models, Fernández-Villaverde and Rubio-Ramírez [2007] proposed the use of the Particle filter.[2] Alternatively, Ruge-Murcia [2007, 2012], Christiano et al. [2010b], Christiano et al. [2015] apply methods of moments to estimate DSGE models. In particular, Ruge-Murcia [2012] applies the Simulated Method of Moments (Duffie and Singleton [1990]) to estimate a non-linear DSGE model of an economy subject to asymmetric productivity shocks. In order to overcome weak identification issues and add extra-data information, Christiano et al. [2010b], Christiano et al. [2015] adopt the Limited information method (BLI) by

a non-linear DSGE model. Furthermore, Grazzini et al. [2017] apply the ABC to estimate an agent-based model.

[2]Despite being a fascinating device, this alternative tool is often considered too computationally burdensome to handle medium or large-scale DSGE models. Besides, for the application of the Particle filter, measurement errors must be added to the observable variables,and in most cases, given the size of the model, the standard deviation of the measurement errors is fixed in advance.

Kim [2002] to estimate linear DSGE models with financial frictions. This paper contributes to this literature along more dimensions. As a main contribution, the paper focuses on ABC methods, which are relatively new in Economics, and runs a Monte Carlo analysis to compare its small sample performance with the one of of the BLI, the state-of-the-art Bayesian method of moments popular in the DSGE literature. Furthermore, the paper highlights the consistency and identification issues arising when using the frequentist methods of moments (GMM, SMM, Indirect Inference) in the context of non-linear DSGE models, in particular by focusing on the identification issues related to the use of unconditional moments in the context of models featuring Occasionally Binding Constraints (OBC). Finally, the paper shows that those issues can be partially fixed by matching the conditional moments and the Bayesian versions of the methods of moments (ABC and BLI).

I contribute to a second stream of literature bridging the gap between the ABC literature developed in population genetics and the Economics literature. Creel and Kristensen [2011, 2016], Forneron and Ng [2015], Creel et al. [2015] explored the use and the asymptotic properties of ABC methods in the estimation of economic models. In particular, Creel and Kristensen [2011] present the Indirect Likelihood Inference estimator, whose simulated Bayesian (SBIL) version coincides with the ABC, to a DSGE models. Creel and Kristensen [2011] provide asymptotic results for the estimator and show that Indirect Likelihood Inference estimators have a higher order efficiency with respect to the GMM style estimators. In their paper, using a DSGE model solved by perturbation methods, they compare the small sample performance of the SBIL estimator with the one of the Simulated Method of Moments. The Monte Carlo experiments are run from a frequentist perspective, by computing the RMSEs of the SMM and of the SBIL estimator with respect to the true values. In this paper, the comparison is run between ABC and the BLI and is made from a Bayesian perspective, by assessing to which extent the posteriors obtained from the BLI and the ABC approximate the full likelihood posterior distributions. Also, I apply ABC to a real life application: a DSGE with an occasionally binding Zero Lower Bound, solved by piecewise linear approximation (Guerrieri and Iacoviello [2015]) and estimated on US data using different sample sizes. In doing that, I focus on the ability of ABC to handle non-regular objective functions, coming from the use of moments computed in a non-linear framework.

The third stream of literature focuses on DSGE models with the Zero Lower Bound (Eggertsson and Woodford [2003], Fernández-Villaverde et al. [2015], Gust et al. [2012], Aruoba et al. [2013]). Gust et al. [2012] estimate a new-Keynesian DSGE model with the ZLB, by estimating the model with three observables (inflation, output and interest rate). In order to compute the likelihood they use the particle filter. In this paper, I estimate a DSGE with the ZLB (Fernández-Villaverde et al. [2015]) by using a matching moments procedure, with conditional moments. Besides, Aruoba et al. [2013] and Christiano et al. [2015] estimate a non-linear model by using data pre-2008 (i.e. excluding the ZLB from the sample). In this paper, the benchmark estimation is run on a sample which includes the ZLB while conditional moments on the regime of monetary policy are used to compute the objective function. Data after 2008 are shown to provide crucial information to help the model to replicate some of the essential features observed in macroeconomic aggregates during the Great Recession and the slow recovery period.

The reminder of the paper is the following. Section 2 presents the ABC techniques. Section 3 runs the comparison between ABC estimator and the BLI estimator. Section 4 exposes the challenges related to the use of method of moments in the estimation of the DSGE with the ZLB and the estimate results by ABC-SMC. Section 5 concludes.

2 Approximate Bayesian Computation.

Approximate Bayesian Computation (ABC) is a set of statistical techniques developed in population genetics at the end of the 90's: Pritchard et al. [2000] developed the basic algorithm (ABC-rejection), while a series of computational refinements have been introduced over time (Beaumont et al. [2002], Marjoram et al. [2003], Sisson et al. [2007]). In the last decade, the methodology spread across all natural sciences, namely epidemiology, ecology and biology. In Economics, ABC methods have been applied by Grazzini et al. [2017], to estimate an agent-based macroeconomic model.

ABC is a likelihood-free method through which the prior distribution is updated by the information provided by the moments of the observed sample (e.g. variances). The core of ABC generally presents three steps. First, the model is simulated a large number of times. Importantly, each simulation has the sample size of the observed sample and is run by using different parameters. Second, a measure for the distance between the simulated data and the observed ones is computed for each simulation. Third, a selection step is made on the simulations, on the basis of this measure. The parameters associated to the accepted simulations provide an approximation to the full likelihood posterior distribution.

ABC simplest form -the ABC-rejection- features the following pseudo-algorithm:

- Draw θ_i from the prior distribution $p(\theta)$;

- Simulate the model and get the variable $\mathbf{y}_i$;
- Compute the summary statistics $\mathbf{s}_i$;
- If the Euclidean distance $\rho||\mathbf{s}_i - \mathbf{s}|| < \epsilon$ accept θ_i otherwise reject it;
- Repeat the procedure for N times;

where $\mathbf{s}_i$ is the vector of moments from the simulated sample, $\mathbf{s}$ is the vector of moments of the observed data, ϵ is the tolerance level. In this method, the initial parameters are drawn from the prior distribution $p(\theta)$. For each simulation the selection criterion is the Euclidean distance ρ_i between the summary statistics of the simulations (i.e. moments) $\mathbf{s}_i$ and the ones from the observed sample $\mathbf{s}$. The selection step is an accept-reject: when the Euclidean distance is smaller than a fixed threshold ϵ, the simulation is accepted. The parameters of the accepted posterior distributions are a sample of the approximate posterior distribution $p(\theta|\mathbf{s_i} = \mathbf{s})$. The Bayes Rule of the Bayesian statistics is approximated in the following way:

$$P(\theta|\mathbf{y}) \propto L(\mathbf{y}|\theta)P(\theta) \rightarrow P(||\mathbf{s}_i - \mathbf{s}|| < \epsilon)P(\theta) \tag{1}$$

where $P(\theta|\mathbf{y})$ is the exact likelihood posterior distribution, $L(\mathbf{y}|\theta)$ is the exact likelihood, $P(\theta)$ is the prior distribution. In particular, the likelihood function is approximated by the accept-reject step on the Euclidean distances criterion, while the ABC posterior distribution is an approximation to the full likelihood posterior distribution.

To this extent, asymptotic theory for ABC methods studies the conditions under which the ABC posterior distributions converge to the full likelihood posterior distribution. In particular, Blum [2010], Biau et al. [2015] and Barber et al. [2015] show that, under very weak assumptions, if the moments used in the estimation are sufficient statistics, the approximate posterior distribution converge to the full likelihood posterior distribution for $\epsilon \rightarrow 0$ and $N \rightarrow \infty$. Besides, Barber et al. [2015] quantify the bias introduced by the presence of a positive tolerance level $\epsilon > 0$. Given a threshold level equal to ϵ, the ABC bias is asymptotically proportional to ϵ^2 as $\epsilon \rightarrow 0$.

Concerning the asymptotic efficiency, Creel and Kristensen [2011] show that ABC-style estimators have a higher order efficiency with respect to the GMM-style estimators.[3] The higher efficiency is the result of the fact that in ABC the simulations have the same size of the observed sample. Thanks to this feature, ABC

[3] In fact, Creel and Kristensen [2011] provide asymptotic results for the Indirect Likelihood inference estimator. In this paper, the authors propose a series of estimators for which inference is based on exploiting the information provided by the likelihood of the moments: the indirect

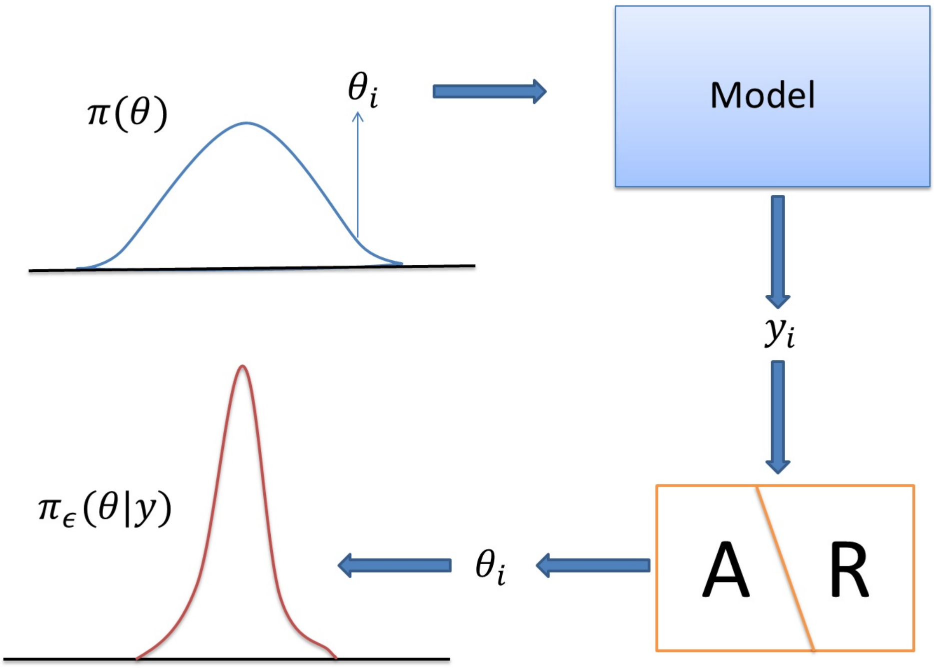


Figure 1: Illustration of the ABC-rejection algorithm.

exploits the simulated moments distribution, in order to update the prior. This simulated distribution takes into account the actual simulated distribution of the moments , i.e. taking into account the sample size, and represents one of the main points of strength of ABC with respect to the other methods of moments. This point is analyzed and presented in more depth in the following section.

ABC posterior differs from the full likelihood posterior distribution for three issues. First, the positive tolerance level ϵ which cannot be equal to 0, introduces an approximation error: accepted simulations are not associated to the minimal distance between the simulated and the observed moments. Second, the number of simulations cannot be infinite: this creates a simulation error (called also sampling error), deriving from the fact that we learn the approximate posterior through a finite sample of accepted simulations. Third, the moments used in the estimation are hardly sufficient: a part of the information contained by the full data is partially lost when using moments.

likelihood. The Maximum Indirect Likelihood (MIL), the Bayesian Indirect Likelihood (BIL) and their simulated versions are illustrated (the Simulated Maximum Indirect Likelihood, SMIL, and and the Simulated Bayesian Indirect Likelihood, SBIL). The SBIL estimator coincides with the ABC. In the paper, they provide asymptotic results for these estimators. In the proposition 1, they affirm that MIL and BIL are consistent estimators. They prove the asymptotic equivalence of the simulated version of the MIL and BIL. They show the first order equivalence relative between the MIL, the BIL and the GMM-style estimators (GMM, SMM, Indirect Inference). Importantly, they also show that the IL estimators are higher order efficient relative to the moments based estimators. In fact, high order expansions reveal that the second order variance of the estimator is smaller than the one obtained by the GMM-style estimators.

Moreover, concerning the first two points, a trade-off exists between the approximation and the simulation error: given a fixed number of simulations, if we lower the tolerance level, the approximation error decreases while the simulation error increases. In order to reduce the simulation error without increasing the approximation error, it is necessary to increase the number of simulations N, while keeping the tolerance level fixed. [4] To this extent, Barber et al. [2015] compute the optimal threshold, by minimizing a loss function which takes into account the bias introduced by the threshold, and the computational cost associated to the production of the simulations.

