0.0.1 DSGE Model

Next, we report results for estimation of a simple DSGE model. Full likelihood-based estimation and inference of such models is complicated by unobserved state variables, which necessitate use of nonlinear filtering methods (Fernández-Villaverde and Rubio-Ramírez, 2005; An and Schorfheide, 2006). Also, models often contain fewer shocks than state variables, which can lead to stochastic singularities. Estimation by ABC avoids both of these difficulties. There is no need for filtering; we only need to simulate the chosen statistic, and stochastic singularities can be avoided by choosing (possibly with the aid of an algorithm) individual statistics that are not perfectly collinear. Our approach is related to Ruge-Murcia (2012) who employs the simulated method of moments (SMM) for the estimation of a DSGE model. Recall that SMM requires numerical optimization, which can be computationally demanding when the parameter space is large. In a simulation study, Ruge-Murcia (2012) treats a number of the parameters as known, while here we estimate all of the model's parameters.

The model that we consider is as follows: A single good can be consumed or used for investment, and a single competitive firm maximizes profits. The variables are: y output; c consumption; k capital; i investment, n labor; w real wages; r return to capital. The household maximizes expected discounted utility

$$E_t \sum_{s=0}^{\infty} \beta^s \left(\frac{c_{t+s}^{1-\gamma}}{1-\gamma} + (1 - n_{t+s})\eta_t \psi \right)$$

subject to the budget constraint $c_t + i_t = r_t k_t + w_t n_t$ and the accumulation of capital $k_{t+1} = i_t + (1 - \delta k_t)$. There is a preference shock, η_t, that affects the desirability of leisure. The shock evolves according to

$$\ln \eta_t = \rho_\eta \ln \eta_{t-1} + \sigma_\eta \epsilon_t \tag{1}$$

The competitive firm produces the good y_t using the technology $y_t = k_t^\alpha n_t^{1-\alpha} z_t$. Technology shocks z_t also follow an AR(1) process in logarithms: $\ln z_t = \rho_z \ln z_{t-1} + \sigma_z u_t$. The innovations to the preference and technology shocks, ϵ_t and u_t, are mutually independent i.i.d. standard normal random variables. The good y_t can be allocated by the consumer to consumption or investment: $y_t = c_t + i_t$. The consumer provides capital and labor to the firm, and is paid at the rates r_t and w_t, respectively.

Following Ruge-Murcia (2012), we estimate steady state hours, $\bar{n}$, along with the other parameters, excepting ψ, because it is comparatively easy to set priors on $\bar{n}$. Then ψ can be recovered using the estimates of the other parameters.The true parameters values are given in the fourth column of Table 3. True steady state hours, $\bar{n}$, is set to $1/3$ of the time endowment. The other parameters are set to values that are intended to be representative of the DSGE literature. Our pseudo-prior $\pi(\theta)$ is chosen as a uniform distribution over the hypercube defined by the bounds of the parameter space, which are found in columns 2 and 3 of Table 1. The chosen limits cause the pseudo-prior means to be biased for the true parameter values (see Table 1, column 5). The chosen limits are intended to be broad, so that the prior mean is quite uninformative as an estimator of the true parameter values (see Table 1, column 6). The DSGE literature sometimes makes use of fairly strongly informative priors, or fixes certain parameters (see Fernández-Villaverde, 2010, for discussion). Our intention here is to try to estimate all parameters of the model, using biased and weakly informative priors, to show that the estimation procedure is able to extract information about all parameters from the sample data.

Table 1: DSGE models, support of uniform priors.

Parameter	Lower bound	Upper bound	True value	Prior bias	Prior RMSE
α	0.2	0.4	0.330	-0.030	0.065
β	0.95	1	0.990	-0.015	0.021
δ	0.01	0.1	0.025	0.030	0.040
γ	0	5	2.000	0.500	1.527
ρ_z	0	1	0.900	-0.400	0.493
σ_z	0	0.1	0.010	0.030	0.042
ρ_η	0	1	0.700	-0.200	0.351
σ_η	0	0.1	0.005	0.040	0.049
$\bar{n}$	6/24	9/24	1/3	-0.021	0.042

The model is solved and simulated using Dynare (Adjemian et al., 2011), using a third order perturbation about the steady state. We assume, in line with much empirical work (e.g., Smets and Wouters, 2007; Guerrón-Quintana, 2010), that all variables except the capital stock are observed and available to use in the computation of statistics. The candidate auxiliary statistics include variable sample means, means of ratios of variables, standard deviations, coefficients of first order autoregressions for each variable in turn, and sample variances and covariances, across equations, of the residuals of first order autoregressions. The first order conditions of the model suggest some statistics that may be informative. For example, the model implies that $w = \psi\eta c^\gamma$, so

$$\log w = \log \psi + \gamma \log c + \log \eta, \tag{2}$$

where the preference shock $\log \eta$ follows an AR(1) process (see eq. 1). Because w and c are observable, equation 2 can be estimated, and the residuals of the estimated model may be used to construct estimators that should be informative for ρ_η and σ_η. In total, the set of candidate statistics has 40 elements. The statistics chosen for the final estimation were selected using the cross validation procedure of Creel and Kristensen (2015) The final set of selected statistics has 22 elements, and is summarized in Table 2.

Given the selected statistics, the ABC estimator is computed using the adaptive im-

Table 2: Selected statistics, DSGE model. For statistics 11-20, σ_{xy} indicates the sample covariance of the residuals of the AR1 models for the respective variables x and y.

Statistic	Description	Statistic	Description
1	$\widehat{\log\psi}$ from eq. 2	12	σ_{qq}
2	$\widehat{\gamma}$ from eq. 2	13	σ_{qn}
3	$\widehat{\rho}_\eta$, residuals of eq. 2	14	σ_{qr}
4	sample mean c	15	σ_{qw}
5	sample mean n	16	σ_{cc}
6	sample std. dev. q	17	σ_{cn}
7	sample std. dev. c	18	σ_{cr}
8	sample std. dev. n	19	σ_{cw}
9	sample std. dev. r	20	σ_{nn}
10	sample std. dev. w	21	σ_{ww}
11	estimated AR1 coef., r	22	$\overline{c/n}$

portance sampling methods described in Algorithms 2 and 3 of Creel and Kristensen (2015). The importance sampling distribution is generated separately for each Monte Carlo replication. Once the importance sampling distribution is generated, 5000 draws from the importance sampling distribution are made, to perform the final nonparametric fitting step.

The final nonparametric fitting step requires setting the bandwidths of the nonparametric fitting and quantile estimation procedures. We present two sets of results. The first results use bandwidths which were selected experimentally, separately for each parameter, to minimize out of sample RMSE and to optimize 90% confidence interval coverage, over 100 "true" parameter values which were drawn randomly from the prior. This is an entirely feasible procedure, which makes use of only pre-sample information. Then these bandwidths were used to do the nonparametric fitting and quantile estimation, using the 1000 Monte Carlo draws for the true parameter values given in Table 1. Software to perform all of these steps, and to replicate the Monte Carlo results reported here, is available at https://github.com/mcreel/ABCDSGE.

Table 3 gives the ABC estimation results for the 1000 Monte Carlo replications. We report results using local constant, local linear, and local quadratic (omitting cross products) nonparametric fits for the posterior mean. Results using the estimated posterior median are very similar, and are therefore not reported here. The table also gives the proportion of times that the true parameter values lie within the estimated 90% confidence interval, based upon nonparametric estimation of the 0.05 and 0.95 conditional quantiles, using a local constant nonparametric quantile estimator. We see that all versions of the ABC estimator reduce bias and RMSE considerably, compared to the prior biases and RMSEs given in Table 1. The local linear and local quadratic versions perform considerably better, overall, than does the local constant version. The magnitude of the biases of the local linear and local quadratic versions is small, compared to the true parameter values, in column 4 of Table 1. Between the local linear and local constant versions, performance is very similar, except that the local quadratic version has a bit less bias for several parameters. With regard to confidence interval accuracy, we see, in the 8th column of Table 3, that it is problematic. For the parameters σ_η and $\bar{n}$, confidence intervals are too narrow, on average, while for the parameters β, δ, γ, and ρ_η, they are too broad.