As highlighted by (Biau et al. [2015]), in practice ABC is a a k-nearest neighbour estimator: instead of fixing a threshold level, practitioners prefer to fix in advance the fraction of the simulations which are accepted. By doing this, the tolerance level is automatically determined by the largest Euclidean distance associated to the accepted simulations. On an exploratory stage, the estimation can be run with different tolerance level and results can be compared. From the comparison of the results obtained with tolerance levels, we can assess whether the tolerance level is sufficiently low to identify our parameters of interests or if we need to reduce the tolerance level further, in order to tackle the approximation error. As a rule of thumb, very irregular shapes of the approximate marginal posteriors for the different parameters can be associated to large simulation errors, due to a tolerance level which is too low with respect to the number of simulations N. Conversely, an important approximation error can arise when the tolerance level is too large to correctly identify parameters that are weakly identified (i.e. parameters for which the curvature of the objective function is almost flat). This happens because weakly identified parameters have a low weight in the determination of the Euclidean distance and the selection step is not able to exploit the information provided by the moments and update the marginal prior distribution for those parameters. As a general practice, assessing the curvature of the Euclidean distance with respect to those parameters allows to have an intuition of the magnitude of the approximation error which can be tolerated without compromising identification.

Alternatively, in order to improve identification, additional moments can be added, to improve the identification for certain parameters. More in general, the set of moments needs to get as close as possible to set of information provided by the sufficient statistics (i.e. the set of moments resuming the information provided

[4] The computational cost for generating ABC sample is like δ^{-q}, where q is the dimension of the observations. They also provide a criterion to optimally choose the tolerance level, balancing the MSE, depending on the simulation error, with the computational cost, related to the costs associated to producing a number of simulations.

by the data). On the one hand, the moments selected will be hardly sufficient. Moreover, it is hard to determine to which extent they miss the sufficiency conditions. On the other hand, this flexibility can be exploited to include moments that are particularly informative on some parameters. Finally, it is worth to stress that these same issues related to the moments selection are found also in the other methods using limited information (GMM, SMM and Indirect Inference estimators (Ruge-Murcia [2012], Christiano et al. [2016]). To this extent, Creel and Kristensen [2016] developed an algorithm for the selection of moments.

Concerning the estimation exercises for the paper, DSGEs are estimated by matching the variances, the covariances and the autocovariances up to order 2, the skewness and the courtosis of the observable series (consumption, interest rate, income). In the main estimation exercise, the moments are computed conditionally on the state of the monetary policy regime (unconstrained monetary policy versus Zero Lower Bound binding). These moments are not necessarily sufficient to ensure the convergence to the full likelihood posterior distribution.To this extent, keeping a Bayesian perspective, ABC can be interpreted as a procedure through which the extra-data information (prior) is updated by the information conveyed in the observed moments, without necessarily claiming the convergence to the full likelihood posterior. Overall, increasing the number of moments can be a good strategy to converge towards the sufficiency. However, increasing the number of moments decreases the asymptotic rate through which the approximated posterior converges to the full likelihood posterior (Beaumont et al. [2002], Barber et al. [2015]). As explained by Grazzini et al. [2017], this finding derives from the trade-off between the approximation error and the simulation error: when the number of moments increases, for a fixed number of simulations, either the tolerance level becomes too high (important approximation error), or the acceptance ratio too low (important simulation error). In other words, ABC encounters the curse of dimensionality, especially when the number of parameters to estimate is large. This phenomenon is amplified when the prior and the posterior distribution strongly differ: because of the low acceptance ratio, a large number of simulations must be produced to contain the simulation error.

In order to improve the computational efficiency and ease the curse of dimensionality, several types of types of refinements have been developed. One approach consists in refining the algorithms to better exploit the information provided by the Euclidean distance and by the different moments. For example, accepted simulations can be assigned a weight according to a kernel function. The argument of the kernel is the Euclidean distance: the smaller the distance, the larger will be the weight. Puit it differently, the indicator function which is implicitly applied in the ABC algorithm, is substituted by other types of kernel functions (Normal kernel, Epanechnicov etc.). In the remainder of the paper, we call this specification *ABC-kernel*.[5] Furthermore, a post-sampling correction step can be added. In ABC-regression (Beaumont et al. [2002]), the accepted parameters are regressed against the mismatch between the simulated and observed moments. The estimated regression is then exploited to correct the accepted parameters.

Alternatively, a more efficient way of drawing the parameters can be obtained by ABC-Importance Sampling (Creel and Kristensen [2016]), ABC-MCMC (Marjoram et al. [2003]) and ABC-SMC (Sisson et al. [2007]). These algorithms exploit the core mechanism of ABC-rejection in a Monte Carlo structure. In Section 4, a DSGE with ZLB is estimated by ABC-rejection. For ABC-MCMC ad ABC-Regression and ABC-SMC, more details on can be found in the Appendix.

3 A comparison with the Bayesian Limited Information Method

In this section I compare the small sample performance of the ABC estimators with the one of Limited Information Method (Kim and Kim [2003]). The BLI has been chosen to run the comparison for two reasons. First, it is the state-of-the-art method of moments used in DSGE estimation (Christiano et al. [2010a, 2016]). Second, being Bayesian, it can be directly compared to the ABC, in that both the methods aim at providing an approximation of the full likelihood.

In BLI, the prior distribution is updated by the likelihood of the *moments*, obtained by relying on moments asymptotic distribution, via Central Limit Theorem. Given the vector of parameters θ, the sample moments $\hat{\gamma}$ and the estimated variance of the moments $\hat{V}$, the Approximate Posterior distribution $P(\theta|\hat{\gamma}, \hat{V})$ is obtained according to the Bayesian updating rule:

$$P(\theta|\hat{\gamma}, \hat{V}) = \frac{P(\hat{\gamma}|\hat{\theta}, \hat{V})P(\theta)}{P(\hat{\gamma}|\hat{V})} \tag{2}$$

where T is the number of moments,$\hat{\gamma}$ is the vector of sample moments, $\gamma(\theta)$ is the vector of analytical moments depending on the parameter θ, $P(\theta)$ is the prior

[5] The ABC kernel coincides with the simulated Bayesian version of the estimator proposed by Creel and Kristensen [2011] (i.e. the Bayesian Indirect Likelihood estimator).

distribution. The likelihood $P(\gamma|\hat{\theta},\hat{V})$, conditional on $\hat{V}$, is computed according to:

$$P(\gamma|\hat{\theta},\hat{V}) = \frac{1}{(2\pi)^{(\frac{N}{2})}}|\hat{V}|^{-\frac{1}{2}} \exp\left\{-\frac{T}{2}\left(\hat{\gamma}-\gamma(\theta)\right)' \hat{V}^{-1}\left(\hat{\gamma}-\gamma(\theta)\right).\right\}. \tag{3}$$

In BLI, the likelihood of the moments, thanks to their asymptotic normality, is computed by focusing just on the first and the second moments of their distribution. This marks a key difference with ABC methods, where the simulations have the same size of the observed sample and the simulated distribution of moments takes into account the small sample bias. This difference may result in an important comparative advantage of ABC with respect to the BLI estimator, when the actual and the asymptotic distribution of the moments substantially differ, as it can be the case of small samples.

The relative performance of the ABC estimator with respect to the BLI method is measured in two Monte Carlo exercises. The exercise is run from a Bayesian perspective: the goal is to understand to which extent the two methods approximate the full likelihood posterior distributions under small samples. To do that, I compute two different objects. First, I compute the Root Mean Square Error (RMSE) with respect to the Full likelihood Posterior Mean. RMSE measures how close are the two estimators to the Full likelihood Bayesian estimator (the Posterior Mean). In particular, the RMSE is obtained by:

$$RMSE = \frac{1}{N}\frac{\sum\left(\hat{\theta}_{app}-\hat{\theta}_{full}\right)^2}{\theta}, \tag{4}$$

where $\hat{\theta}_{app}$ is the mean of the posterior of one of the two approximating methods, $\hat{\theta}_{full}$ is the full likelihood posterior mean.

Second, I compute the Overlapping Ratio, capturing which of the two methods deliver a better approximation of the posterior distributions. The Overlapping Ratio is obtained by:

$$OR = \frac{CI_{90\%,App} \cap CI_{90\%,Fl}}{CI_{90\%,App} \cup CI_{90\%,Fl}} \tag{5}$$

where $CI_{i-\%,App}$ is the $i-th$ Percentile of the Approximate Posterior distribution, $\cap$ stands for Intersection and $\cup$ for Union. The Overlapping Ratio is always included in the interval $[-1,1]$. For example if the two intervals perfectly coincide the Overlapping Ratio equals 1, whereas if two degenerate posterior distributions do not overlap at all, the Overlapping Ratio equals -1.

I run the Monte Carlo experiment on two different models: an AR(1) model and a RBC model subject to some weak identification issues.[6]

3.1 Case 1: AR(1)

In this subsection I run a Monte Carlo experiment on an AR(1) process and I find that the ABC methods outperform the BLI method in approximating the full likelihood posterior distribution. The exercise is run for for different auto-correlations in the Data Generating Process (DGP) and for different sample sizes. According to the results, the advantage of ABC increases with the auto-correlations in the DGP. Also, the smaller the sample, the larger the advantage of ABC with respect to BLI.

Before presenting the experiment in more depth, I provide an intuition on why the sample size and the persistence concur in this way to the results. In Fig. 2 and 3, I report the distributions of the sample autocovariances obtained by simulating the AR(1) model, respectively when the autocorrelation ϕ of the DGP 0.5 and when 0.99. The distributions are reported for different sample sizes: from 50 to 1000 observations. In the case of lower auto-correlation (Fig. 2), when the sample size increases, the distribution of the autocovariances converges to a normal distribution with mean equal to the population autocovariance ($\gamma = \phi/(1-\phi^2)$), (the pink plane), whereas in small sample the distribution is skewed and not centered around the analytic moment. Fig.3 reports the case when $\phi = 0.99$, for which even when the sample contains 5000 observations, the sample distribution substantially differs from the asymptotic one. In the updating step ABC relies on the simulated distribution in order to update the prior distribution, whereas BLI relies on the asymptotic normal distribution of the moments. For small samples and high autocorrelation in Data Generating Processes, this difference plays in favour of a substantial comparative advantage of ABC over BLI.

For each couple of autocorrelations and sample sizes, the estimation for each AR(1) process is run 1000 times. The sample sizes considered are 100, 300, 1000 observations. The autocorrelation factor tuning the persistence assumes the following values $\phi = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99]$. The moment to match is the first order autocovariance, whereas the prior distribution is a Uniform prior $\sim U[0,1]$. For the ABC, 10000 simulations are produced drawing from the prior,

[6] In a similar exercise, Creel and Kristensen [2011] run Monte Carlo experiments to assess the small sample performance of the Indirect Likelihood Inference with respect to Simulated Method of Moments. Their comparison is run from a frequentist perspective, since they compare the performance of the estimators in inferring the true parameter of the Data Generating process. The following exercises can be thought as the Bayesian version of their exercise, given the Bayesian nature of the estimators which are compared and the Bayesian criteria adopted in assessing the performance.

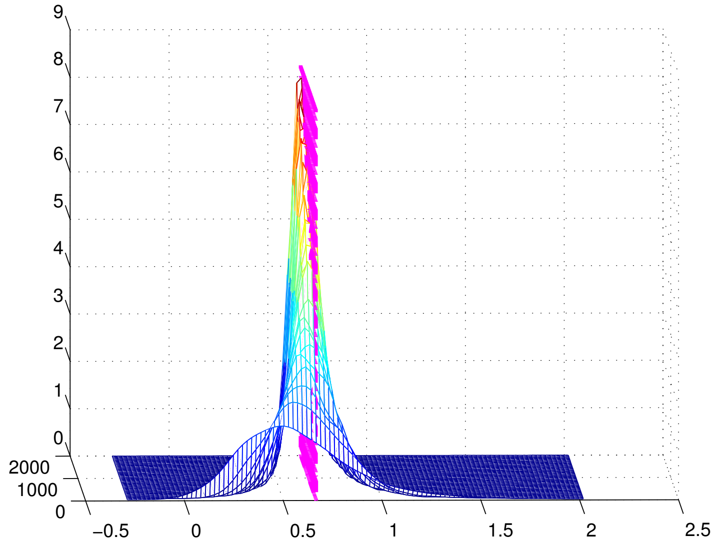


Figure 2: Distribution of the sample autocovariance for an AR(1) process with $\phi = 0.50$ for different sample sizes: from 50 to 2000 observations. The pink plane represents the population autocovariance.

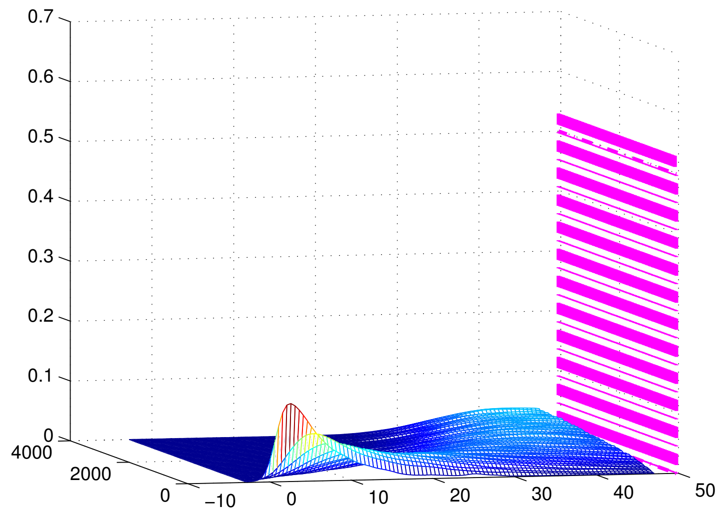


Figure 3: Distribution of the sample autocovariance for an AR(1) process with $\phi = 0.99$ for different sample sizes: from 50 to 4000 observations. The pink plane represents the population autocovariance.

and the 1% of the simulations are accepted, according to the Euclidean distance. The curse of dimensionality does not affect the estimation since 10000 simulations are enough to make the simulation error negligible. For this reason, the correction step of the ABC-regression and the Kernel Weighting do not improve the estimation results upon the ABC-rejection procedure. Only results for ABC-regression are reported for the sake of brevity.

For the Bayesian Limited Information method, the likelihood of the autocovariance is computed and the prior is updated. The posterior distribution is studied with the Importance Sampling algorithm: as importance distribution the prior distribution is used and 10000 samples are drawn for each estimation. The variance covariance matrix $\hat{V}$ is computed with two alternatives: the HAC Variance Covariance Estimator or a bootstrapping procedure.[7]

Fig. 4, 5 and 6 present the evolution of the RMSEs with respect to the Full likelihood posterior mean, obtained using three different sample sizes: respectively 100, 300 and 1000 observations. For each figure, on the horizontal axis, the autocorrelation of the DGP varies from $\phi = 0.1$ up to $\phi = 0.99$. We find that ABC has a smaller RMSE with respect to the BLI. Furthermore, the difference widens in smaller sample and when the DGP is more persistent. Among the different approaches to estimate the variance covariance matrix of the moments $\hat{V}$, the HAC Newey-West estimator ensures smaller RMSEs especially in highly persistent cases and small samples, while the bootstrapping methods has smaller RMSE with low autocorrelations. In large samples, the RMSEs converge, at least for the cases where the autocorrelations ϕ of the DGP is smaller than 0.95.

In order to assess to which extent the posterior distributions approximate the full likelihood posterior, we compute the Overlapping Rations (OR) between the 90th credible intervals of ABC and BLI with respect to the one of the Full Likelihood. Figs. 7, 8 and 9 report the evolution of the Overlapping Ratios. Again, the samples are made of 100, 300 and 1000 observations. The OR difference between ABC methods and BLI is larger in general for highly persistent processes proving that ABC outperforms BLI in approximating the posterior distributions under small

[7] In the first case, the Newey-West estimator is computed, using a Bartlett Kernel, having a bandwidth equal to $B(T) = floor(4 * (T/100)^{(2/9)})$, where T is the sample size. In the second case, the bootstrapping is applied in two steps. A first step estimator is computed to minimize a quadratic objective function using the identity matrix as variance covariance matrix. Afterwards, the AR(1) process is simulated for 1000 times (bootstrapping) to compute the autocorrelation for each bootstrap and the covariance of the moments $\hat{V}$ to compute the likelihood. In this simple case, the identity matrix at the initial step is simply the unity scalar, whereas the the covariance matrix of the second step is the variance of the autocovariances computed in the firt one. This method is inspired to the solution proposed in Christiano, Trebandt ad Walentin, where the covariance matrix is estimated through a bootstrap step.

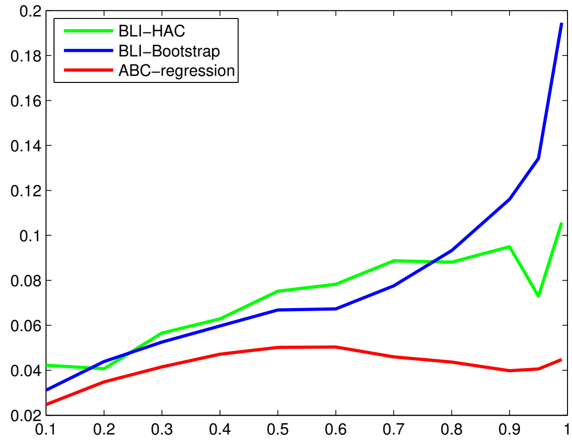


Figure 4: RMSE of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=100. Different autocorrelations on the horizontal axis.

sample. Also from this standpoint, results suggest that among the BLI estimators, the HAC estimation of the variance covariance matrix has a larger OR values than the Bootstrapping Procedure for persistent processes, while the opposite is true for the low persistent cases. [8]

3.2 Case 2: A RBC with identification issues

In this second section, the performance of the two estimators is studied in a more complex application. The experiment is run on a linear RBC model with three structural shocks and three observables.[9] The model equations are presented in Appendix.

The RBC studied in this section encounters some identification issues concerning the preference parameters, due to the presence of three stochastic processes: a

[8]The difference in the result between the two different versions of the BLI is explained by the way through which the likelihood updates the prior.Indeed, in BLI the variance covariance matrix has two roles. First, as in the GMM-style estimators, it is the weighting function of the moments used in the estimation. Second, the variance covariance determines the weight of the likelihood with respect to the prior distribution. In this estimation exercise, the role of the variance covariance is limited to the latter element, since we use only one moment to update the inference. Generally speaking the estimation of the variance co-variance matrix is one of the trickiest steps for all the methods of moments, especially in a context of small sample. In this respect, concerning the BLI, the variance covariance estimation not only affects the computation of the objective function (i.e. the likelihood), but also the weight that this function has in updating the prior distribution.

[9]The comparison is run on a linear solution: the likelihood is computed by Kalman filter. Running a comparison with respect to the likelihood in a non-linear version would have implied the use of Particle filtering for the likelihood computation. This would have introduced an additional computational error in the estimations related to the use of the particle. The use of the linear model allows to avoid this additional source of error in the Monte Carlo exercise.

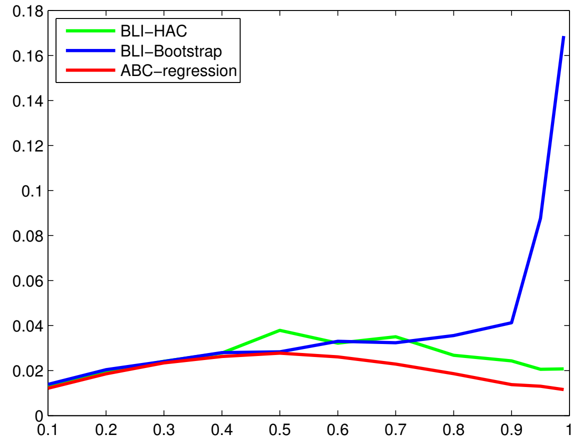


Figure 5: RMSE of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=300. Different autocorrelations on the horizontal axis.

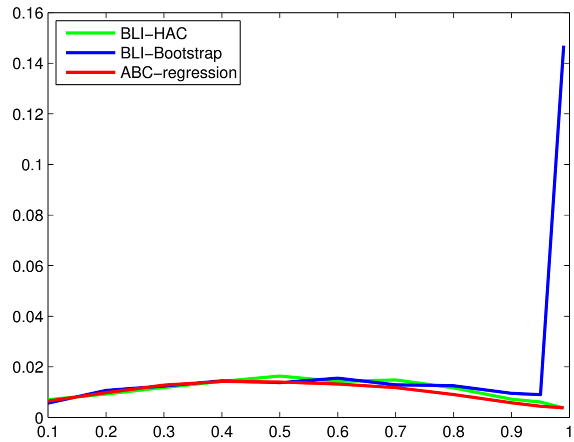


Figure 6: RMSE of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=1000. Different autocorrelations on the horizontal axis.

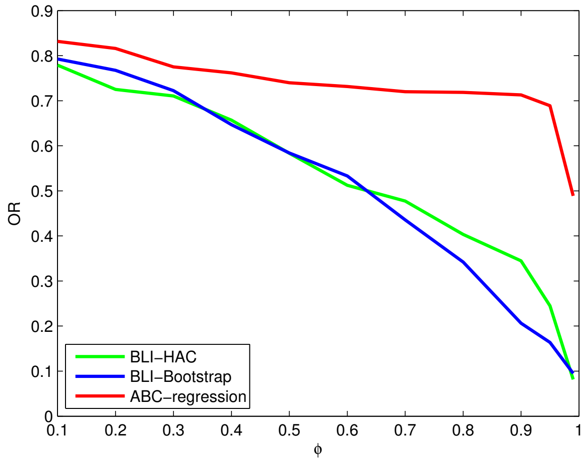


Figure 7: Overlapping Ratios of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=100. Different autocorrelations on the horizontal axis.

productivity shock on the production function, a shock on the preference affecting the labour supply and a shock on the interest rate requested by the household. The presence of these three shocks allows to estimate the full likelihood distribution using three observable variables, without the need of adding measurement errors.

Each Monte Carlo experiment is made of 100 repetitions. The RMSE and the Overlapping Ratio are computed using different sample sizes: 100, 200, 500 observations. The data generating process parameters are the following the subjective discount: $\beta = 0.95$, the utility function parameter $\gamma = 2$, the autocorrelations for the TFP process, for the labour supply and for the interest rate are respectively $\rho_a = 0.95$, $\rho_b = 0.95$, $\rho_d = 0.95$, while the respsective standard deviation of the shocks are $\sigma_a = 0.01$ $\sigma_b = 0.01$, $\sigma_d = 0.01$. The moments used in the estimation are the covariances and the first order autocovariances of three observables: income Y_t, hours H_t and investments I_t.

The prior distribution is reported in Table 1. Concerning the ABC methods, the RBC is simulated 5000 times, the tolerance level is such that the acceptance ratio of the simulations is equal to 5%. We report the results for ABC-rejection, ABC-kernel, ABC-regression. The variance covariance matrix of the BLI estimator is obtained through the HAC Newey-West estimator. For each Full likelihood and BLI estimation, posteriors are studied by MCMC methods (An and Schorfheide [2007]). Each chain contains 10000 draws with a burn-in period of 1000 draws.

Table 2 contains the results of the RMSE for the case of 100 observations, informative prior and high persistence of the process. Overall, RMSEs with ABC are smaller than the RMSEs with BLI. Tables 3 and 4 report the RMSEs respectively

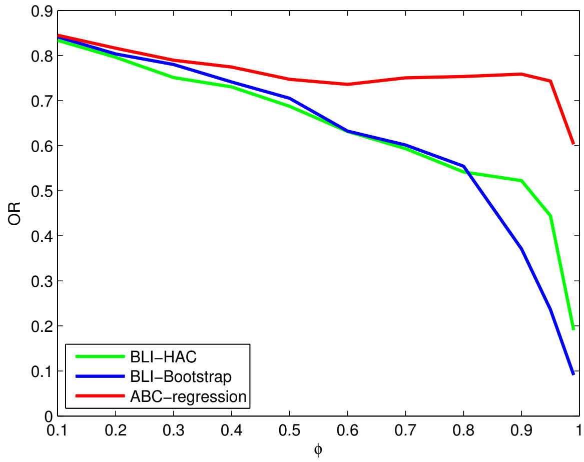


Figure 8: Overlapping Ratios of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=300. Different autocorrelations on the horizontal axis.

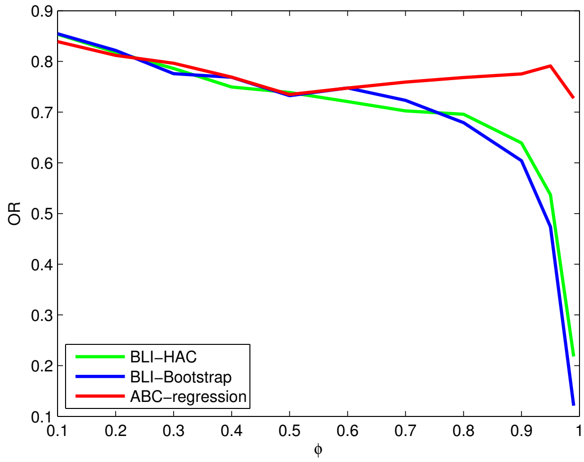


Figure 9: Overlapping Ratios of the Monte Carlo experiment: AR(1) process. Comparison among the ABC, HAC-BLI, Bootstrapping-BLI estimators, sample size=1000. Different autocorrelations on the horizontal axis.

for 200 and 500 observations. Also, under the long samples, the gap between the estimators is still in favour of the ABC. Overlapping Ratios of the 90% credible intervals of the approximate posterior distributions and the Full likelihood posterior distribution are compared. The results are respectively reported in Tables 5,6,7.

Overall, ABC methods outperforms BLI in approximating the full likelihood posterior distribution under the three different sample sizes.

4 A real life application: a DSGE with ZLB

In the reminder of the paper, I estimate a non-linear DSGE with an occasionally binding Zero Lower Bound (Fernández-Villaverde et al. [2015]).

The Zero Lower Bound is often pointed as one of the main amplification factors of the financial crisis, helping to explain the magnitude of the downturn observed during the Great Recession and the slow recovery. When the monetary policy is constrained, economies are more sensitive to demand shocks since interest rates cannot be lowered to counteract the negative demand shocks. Despite the centrality of this topic in contemporary policy analysis, few papers tackle the estimation of new-Keynesian models featuring the Zero Lower bound. [10]

This paper borrows a standard new-Keynesian model with an occasionally binding positivity constraint by Fernández-Villaverde et al. [2015]. A household maximizes her utility consuming and providing labour (the unique productive factor) to intermediate firms that operate in monopolistic competition and readjust prices according to Calvo type of contracts. The differentiated products are then assembled by retail firms operating in perfect competition. The models equations are housed in the Appendix.

The presence of the lower bound generates state-dependency in the solution, due to the fact that agents consider that monetary policy will not be able to set the interest rate below its effective lower bound. In our case, the presence of the occasionally binding constraint generates two alternative regimes: a slack regime where the monetary policy sets the interest rate in accordance with the Taylor rule, and a constrained one where the ZLB binds. In order to take this non-linearty into account, the model is solved by Piecewise linear approximation Guerrieri and Iacoviello [2015].[11]

[10]Most of the estimated models use samples which exclude the Zero Lower Bound (Christiano et al. [2015], Aruoba et al. [2013]. Gust et al. [2012] estimate a new-Keynesian model with a binding constraint on the interest rate using the particle filter. Their sample contains three observable variables to make inference on the structural parameters. They solve the model with a fully non-linear method.

[11]The piecewise linear apporximation is a solution method that allows handling occasionally

In the reminder of the section, first, I present the main issues related to the use of the methods of moments in in this type of models and show why ABC is the natural candidate to tackle those issues. Second, I apply the ABC method to estimate the DSGE.

4.1 Methods of moments and the ZLB: some challenges

The estimation of non-linear DSGE by methods of moments encounters one practical limitation: mapping functions between the parameters and the moments are not necessarily as regular as in the linear versions. This lack of regularity can cause two main problems: 1) identification issues for the parameters that control the degree of non-linearity of the model; 2) computational problems during the minimization of the objective function. The identification issue can be partially eased by the use of conditional moments (moments computed conditionally on the state of the economy). The computational problems can be tackled by the use of Bayesian methods (ABC, BLI). Also, the use of conditional moments amplifies the small sample issue. To this extent ABC appears as the natural candidate to be used, given its good small sample performance relative to the alternative Bayesian methods, as seen in section 3.

In linear models, variable dynamics are not affected by the state of the economy. This allows computing moments unconditionally on the state of the economy (i.e. *unconditional moments*). Instead, in non-linear solutions, dynamics are affected by the state of the economy. To this extent, unconditional moments convey state-dependent dynamics in a unique object. This can generate irregular and strongly-non monotonic mapping functions, causing substantial identification issues in the estimation. For an illustration of this point, let us consider the case where we want to make an estimate by matching the unconditional moments of the model with the ones observed: i.e. moments, as covariances and variances computed unconditional on the state of the economy. In fig. 10, I report the covariance between income

binding constraints. With repsct to alternative non-linear solution method (value function iteration, Chebyshev polynomial) the solution method reduces the compuational burden expecially in case of medium-size DSGE models. The piecewise linear solution allows to obtain a large number of simulations and tackles the curse of dimensionality encountered when dealing with medium-scale models. The piecewise solution method delivers a first order perturbation solution in a piecewise fashion. The solution is not just the juxtaposition of two linear solutions: the policy coefficients depend on how long the regime is expected to last. How long the model lasts is influenced by the state vector. This feedback effect can produce an important non-linearity. However, a drawback of this solution method is that it assumes that agents do not expect future shocks hitting the economy in the following periods. Hence precautionary savings are not considered. To solve the model two conditions must be met. First, Blanchard-Khan conditions must hold in the reference regime. Second, if the shocks hitting the economy move the model away from the reference regime to the alternative regime, in absence of future shocks the model must return to the reference regime.

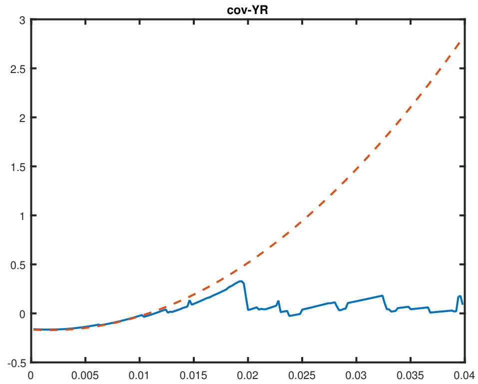


Figure 10: y axis: Covariance between income and interest rate. x-axis: standard deviation of the preference shock. The dashed line represent the covariance when the solution of the model is linear. The continuous line is the covariance when the model is solved non-linearly

and interest rate, with respect to the standard deviation of the preference shock (i.e. a demand shock). On the horizontal axis, I report the standard deviation of the preference shock. The dashed line represent the covariance when the solution of the model is linear. The solid line is the covariance when the model is solved non-linearly. The figure shows that the when the model is linear, the function relating the parameter to the covariance is regular and monotonic. Instead, when the model is solved non-linearly, the relation becomes less regular and non-monotonic. This irregularity derives from the fact when the standard deviation of the shock increases, the model hits the ZLB more often. Under ZLB, the covariance between the interest rate and income is close to zero. Therefore, when using this mapping function to build the objective function, this irregularity can lead to a strong identification issues (e.g. in this case leading to under-estimating the standard deviation of the model).

The use of conditional moments (Gospodinov and Otsu [2012]) can ease this problem. For example, moments can be computed and matched conditionally on the state of the economy: e.g. covariances and variances can be computed with respect to condition holding for the constraint (ZLB versus unconstrained monetary policy). By this fashion, we can convey the information provided by the data in a more efficient way. For example, this will decrease the underestimation obtained in the unconditional case.

A second issue generated by the use of non-regular mapping function is the non-regularity of the objective function itself. The use of conditional moments can only ease the problem, since the mapping functions themselves of the conditional moments remain less regular than the ones produced in the linear case. This is due to the fact that conditional moments can only partially capture the state-dependent dimension of the moments. Additionally, non-linear solution methods can also produce non-regular mapping functions due to possible instability in the solution algorithms. To this extent, frequentist methods, involving minimization of the objective functions, can encounter problems in finding the minimum of the objective function. For example, the minimization step can be severely affected by the choice of the initial point in the algorithm. In this case, ABC and BLI can ease the problem, avoiding the minimization step and studying the posterior distribution through Bayesian techniques.

Finally, it is worth to notice that the use of conditional moments can amplify the small sample bias, due to the fact that conditional moments are computed on subset of the sample, further decreasing the information on which the moments are computed. To this extent, ABC appears as the natural candidate to be used in the estimation with conditional moments, given its good small sample performance, as shown in the Monte Carlo exercises run in the Section 3.

4.2 Estimation

In this subsection, ABC-rejection is applied to the estimation of a new-Keynesian model with an occasional binding constraint on the zero lower bound. I focus on two main aspects: i) the importance of adding data observed during the ZLB in affecting the dynamics of the estimated model; ii) the role of different types of moments in exploiting data and update prior information. In presenting the results, I also compare the estimated model with its original calibrated version in Fernández-Villaverde et al. [2015].

In order to estimate the model, 30000 simulations are produced. Priors are standard and presented in table 8. Of the simulations, 0.33% with the smallest Euclidean distance are accepted. Moments are rescaled by their observed value. The observables are: consumption, output, interest rate, inflation and wages. Estimation is performed using different sample sizes: i) the benchmark estimation 1966Q1-2015Q4, ii) 1966Q1-2009Q1, iii) 1983Q1-2009Q1, iv) 1983Q1-2015Q4, v) 1994Q1-2015Q4. Except for the second and the third set, all the others include the period with the ZLB.

Estimation is performed using different combinations of moments. Each combination contains variances, covariances, and autocorrelations of order 1 and 2. Variances and covariances can be computed unconditionally (i.e. computed considering the whole sample) and conditionally (computed conditional on the monetary policy regime). Conditional moments are expected to better exploit the information provided by the data, given the different dynamics of economic variables when the ZLB binds. Finally, estimation is performed with and without including higher order moments (asymmetry and kurtosis). These moments are expected to provide useful information, given the asymmetries produced by the constraints.

Concerning the estimation across periods, including the period of ZLB in the model helps to replicate some important facts observed during and after the financial crisis: 1) a higher volatility of data with respect to the standard Great Moderation fluctuation; 2) a larger persistence of economic shocks; 3) a strong persistence of inflation despite the negative downturn and the long duration of the ZLB. Table 9 reports estimated parameters using the different sample sizes. Results across the different periods are robust. The estimates based on sample including the ZLB (1966Q1-2015Q4, 1983Q1-2009Q1, 1999Q1-2015Q4) are characterized by larger standard deviation for the preference shocks (σ_U=0.0017), whereas the same parameter is at 0.0012 and 0.0015 respectively for 1966Q1-2009Q1 and 1983Q1-2009Q1. The same holds for the autocorrelation parameter for the preference process: for the exercises excluding the period ZLB $\rho_U = 0.73$, whereas the same parameter is around 0.79 in the cases where the ZLB is included. Besides, concerning the parameter affecting the inflation dynamics -and its interaction with the real economy-, the estimates for the Great Moderation period are characterized by higher price stickiness ($\theta = 0.68$) and lower inflation reaction coefficient ($\phi_\pi = 2.15$), relatively to the estimated parameters obtained considering the other sample. When the ZLB is included, price stickiness is generally lower (around 0.63) and the inflation reaction coefficient is generally above 2.25. Concerning the reaction coefficient to output variations, the parameter $\phi_y = 0.10$, whereas it is higher for the other cases. Finally, estimates for the long term inflation are higher for the samples than without the ZLB ($\pi = 0.50$). No significant variations are found for the other parameters.

Concerning the role of the different moments in adding information, in Table 10 I compare the estimated parameters for four different combinations of moments by alternatively using: i) conditional moments (i.e. moments conditional on the monetary policy regime) or unconditional moments; 2) by including or excluding higher order moments (asymmetries and kurtosis). Overall, conditional moments play a more important role in identification with respect to the higher order moments. In particular, when conditional moments are used, autocorrelation ρ_U for the preference processes is higher, being around 0.78 in the first case and 0.73. The same holds for the standard deviation (around 0.0016 when conditional moments are used) and 0.0012 in the case of unconditional moments. Also, the use of conditional moments increases the estimates for the standard deviations of the monetary policy shock and for the productivity shocks. No significant differences emerge by adding higher order moments. These results play in favour of the use of conditional moments in the estimation in order to exploit information that is not efficiently used when using standard unconditional moments.

These estimated parameters ease one limitation of the New-Keynesian models with ZLB, highlighted by Fernández-Villaverde et al. [2015]: the variation of consumption when the economy hits the ZLB is too large with respect to what found in the data. To this extent, a larger inflation reaction coefficient limits the variation of the price dispersion and its negative consequences on the real economy. When the model is estimated using data related to the financial crisis and to ZLB, the estimation tries to match the moments observed with the one predicted by the model. With respect to the calibration, the estimated Taylor rule coefficient for inflation increases while the price stickiness decreases to help the model to reproduce a smaller fluctuation of consumption and output during the ZLB. In the calibrated model inflation persistence is mainly explained by the high Calvo parameter, whereas in the estimated model the stronger inflation persistence is more related to the monetary policy coefficients.

In Figure 11, we plot the impulse responses to a negative preference shock. This shock can be interpreted as an exogenous decrease in demand, pushing down output, inflation and interest rates. We apply the shock to the estimated model, according to the different sample sizes. The shock has the same size (0.0096) for all the estimated versions. For the sake of comparison, the size of the shock is the same as the one used by Fernández-Villaverde et al. [2015] to show the non-linear dynamics of their calibrated model. In figure 11, the impulse responses for interest rates, consumption, output and inflation are reported. Impulse responses are obtained for the calibrated version (in blue dotted line), for the model estimated with the sample including the Zero Lower bound (1966-2015, the solid green light), and for the model estimated excluding the Zero Lower Bound (1966-2009, in red dashed line). In the latter case, the shock sends the economy on the Zero Lower bound for 3 periods. In case the estimation sample includes includes the Zero Lower Bound, shock sends the interest rate to zero for four periods. The calibrated version features the longest spell, mainly due to the fact that in the calibrated version the auto-correlation parameter ρ_R of the Taylor rule equals zero, whereas it is positive in estimated versions. In the estimated version 1966-2015, a larger value for ϕ_π and a smaller value for the price stickiness θ make inflation and output more resilient in the period of ZLB with respect to the calibrated version: the percentage deviation of output, consumption and inflation is smaller for the estimated model compared to the ones obtained for the calibrated version.

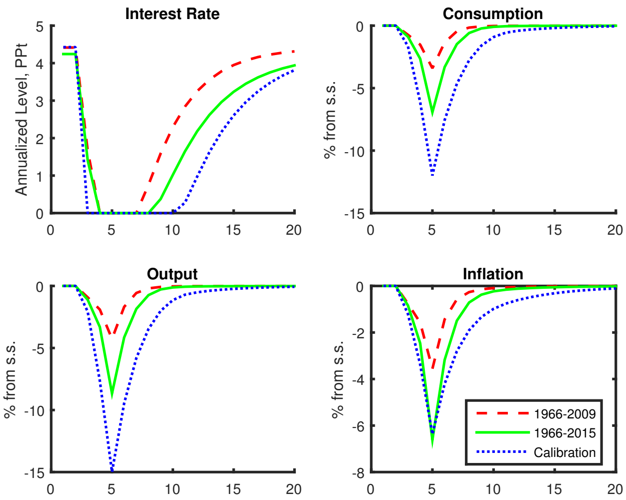


Figure 11: Impulse responses to a negative preference shock for interest rate (in levels), consumption, output, inflation (in percentage deviation from the steady state value). Impulses are reported for the calibrated version (dotted blue line) and for the estimated versions using: i) the sample 1966-2015 (red lines); ii) the sample 1966-2009 (dashed red line)

5 Conclusion

In this paper, I pledge for the use of Approximate Bayesian Computation techniques in the esimation of non-linear DSGE model.

In particular, through two Monte Carlo exercises, I show that ABC has a better small sample performance, compared to the other GMM-style estimators. In particular I show that ABC have a better small sample performance with respect to the Bayesian Limited Information Method (BLI), Kim [2002], the state-of-the-art Bayesian method of moments used in DSGE estimation. This result hinges on the fact that in ABC methods, the distribution of moments is simulated taking into account the actual size of the sample, rather than focusing on the asymptotic one, as it is done for the GMM-style estimators.

ABC is also tested on a real life application: a new-Keynesian model with a Zero Lower Bound. The non-linear solution displays some of the common challenges arising when applying the methods of moments in non-linear DSGE estimation. ABC is shown to deal with those issues. In particular, I show that the presence of non-linearities, as the ones generated by the Zero Lower Bound, can generate non-regular mapping functions between the structural parameters and the moments, limiting identification of the parameters controlling the degree of non-linearity. This issue can be eased by recurring to conditional moments, i.e. moments computed conditionally on the state of the economy. Another issue is related to the fact that non-regular mapping function are likely to generate non-regular objective functions, whose minimum can be harder to find. To this extent, the use of a Bayesian method of moments, as the ABC and BLI, can circumvent problems related to the minimization. Besides, ABC can also tackle the small sample issue amplified by the use of conditional moments. Besides, I estimate the model with simple ABC-rejection and I find three main results. First, including the period of the ZLB in the sample provides important information help the model to replicate the higher volatility of the macroeconomic variables, the persistence of inflation and of the Zero Lower Bound itself. Second, conditional moments tackle identification issue more than higher order moments, as such as asymmetry and kurtosis. Third, the estimated version of the model eases the problem identified in its calibrated version by Fernández-Villaverde et al. [2015], related to the predicted over-reaction of macroeconomic aggregates durig the period of ZLB. To this extent higher Taylor rule coefficient in response of inflation and lower price stickiness look crucial in reducing the the sensitivity of consumption and output to negative demand shocks inducing Zero Lower Bound.

References

Sungbae An and Frank Schorfheide. Bayesian analysis of dsge models. *Econometric reviews*, 26(2-4):113–172, 2007.

S Borağan Aruoba, Pablo Cuba-Borda, and Frank Schorfheide. Macroeconomic dynamics near the zlb: A tale of two countries. Technical report, National Bureau of Economic Research, 2013.

Stuart Barber, Jochen Voss, Mark Webster, et al. The rate of convergence for approximate bayesian computation. *Electronic Journal of Statistics*, 9(1):80–105, 2015.

Mark A Beaumont. Approximate bayesian computation in evolution and ecology. *Annual review of ecology, evolution, and systematics*, 41:379–406, 2010.

Mark A Beaumont, Wenyang Zhang, and David J Balding. Approximate bayesian computation in population genetics. *Genetics*, 162(4):2025–2035, 2002.

Gérard Biau, Frédéric Cérou, Arnaud Guyader, et al. New insights into approximate bayesian computation. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 51, pages 376–403. Institut Henri Poincaré, 2015.

Michael GB Blum. Approximate bayesian computation: a nonparametric perspective. *Journal of the American Statistical Association*, 105(491):1178–1187, 2010.

Michael GB Blum and Olivier François. Non-linear regression models for approximate bayesian computation. *Statistics and Computing*, 20(1):63–73, 2010.

Lawrence J Christiano, Roberto Motto, and Massimo Rostagno. Financial factors in economic fluctuations. 2010a.

Lawrence J Christiano, Mathias Trabandt, and Karl Walentin. Dsge models for monetary policy analysis. Technical report, National Bureau of Economic Research, 2010b.

Lawrence J Christiano, Martin S Eichenbaum, and Mathias Trabandt. Understanding the great recession. *American Economic Journal: Macroeconomics*, 7(1): 110–167, 2015.

Lawrence J Christiano, Martin S Eichenbaum, and Mathias Trabandt. Unemployment and business cycles. *Econometrica*, 84(4):1523–1569, 2016.

Michael Creel and Dennis Kristensen. Indirect likelihood inference. 2011.

Michael Creel and Dennis Kristensen. On selection of statistics for approximate bayesian computing (or the method of simulated moments). *Computational Statistics & Data Analysis*, 100:99–114, 2016.

Michael Creel, Jiti Gao, Han Hong, and Dennis Kristensen. Bayesian indirect inference and the abc of gmm. *arXiv preprint arXiv:1512.07385*, 2015.

Darrell Duffie and Kenneth J Singleton. Simulated moments estimation of markov models of asset prices, 1990.

Gauti B Eggertsson and Michael Woodford. Optimal monetary policy in a liquidity trap. Technical report, National Bureau of Economic Research, 2003.

Jianqing Fan and Qiwei Yao. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85(3):645–660, 1998.

Jesús Fernández-Villaverde and Juan F Rubio-Ramírez. Estimating macroeconomic models: A likelihood approach. *The Review of Economic Studies*, 74(4):1059–1087, 2007.

Jesús Fernández-Villaverde, Grey Gordon, Pablo Guerrón-Quintana, and Juan F Rubio-Ramirez. Nonlinear adventures at the zero lower bound. *Journal of Economic Dynamics and Control*, 57:182–204, 2015.

Jean-Jacques Forneron and Serena Ng. The abc of simulation estimation with auxiliary statistics. *arXiv preprint arXiv:1501.01265*, 2015.

Nikolay Gospodinov and Taisuke Otsu. Local gmm estimation of time series models with conditional moment restrictions. *Journal of Econometrics*, 170(2):476–490, 2012.

Jakob Grazzini, Matteo G Richiardi, and Mike Tsionas. Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control*, 77:26–47, 2017.

Luca Guerrieri and Matteo Iacoviello. Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, 70:22–38, 2015.

Christopher J Gust, David López-Salido, and Matthew E Smith. The empirical implications of the interest-rate lower bound. 2012.

Jae-Young Kim. Limited information likelihood and bayesian analysis. *Journal of Econometrics*, 107(1):175–193, 2002.

Jinill Kim and Sunghyun Henry Kim. Spurious welfare reversals in international business cycle models. *journal of International Economics*, 60(2):471–500, 2003.

Finn E Kydland and Edward C Prescott. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370, 1982.

Paul Marjoram, John Molitor, Vincent Plagnol, and Simon Tavaré. Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328, 2003.

Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.

Jonathan K Pritchard, Matthew Stephens, and Peter Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959, 2000.

José-Víctor Ríos-Rull, Frank Schorfheide, Cristina Fuentes-Albero, Maxym Kryshko, and Raül Santaeulàlia-Llopis. Methods versus substance: Measuring the effects of technology shocks. *Journal of Monetary Economics*, 59(8):826–846, 2012.

Francisco Ruge-Murcia. Estimating nonlinear dsge models by the simulated method of moments: With an application to business cycles. *Journal of Economic Dynamics and Control*, 36(6):914–938, 2012.

Francisco J Ruge-Murcia. Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics and Control*, 31(8):2599–2636, 2007.

Scott A Sisson, Yanan Fan, and Mark M Tanaka. Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6):1760–1765, 2007.

Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175, 2003.

Table 1: Prior distribution for the RBC parameters

Parameter	**Distribution**	**1**	**2**
β	*Beta*	0.95	0.02
γ	*Normal*	2	0.50
ρ_a	*Beta*	0.95	0.04
ρ_b	*Beta*	0.95	0.04
ρ_d	*Beta*	0.95	0.04
σ_a	*Gamma Inverse*	0.01	4
σ_b	*Gamma Inverse*	0.01	4
σ_d	*Gamma Inverse*	0.01	4

Prior distribution: Informative Prior

Table 2: RMSE, sample size=100 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.01395	0.04079	0.01812	0.01566	0.01609	0.27268	0.22648	0.12772
ABC-ker	0.01456	0.04394	0.01871	0.01596	0.01666	0.27522	0.22939	0.13000
ABC-OLS	0.01406	0.06961	0.02131	0.02157	0.02014	0.27040	0.26608	0.16532
ABC-regr	0.01415	0.07220	0.02195	0.02180	0.02079	0.27223	0.26811	0.16567
ABC-HC	0.01920	0.10839	0.02755	0.03006	0.02448	0.26240	0.28406	0.22597
BLI	0.03729	0.05116	0.04154	0.03172	0.02695	0.67502	0.87365	0.30317

RMSE obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 100 observations. Case: High peristency and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Table 3: RMSE, sample size=200 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.01231	0.05162	0.01738	0.01691	0.01543	0.25187	0.22934	0.10386
ABC-ker	0.01332	0.05151	0.01798	0.01763	0.01606	0.25104	0.23001	0.10843
ABC-OLS	0.01237	0.06588	0.02006	0.02174	0.02197	0.24180	0.26175	0.13565
ABC-regr	0.01269	0.06876	0.01998	0.02186	0.02150	0.24271	0.26266	0.14553
ABC-HC	0.01655	0.09258	0.02385	0.02764	0.02650	0.22664	0.26542	0.20675
BLI	0.03418	0.10956	0.04294	0.05040	0.02682	0.58462	0.72849	0.59571

RMSE obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 200 observations. Case: High peristency and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Table 4: RMSE, sample size=500 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.01093	0.05065	0.02034	0.01764	0.01665	0.22934	0.27019	0.13820
ABC-ker	0.01110	0.05388	0.02042	0.01761	0.01669	0.22629	0.26472	0.14590
ABC-OLS	0.01081	0.08508	0.01765	0.01776	0.02078	0.22260	0.28040	0.19021
ABC-regr	0.01068	0.08764	0.01759	0.01755	0.02098	0.22184	0.28060	0.19879
ABC+HC	0.01205	0.11679	0.01875	0.01930	0.02520	0.20846	0.27512	0.26526
BLI	0.03742	0.06969	0.03863	0.03228	0.04938	0.58462	0.93056	0.53128

RMSE obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 500 observations. Case: High peristency and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Table 5: OR100, sample size=100 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.60627	0.81372	0.75159	0.80626	0.70472	0.44003	0.50961	0.70015
ABC-ker	0.67059	0.87735	0.76656	0.80266	0.80791	0.49469	0.58173	0.76644
ABC-OLS	0.36064	0.81120	0.72085	0.72093	0.75838	0.28645	0.39473	0.75102
ABC-regr	0.35319	0.80143	0.72284	0.72073	0.75071	0.28332	0.38944	0.74532
ABC-HC	0.55281	0.73430	0.66426	0.63165	0.68748	0.40299	0.42275	0.71252
BLI	0.04772	0.88188	0.66390	0.33132	0.34849	0.05231	-0.03619	0.43781

Overlapping Ratio obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 100 observations. Case: High peristency and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Table 6: OR200, sample size=200 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.57967	0.79299	0.73966	0.79252	0.67809	0.42241	0.47965	0.69985
ABC-ker	0.65300	0.87342	0.77590	0.78717	0.79739	0.48200	0.54320	0.76554
ABC-OLS	0.29644	0.82433	0.69746	0.68340	0.72241	0.22776	0.34818	0.74757
ABC-regr	0.29253	0.81414	0.69455	0.67834	0.72318	0.22574	0.34598	0.74664
ABC-HC	0.51729	0.77607	0.68972	0.67147	0.69955	0.39251	0.40124	0.71773
BLI	0.31990	0.77961	0.66926	0.62000	0.19545	0.12139	0.26556	0.43072

Overlapping Ratio obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 200 observations. Case: High peristency and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Table 7: OR500, sample size=500 obs.

Methods	β	γ	ρ_a	ρ_b	ρ_d	σ_a	σ_b	σ_d
ABC-rej	0.50557	0.79718	0.68900	0.76053	0.64924	0.36181	0.44401	0.64670
ABC-ker	0.55337	0.86392	0.73684	0.76653	0.75746	0.40674	0.49992	0.73275
ABC-OLS	0.20349	0.79516	0.56611	0.58313	0.66112	0.13250	0.31733	0.69377
ABC-regr	0.20260	0.78619	0.56501	0.58256	0.65981	0.13104	0.31516	0.68630
ABC-HC	0.47321	0.76688	0.70070	0.74279	0.63753	0.34438	0.44529	0.67628
BLI	0.05993	0.83720	0.70892	0.31765	0.68671	0.11578	-0.04109	0.59651

Overlapping Ratio obtained in a Monte Carlo experiment, 100 repetitions. The sample contains 500 observations. Case: High persistence and Informative Priors. ABC-rej= ABC-rejection, ABC-ker=ABC-rejection + kernel weighting, ABC-OLS= ABC + OLS Regression Step; ABC-regr= ABC-regression with Local Linear Regression, ABC-HC=ABC-regression + Correction for Heteroskedasticity

Par	Prior Distr	Prior Mean	Prior St.Dev.
θ	*Beta*	0.7	0.1
ϕ_y	*Normal*	0.12	0.1
ϕ_π	*Normal*	2.0	0.5
ρ_R	*Beta*	0.75	0.1
π	*Uniform*	1.005	0.001
ρ_A	*Beta*	0.80	0.10
ρ_G	*Beta*	0.80	0.10
ρ_U	*Beta*	0.80	0.10
σ_A	*InvGamma*	0.005	0.01
σ_G	*InvGamma*	0.005	0.01
σ_M	*InvGamma*	0.005	0.01
σ_U	*InvGamma*	0.005	0.01

Table 8: Prior distribution for the estimation of the newkeynesian model with the occasionally binding ZLB

Sample	Benchmark: 1966Q1-2015Q4	1966Q1-2009Q1	1983Q1-2009Q1	1983Q1-2015Q4	1999Q1-2015Q4
ρ_A	0.806 (0.629 0.942)	0.812 (0.582 0.942)	0.796 (0.613 0.928)	0.804 (0.606 0.942)	0.802 (0.583 0.929)
ρ_G	0.784 (0.560 0.945)	0.798 (0.554 0.927)	0.794 (0.560 0.953)	0.811 (0.601 0.943)	0.802 (0.538 0.962)
ρ_U	0.786 (0.610 0.929)	0.737 (0.537 0.923)	0.738 (0.534 0.923)	0.793 (0.610 0.923)	0.799 (0.570 0.932)
ϕ_y	0.135 (-0.094 0.332)	0.154 (-0.011 0.348)	0.108 (-0.063 0.274)	0.139 (-0.056 0.367)	0.127 (-0.063 0.316)
ϕ_π	2.25 (1.626 3.238)	2.35 (1.524 3.248)	2.15 (1.358 3.200)	2.28 (1.610 3.207)	2.248 (1.471 3.050)
ϕ_R	0.654 (0.439 0.817)	0.645 (0.434 0.792)	0.685 (0.466 0.842)	0.627 (0.439 0.822)	0.636 (0.439 0.787)
σ_A	0.00314 (0.001 0.009	0.00311 (0.0007 0.0070)	0.0061 (0.0025 0.0134)	0.0030 (0.0008 0.0072)	0.0036 (0.0008 0.0099)
σ_G	0.0040 (0.0010 0.0135)	0.0042 (0.0008 0.0188)	0.0040 (0.0009 0.01225)	0.0041 (0.0010 0.015)	0.0036 (0.0009 0.0105)
σ_M	0.0019 (0.0008 0.0039)	0.0012 (0.0005 0.0021)	0.0014 (0.0007 0.0030)	0.0017 (0.0007 0.0034)	0.0017 (0.0007 0.0035)
σ_U	0.0017 (0.0007 0.0028)	0.0012 (0.0005 0.0020)	0.0015 (0.0006 0.0031)	0.0016 (0.0007 0.0027)	0.0017 (0.0007 0.0028)
θ	0.6311 (0.4401 0.7833)	0.5978 (0.4313 0.7960)	0.6886 (0.4591 0.8562)	0.6290 (0.4515 0.7854)	0.6497 (0.4653 0.8204)
$\bar{\pi}$	1.0045 (1.0030 1.0067)	1.0049 (1.0031 1.0068)	1.0050 (1.0031 1.0067)	1.0045 (1.0030 1.0067)	1.0046 (1.0030 1.0070)

Table 9: Estimate results for the structural parameters of the New-Keynesian model with the occasionally binding Zero Lower Bound.The estimates are obtained by usng the information provided by conditional moments and higher order moments. Estimates are reported for the different sample sizes: a) 1966Q1-2015Q4; b) 1966Q1-2009Q1; c) 1983Q1-2009Q1; d) 1983Q1-2015Q4; e) 1993Q1-2015Q4. 5% credible intervals are reported between parenthesis.

	Conditional + higher order moments		Conditional without higher order moments		Unconditional + higher order moments		Unconditional without higher order moments	
ρ_A	0.806		0.8064		0.7867		0.7871	
	(0.6296	0.9426)	(0.6296	0.9426)	(0.5452	0.9424)	(0.5452	0.9424)
ρ_G	0.7846		0.7825		0.8000		0.7938	
	(0.5607	0.9457)	(0.5607	0.9457)	(0.5238	0.9414)	(0.5167	0.9273)
ρ_U	0.7867		0.7874		0.7383		0.7338	
	(0.6102	0.9296)	(0.6102	0.9296)	(0.4511	0.8717)	(0.4511	0.8954)
ϕ_y	0.1359		0.1333		0.1549		0.1571	
	(-0.0944	0.3322)	(-0.0944	0.3322)	(-0.0289	0.3370)	(-0.01188	0.3251)
ϕ_π	2.2531		2.2505		2.3090		2.291	
	(1.6268	3.2384)	(1.6268	3.2384)	(1.5216	3.1488)	(1.4102	3.1141)
ϕ_R	0.6545		0.6556		0.6605		0.6670	
	(0.4396	0.8175)	(0.4396)	0.8270	(0.4344	0.8407)	(0.4344	0.8407)
σ_A	0.0031		0.0031		0.0022		0.0022	
	(0.0008	0.0090)	(0.0008	0.0090)	(0.0007	0.0044)	(0.0007	0.0049)
σ_G	0.0040		0.0041		0.0044		0.0044	
	(0.0010	0.01353)	(0.0010	0.0135)	(0.0010	0.0188)	(0.0010	0.01889)
σ_M	0.0019		0.0019		0.0012		0.0011	
	(0.0008	0.0039)	(0.0008	0.0039)	(0.0005	0.0022)	(0.0005	0.0019)
σ_U	0.0016		0.0017		0.0012		0.0012	
	(0.0007	0.0028)	(0.0007	0.0028)	(0.0005	0.0024)	(0.0005	0.0022)
θ	0.6311		0.6303		0.5984		0.5997	
	(0.4401	0.7833)	(0.4401	0.7833)	(0.4313	0.7570)	(0.4162	0.7570)
$\bar{\pi}$	1.0045		1.0045		1.0051		1.005	
	(1.0030	1.0067)	(1.0030	1.0067)	(1.0031	1.0067)	(1.0032	1.0067)

Table 10: Estimate results for the structural parameters of the New-Keynesian model with the occasionally binding Zero Lower Bound.The sample contains observations from 1966Q1 to 2015Q4. Estimates are reported for the different combinations of moments: a) Conditional moments and higher order moments; b) conditional moments without higher order moments; c) unconditional moments and higher order moments; d) unconditional moments. 5% credible intervals are reported between parenthesis.

Appendix 1: ABC refinements

ABC-regression

ABC-rejection is affected by the curse of dimensionality: to estimate a large set of parameters, we need to increase the number of summary statistics in the Euclidean distance computation. The probability of the simulated parameters to be accepted decreases and a higher number of simulations have to performed. This may have a huge impact on the feasibility of the estimation procedure. Besides, to increase the tolerance level can strongly compromise the approximation of the posterior distribution due to a larger simulation error. ABC-regression increases the efficiency of ABC through a post-sampling correction.

Two main refinements are introduced after the selection step:

- Each accepted simulation is assigned a weight according to its euclidean distance: the smaller the distance ρ_i, the larger the weight W_i. An Epanechnicov weighting function is generally used, but the algorithm is compatible with other kinds of kernel (normal, triangular and so forth). [12].
- The accepted parameters are corrected exploiting the result of a regression run after the selection step (hence the name ABC-regression). Each parameter is updated according to the result of a local linear regression of the accepted parameters on the discrepancies between simulated moments and observed ones (Beaumont [2010]).

In ABC regression (Beaumont et al. [2002]), ABC is equivalent to a problem of conditional density estimation, where a joint density distribution $P(\mathbf{s}_i, \theta_i)$ is updated through an accept-reject algorithm:

$$P(\theta|\mathbf{s}) = \frac{p(\mathbf{s}_i, \theta)}{I\left\{\rho|\mathbf{s}_i - \mathbf{s}| < \epsilon\right\}} \tag{6}$$

For this reason, conditional density estimation techniques (Fan an Gijbels, 1992) estimation are borrowed and incorporated in the ABC algorithms.

The ABC-regression pseudo-algorithm is:

- Draw θ_i from the prior $P(\theta)$;
- Simulate the model and obtain the observable variables $\mathbf{y}_i$;
- Compute the simulated moments $\mathbf{s}_i$ and the absolute standard deviation for each moment k_j;

[12]This correction coincides with the Indirect Likelihood Inference by Creel and Kristensen [2011]

- Compute the Euclidean distance for each simulation;

$$\rho|s_i, s| = \sqrt{\sum_{j=1}^{s}(s_i/k_j - s/k_j)^2} \tag{7}$$

- Select the tolerance level such that a fraction of the simulated parameters is accepted $P_\epsilon = N/M$;

- Each accepted draw is assigned a weight according to the Epanechnikov kernel:

$$K_\epsilon(\rho_i) = \begin{cases} \epsilon^{-1}(1 - (\frac{\rho_i}{\epsilon})^2) & \rho_i \leq \epsilon, \\ 0, & \rho_i > \epsilon; \end{cases}$$

- Apply a local linear regression to the linear model:

$$\theta_i = \alpha + (\mathbf{s}_i - \mathbf{s})'\beta + \epsilon_i, \tag{8}$$

  for $i = 1, ..., N$.

- Adjust the parameter given the results of the local linear regression:

$$\theta^* = \theta - (\mathbf{s}_i - \mathbf{s})'\hat{\beta}, \tag{9}$$

  which is equivalent to compute: $\theta_i^* = \hat{\alpha} + \hat{\epsilon}_i$.

The adjusted parameters associated to their kernel weights are random draws of the approximate posterior distribution. The initial part of the ABC-regression is the simple ABC rejection. The accepted parameters are corrected given two assumptions on the relation between the parameters drawn and the summary statistics simulated:

- Local linearity: a local linear relationship between the discrepancies of the moments and the parameters holds in the vicinity of the observed moment s such that the parameters can be expressed by the following equation:

$$\theta_i = \alpha + (\mathbf{s}_i - \mathbf{s})'\beta + \epsilon_i; \tag{10}$$

- Errors ϵ_i's have zero mean, are uncorrelated and homoskedastic.

In general, linearity only in the vicinity of $\mathbf{s}$ is a more palatable assumption than *global* linearity. In the local linear regression to estimate the coefficients for α and

β, the minimized object is:

$$\sum_{i=1}^{m} \left\{ \theta_i - \alpha - (\mathbf{s}_i - \mathbf{s})^T \beta \right\}^2 K_\delta(||s_i - s||). \tag{11}$$

In ABC literature, Epanechnikov kernel function is the one more common but others are feasible. In Eq.(11), the only difference with respect to the standard OLS is that the squared errors are weighted according to the distance ρ_i associated to the parameter θ_i. The solution is given by:

$$(\hat{\alpha}, \hat{\beta}) = (XWX)'(XW\theta), \tag{12}$$

where $X = (\mathbf{s}_i - \mathbf{s})$ for $i = 1, ..., N$ and W is a diagonal matrix, where each non zero element is $K_\delta(||s_i - s||)$.
The estimates for α and β are used in the adjustment step, through the adjustment equation 9. In conditional density estimation terms: $E[\theta|\mathbf{s}_i = \mathbf{s}] = \alpha$. The posterior mean coincides with the Nadaraya-Watson estimator (Nadaraya [1964], Watson, 1964), as suggested by Blum and François [2010] :

$$\alpha = \frac{\sum_i \theta_i^* K_\delta(||(\mathbf{s}_i - \mathbf{s}||)}{\sum_i K_\delta||(\mathbf{s}_i - \mathbf{s})||)}. \tag{13}$$

Blum and François [2010] add further step: a *correction for heteroskedasticity* in the adjustment step with non-linear regression in lieu of the local linear regression. For the sake of simplicity, here the local linearity assumption is maintained allowing the variance of the errors to change with the moments Beaumont [2010]. The heteroskedastic version is:

$$\theta_i = \alpha + (\mathbf{s}_i - \mathbf{s})'\beta + \epsilon_i = \alpha + (\mathbf{s}_i - \mathbf{s})'\beta + \sigma_i \xi_i, \tag{14}$$

where σ_i^2 is the variance of the error conditional on observing the simulated moments $Var[\theta|\mathbf{s}_i]$ and $\xi_i \sim N(0, 1)$. In this new procedure (ABC-regression with correction for heteroskedasticity) estimates α and β remain the same while in a further step the conditional variance for each draw is estimated. Finally, the correction mechanism is applied.

Blum and François [2010] model the conditional variance on the moments discrepancy by a second local linear model, borrowing from Fan and Yao [1998]. A second local linear regression is run and the conditional variance for each draw σ_i is estimated :

$$log(\epsilon_i)^2 = \tau + (\mathbf{s}_i - \mathbf{s})'\pi + \upsilon_i, \tag{15}$$

where υ_i is *iid* with mean zero and common variance. In this second local linear regression, the following object is minimized:

$$\min \left\{ log(\hat{\epsilon}_i)^2 - (\mathbf{s}_i - \mathbf{s})'\pi \right\} K_\delta(||\mathbf{s}_i - \mathbf{s}||), \tag{16}$$

where $\hat{\epsilon}_i$'s are the heteroskedastic errors estimated in the first regression. The variance conditional on the observed moments is $\sigma^2 = Var[\theta|s]$ is obtained according to

$$\hat{\sigma} = \hat{\tau}, \tag{17}$$

while the the variance conditional on each *simulated* moments is

$$\hat{\sigma}_i = \hat{\tau} + (s_i - s)'\hat{\pi}. \tag{18}$$

The Values obtained in equation 18 are used in the new post-sampling correction (equation 19) where the magnitude of each *heteroskedastic* error ϵ_i is corrected by the estimated standard deviation $\hat{\sigma}_i$:

$$\theta^* = \hat{\alpha} + \frac{\hat{\sigma}}{\hat{\sigma}_i}\hat{\epsilon}_i. \tag{19}$$

When the associated variance is higher than the variance conditional on the observed moments, the ratio $\frac{\hat{\sigma}}{\hat{\sigma}_i}$ is lower than 1 and the magnitude of the correction will be decreased with respect to the estimated $\hat{\epsilon}_i$.

ABC-regression allows to increase the tolerance level (i.e. increase the fraction of accepted simulations), making the algorithm computationally more efficient. Nonetheless, when the dimensionality of the parameters increases, the algorithm can deliver unstable results. Besides, some problems in the adjustment step can arise when the local linearity assumption does not hold: when the observed moments lie at the boundary of the simulated moments, adjusted values can be updated outside the support of the prior distribution (extrapolating rather than interpolating). Some refinements have been found by the literature to fix this problem, but a general consensus has not been reached.

Before adopting ABC-regression, drawing scatter plots can be useful to assess the informativeness of the moments regard the parameters to infer. In particular, (local) linear relations between moments and parameters can be found.

ABC-MCMC

ABC-MCMC methods draw parameters from a distribution closer to the posterior. This increases the acceptance rate of the algorithm. The algorithm developed by Marjoram et al. [2003] is the following:

- For $t = 0$, Draw $\theta \sim \pi(\theta)$;
- For $t \geq 1$ draw from:
$$\theta^{'} \sim K(\theta|\theta^{t-1}); \tag{20}$$
- Simulate and produce the moments conditional on θ^t;
- If $\rho(S(x), S(y)) < \epsilon$
  - Draw $u \sim U(0, 1)$,
  - If
$$u \leq \frac{\pi(\theta')}{\pi(\theta)^{t-1}} \frac{K(\theta^{t-1}|\theta')}{K(\theta'|\theta^{t-1})} \tag{21}$$
    then, $\theta^t = \theta^{'}$; otherwise $\theta^t = \theta^{t-1}$

  otherwise $\theta^t = \theta^{t-1}$

The MCMC produced by the algorithm is an approximation of the posterior distribution. Problems associated with ABC-MCMC are mainly related to presence of multimodality and mixing problems.

ABC-Sequential Monte Carlo

ABC-SMC nests ABC into the structure of a SMC technique: in an initial step, the vectors of parameters are drawn from a proposal distribution and a first selection is done on the basis of the Euclidean distance. In the following steps, the distribution of accepted parameters is perturbed and iteratively used in new simulation and selection steps, until convergence to the target distribution.

At each iteration, the accepted particles are perturbed according to a Kernel function. Each particle is accepted or rejected according to the Euclidean distance, choosing a decreasing tolerance level for each iteration such that $\epsilon_t \leq \epsilon_{t-1}$. If accepted, the particle is assigned a weight according to the Kernel function. A resampling procedure is envisaged to avoid sample degeneracy (i.e. few particles ending up hoarding much of the weight).

The algorithm is the following:

1. Initialize the tolerance level sequence: $\epsilon_1, \epsilon_2, \epsilon_3 ... \epsilon_T$ and select a sampling distribution μ_i. Set the iteration indicator $t = 1$.

2. Set the particle indicator $i = 1$ and:

   - If $t = 1$, draw the swarm of particles $\{\theta_1 \theta_2 ... \theta_N\}$ from the importance distribution μ_1.
   - If $t > 1$, sample the new swarm $\left\{\theta^{**}_{i,t-1}\right\} i = 1^N$ with weights $\left\{W^{**}_{i,t-1}\right\}_{i=1}^{N}$ ad perturb each particle according to a transition kernel $\theta^{**} \sim K_t(\theta|\theta^*)$

3. Simulate the model to obtain x^{**} conditional on each particle : if $\rho(S(x^{**}, S(x_0)) < \epsilon_t$ accept the particle, otherwise reject.

4. If accepted, assign the particle a weight:

   - If $t = 1$, $W_{i,1} = \frac{\pi(\theta_{i,1})}{\mu_1(\theta_{i,1})}$.
   - If $t > 1$,
   $$W_{i,t} = \frac{\pi(\theta_{i,t})}{\sum_{j=1}^{N} W_{t-1}(\theta_{t-1,j}) K_t(\theta_{t,i}|\theta_{t-1,j})} \quad (22)$$
   where $\pi(\theta)$ is the prior distribution for θ.

5. Normalize the weights such that $\sum_{i=1}^{N} W_{t,i} = 1$.

6. Compute the Effective Sample Size (ESS):
   $$ESS = \left[\sum_{i=1}^{N} (W_{t,i})^2)\right]^{-1} \quad (23)$$
   If the ESS is below $N\frac{1}{2}$, resample with replacement the particles according to the weights $\{W_{i,t}\}_{i=1}^{N}$ and obtain the new population with new weights $W_{t,i} = \frac{1}{N}$.

7. If $t < T$, return to (2).

ABC-SMC exhibits two interesting properties in the context of non-linear DSGE estimation. First, ABC-SMC is able to explore the whole support of the distribution, also in case of multimodality, which can arise in non-linear models. Second, it eases the computational inefficiency in case of significant mismatch between prior and posterior.

Appendix 2: the Real Business Cycle Model.

The households maximize the following expected sum of the utility functions:

$$maxE_t\left(\sum_{t=0}^{\infty}\beta^t\left(lnC_t - B_t\frac{H_t^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}}\right)\right), \tag{24}$$

subject to the budget constraint:

$$C_t + I_t = W_tH_T + D_tR_tK_t. \tag{25}$$

E_t stands for the expectation operator, C_t is the consumption, H_t are the hours offered by each household, B_t is the shock to the preference, namely the labour supply (Ríos-Rull et al. [2012]) and D_t is the shock to the interest rate requested by the household like in Smets and Wouters [2003]. β is the subjective discount factor and ν is the Frisch elasticity. Capital K_t is cumulated according to the following rule:

$$K_{t+1} = (1-\delta)K_t + I_t, \tag{26}$$

where δ is the depreciation rate and I_t is the investment. Firms choose how much capital and hours to employ in the production function given the technology A_t:

$$Y_t = A_tK_t^{\alpha}H_t^{1-\alpha}. \tag{27}$$

The market clearing is defined by:

$$Y_t = C_t + I_t. \tag{28}$$

The economy is subject to the following three structural shocks:

$$log(A_{t+1}) = \rho_a log(A_t) + \sigma_a\epsilon_a, \tag{29}$$

$$log(B_{t+1}) = \rho_b log(B_t) + \sigma_b\epsilon_b, \tag{30}$$

$$log(D_{t+1}) = \rho_d log(D_t) + \sigma_d\epsilon_d. \tag{31}$$

The technology autoregressive process A_t is the AR(1) of the RBC literature (Kydland and Prescott [1982]). The shock on the preferences B_t perturbs the labour supply hitting the marginal rate of substitution between consumption and leisure (see Ríos-Rull et al. [2012]). The shock on D_t is a shock on the interest rate requested by the households and can be interpreted as a shock to the risk premium.(Smets and Wouters [2003]). The model equilibrium is obtained by the following equations:

$$H_t = \left(\frac{1}{B_t}\frac{W_t}{C_t}\right)^{\gamma}, \tag{32}$$

$$\frac{1}{C_t} = \beta\left(\frac{1}{C_{t+1}}\left((1-\delta) + D_{t+1}R_{t+1}\right)\right), \tag{33}$$

$$Y_t = C_t + I_t, \tag{34}$$

$$K_{t+1} = K_t(1-\delta) + I_t, \tag{35}$$

$$R_t = \alpha A_t K_t^{\alpha-1} H_t^{1-\alpha}, \tag{36}$$

$$W = (1-\alpha) A_t K_t^{\alpha} H_t^{-\alpha}, \tag{37}$$

$$Y_t = A_t K_t^{\alpha} H_t^{1-\alpha}. \tag{38}$$

Eq. 32 expresses the intratemporal choice between consumption and leisure, Eq. 33 is the Euler Equation. Eqs. 34 and 35 are respectively the resource constraints and market clearing conditions completing the equilibrium of the model. Eq. 35 is the law of motion of capital and Eq.s 36, 37, 38 are the exogenous processes.
The experiment adopts an informative prior distribution similar to the one used by Ríos-Rull et al. [2012] to estimate a state-of-the-art Real Business Cycle. Informativeness in the prior distribution eases the identification issues associated to the preferences parameters.

Appendix 3: the new-Keynesian DSGE model with occasionally binding Zero Lower Bound.

Households maximise the following utility function separable in consumption c_t and labour l_t.

$$\sum_{i=0}^{\infty}\left(\prod_{i=0}^{t}\beta_i\right)\left\{logc_t - \psi\frac{l_t^{1+\phi}}{1+\phi}\right\}, \tag{39}$$

where ϕ is the inverse of the Frisch labour supply elasticity and β_t is the subjective discount factor subject to stochastic fluctuations around the mean β:

$$\beta_{t+1} = \beta^{1-\rho_b}\beta_t^{\rho_b}exp(\sigma_b\epsilon_{b,t+1}), \tag{40}$$

with $\epsilon_{b,t+1} \sim N(0,1)$. ρ_b and σ_b are respectively the autocorrelation and the standard deviation of the AR(1) process. The household maximizes her utility subject to the budget constraint:

$$c_t + \frac{b_{t+1}}{p_t} = w_t l_t + R_{t-1}b_t/p_t + T_t + F_t, \tag{41}$$

where b_t is a nominal government bond that pays a nominal interest rate R_t. p_t is the price level, whereas T_t and F_t are respectively the lamp sum taxes and the profits of the firms. Retail firms reassemble intermediate goods y_{it} and the technology:

$$y_t = \left(\int_0^1 y_{it}^{\frac{\epsilon-1}{\epsilon}}di\right)^{\frac{\epsilon}{\epsilon-1}}, \tag{42}$$

with ϵ is the elasticity of substitution. Final producers maximize their profit taking into account intermediate goods prices p_{it}, final prices p_t. The demand for each good will follow:

$$y_{it} = \left(\frac{p_{it}}{pt}\right)^{-\epsilon} y_t, \tag{43}$$

and the price of the final good will be equal to:

$$p_t = \left(\int_0^1 p_{it}^{1-\epsilon}di\right)^{\frac{1}{1-\epsilon}}. \tag{44}$$

The wholesale firms operate according to the production function:

$$y_{it} = A_t l_{it}, \tag{45}$$

where the productivity A_t evolves according to the law of motion:

$$A_t = A^{1-\rho_A} A_{t-1}^{\rho_A} exp(\sigma_A \varepsilon_{A,t}), \tag{46}$$

with $\varepsilon_t \sim N(0,1)$. The marginal costs are $mc_t = \frac{w_t}{A_t}$.

The firms choose their price according to a Calvo rule, where each period just a fraction $1-\theta$ firms can re-optimize their prices p_{it}. Firms will choose their price to maximize the profits:

$$\max_{p_{it}} E_t \sum_{\tau=0}^{\infty} \theta^{\tau} \left(\prod_{i=0}^{\tau} \beta_{t+1} \right) \frac{\lambda_{t+1}}{\lambda_t} \left(\frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau} \right) y_{it+\tau}, \tag{47}$$

s.t.

$$y_{it} = \left(\frac{p_{it}}{p_t} \right)^{-\epsilon} y_t, \tag{48}$$

where λ_{t+s} is the Lagrangian multiplier for the household in period $t+s$. Two auxiliary $x_{1,t}$ and $x_{2,t}$ are used to define the solution to the maximization problem:

$$\epsilon x_{1,t} = (1 - \epsilon x_{2,t}), \tag{49}$$

$$x_{1,t} = \frac{1}{c_t} mc_t y_t + \theta E_t \beta_{t+1} \Pi_{t+1}^{\epsilon} x_{1,t+1}, \tag{50}$$

$$x_{2,t} = \frac{1}{c_t} \Pi_t^* y_t + \theta E_t \beta_{t+1} \Pi_{t+1}^{\epsilon-1} \frac{\Pi_t^*}{\Pi_{t+1}^*} x_{2,t+1} = \Pi_t^* \left(\frac{1}{c_t} y_t + \theta E_t \beta_{t+1} \frac{\Pi_{t+1}^{\epsilon-1}}{\Pi_{t+1}^*} x_{2,t+1} \right), \tag{51}$$

where $\Pi_t^* = \frac{p_t^*}{p_t}$. Inflation dispersion will be equal to:

$$1 = \theta \Pi_t^{\epsilon-1} + (1-\theta)(\Pi_t^*)^{1-\epsilon}. \tag{52}$$

The government sets the nominal interest rate:

$$R_t = max \left[R_t, 1 \right], \tag{53}$$

with the notional interest rate Z_t:

$$Z_t = R^{1-\rho_r} R_{t-1}^{\rho_r} \left[\left(\frac{\Pi_t}{\Pi} \right)^{\phi_\pi} \left(\frac{y_t}{y_{t-1}}) \right)^{\phi_y} \right]^{1-\rho_r} m_t, \tag{54}$$

with m_t being the monetary policy *iid* shock $m_t = exp(\varepsilon_{m,t}\sigma_m)$, $\epsilon_{m,t} \sim N(0,1)$. The gross interest rate is equal to the notional interest rate as long it is larger than 1, since it cannot be set below 1 (the zero lower bound).

The government sets also the spending:

$$g_t = s_{g,t} y_t, \tag{55}$$

$$s_{g,t} = s_g^{1-\rho_g} s_{g,t-1}^{\rho_g} exp(\sigma_g \varepsilon_{g,t}), \tag{56}$$

with $\varepsilon \sim N(0,1)$. Since the agents are ricardian, we can set $b_t = 0$.

After aggregation we obtain:

$$y_t = \frac{A_t}{v_t} l_t, \tag{57}$$

with v_t is the loss of efficiency introduced by the price dispersion:

$$v_t = \int_0^1 \left(\frac{p_{i,t}}{p_t} \right)^{-\epsilon} di, \tag{58}$$

Moreover, following the Calvo pricing properties we can write:

$$v_t = \theta \Pi_t^{\epsilon} v_{t-1} + (1-\theta)(\Pi_t^*)^{-\epsilon}. \tag{59}$$

The Equilibrium

The Equilibrium is given by the sequence

$$\{y_t, c_t, l_t, mc_t, x_{1,t}, x_{2,t}, w_t, \Pi_t, \Pi_t^*, v_t, R_t, Z_t, \beta_t, A_t, m_t, g_t, b_t, s_{g,t}\}_{t=0}^{\infty}. \tag{60}$$

The equilibrium is defined by the following equations. The intertemporal and the intratemporal household F.O.Cs:

$$\frac{1}{c_t} = E_t \left\{ \frac{\beta_{t+1}}{c_{t+1}} \frac{R_t}{\Pi_{t+1}} \right\}, \tag{61}$$

$$\psi l_t^{\phi} c_t = w_t, \tag{62}$$

The solution of the maximization problem of the firms:

$$mc_t = \frac{w_t}{A_t}, \tag{63}$$

$$\epsilon x_{1,t} = (1 - \epsilon x_{2,t}), \tag{64}$$

$$x_{1,t} = \frac{1}{c_t} mc_t y_t + \theta E_t \beta_{t+1} \Pi_{t+1}^{\epsilon} x_{1,t+1}, \tag{65}$$

$$x_{2,t} = \frac{1}{c_t}\Pi_t^* y_t + \theta E_t \beta_{t+1} \Pi_{t+1}^{\epsilon-1} \frac{\Pi_t^*}{\Pi_{t+1}^*} x_{2,t+1} = \Pi_t^* \left(\frac{1}{c_t} y_t + \theta E_t \beta_{t+1} \frac{\Pi_{t+1}^{\epsilon-1}}{\Pi_{t+1}^*} x_{2,t+1} \right). \quad (66)$$

The government equations are:

$$R_t = max\left[R_t, 1\right], \quad (67)$$

$$Z_t = R^{1-\rho_r} R_{t-1}^{\rho_r} \left[\left(\frac{\Pi_t}{\Pi} \right)^{\phi_\pi} \left(\frac{y_t}{y_{t-1}}) \right)^{\phi_y} \right]^{1-\rho_r} m_t. \quad (68)$$

Inflation evolution and price dispersion:

$$1 = \theta \Pi_t^{\epsilon-1} + (1-\theta)(\Pi_t^*)^{1-\epsilon}, \quad (69)$$

$$v_t = \theta \Pi_t^{\epsilon} v_{t-1} + (1-\theta)(\Pi_t^*)^{-\epsilon}. \quad (70)$$

Market clearing conditions:

$$y_t = c_t + g_t, \quad (71)$$

$$y_t = \frac{A_t}{v_t} l_t. \quad (72)$$

The stochastic processes are:

$$\beta_{t+1} = \beta^{1-\rho_b} \beta_t^{\rho_b} exp(\sigma_b \epsilon_{b,t+1}), \quad (73)$$

$$A_t = A^{1-\rho_A} A_{t-1}^{\rho_A} exp(\sigma_A \varepsilon_{A,t}), \quad (74)$$

$$s_{g,t} = s_g^{1-\rho_g} s_{g,t-1}^{\rho_g} exp(\sigma_g \varepsilon_{g,t}), \quad (75)$$

$$m_t = exp(\varepsilon_{m,t} \sigma_m). \quad (76)$$

Observable equations

The observable equations are the following:

$$log\Delta GDP_t = 100(y_t - y_{t-1}) + \gamma, \quad (77)$$

$$log\Delta CONS_t = 100(c_t - c_{t-1}) + \gamma, \quad (78)$$

$$log\Delta WAGES_t = 100(w_t - w_{t-1}) + \gamma, \quad (79)$$

$$logHOURS_t = 100 l_t + \bar{l}, \quad (80)$$

$$log\Delta DEFL_t = 100 * (p_t * \pi + \pi - 1), \quad (81)$$

$$logFEDDUNDS_t = 100(exp(r_t) * RSS - 1) - 1. \quad (82)$$

Where Δ is the difference operator, $RSS = \frac{\pi}{\beta}$.