Table 3: DSGE model. Monte Carlo results (1000 replications). Bandwidths tuned using prior. LC=local constant, LL=local linear, LQ=local quadratic. 90% CI gives the proportion of times that the true value is in the 90% confidence interval.

	Bias			RMSE			90% CI
Parameter	LC	LL	LQ	LC	LL	LQ	LC
α	0.025	0.002	0.001	0.032	0.013	0.012	0.920
β	-0.008	0.001	0.001	0.010	0.003	0.003	0.993
δ	0.007	0.001	-0.000	0.011	0.004	0.003	0.991
γ	0.037	0.037	0.006	0.158	0.103	0.106	0.986
ρ_z	-0.012	-0.003	0.001	0.040	0.012	0.009	0.877
σ_z	-0.001	-0.001	-0.000	0.003	0.002	0.002	0.893
ρ_η	-0.007	-0.011	-0.009	0.054	0.047	0.049	1.000
σ_η	0.001	-0.000	0.000	0.003	0.002	0.001	0.834
$\bar{n}$	0.003	0.001	0.001	0.005	0.004	0.004	0.731

The results in Table 3 are based upon bandwidths that use no local information, as they were tuned using draws from the prior, which is biased and quite dispersed, given the true parameter values. In actual practice, one would prefer to use bandwidths that are tuned locally to the realized value of the statistic. One means of doing this is to do estimation exactly as was done to generate the results reported in Table 3, but then, given the realized estimate of the parameters, implement the experimental bandwidth tuning procedure using samples drawn at using the parameter estimate, rather than draws from the prior. This would provide a feasible, local, bandwidth tuning procedure. Unfortunately, such a procedure is too costly to implement within a Monte Carlo framework, though it is perfectly feasible when performing a single estimation for a real sample. As an approximation, we instead randomly draw 100 "true" parameter values from the 1000 Monte Carlo realized estimates from the first round, and implement the bandwidth tuning method using these. This gives a fixed set of bandwidths to use for each of a new set of Monte Carlo replications, rather that specific bandwidths for each Monte Carlo replication, which would be more desirable, but which is too costly to implement in the Monte Carlo context. Table 4 gives results for 1000 additional Monte Carlo replications, using bandwidths tuned in this way. We see that bias and RMSE are essentially the same as in Table 3, but that confidence interval coverage is considerably improved, overall, though still somewhat problematic for the parameters β, ρ_η and σ_η.

We also estimated true optimal bandwidths, by implementing the tuning procedure using 100 random samples generated at the true parameter values. When such bandwidths are used, 90% confidence interval coverage is correct, within expected error bounds, for all parameters. This procedure is of course not feasible outside of the Monte Carlo context, but it does confirm the theoretical result that confidence intervals have asymptotically correct coverage, and it lends support to performing local bandwidth tuning by drawing random samples at the first round ABC estimate, as this first round estimate is a consistent estimator of the true parameter.

Table 4: DSGE model. Monte Carlo results (1000 replications). Bandwidths tuned locally. LC=local constant, LL=local linear, LQ=local quadratic. 90% CI gives the proportion of times that the true value is in the 90% confidence interval.

	Bias			RMSE			90% CI
Parameter	LC	LL	LQ	LC	LL	LQ	LC
α	0.027	0.003	0.001	0.033	0.013	0.012	0.916
β	-0.008	0.001	0.002	0.011	0.003	0.003	1.000
δ	0.008	0.001	-0.000	0.011	0.004	0.003	0.900
γ	0.031	0.036	0.005	0.145	0.103	0.099	0.922
ρ_z	-0.013	-0.002	0.001	0.040	0.010	0.008	0.900
σ_z	-0.001	-0.001	-0.008	0.003	0.002	0.002	0.863
ρ_η	-0.010	-0.012	-0.010	0.054	0.046	0.049	0.794
σ_η	0.001	0.000	0.000	0.003	0.002	0.001	0.835
$\bar{n}$	-0.006	0.001	0.002	0.006	0.004	0.004	0.921

Given the simplicity and good performance of the ABC estimator, we believe that it provides an interesting alternative to the considerably more complex and computationally demanding methodology of MCMC combined with particle filtering, which can probably be described as the current state of the art for estimation of DSGE models. The practicality of estimation of a complex model using ABC is illustrated by the fact that we have been able to perform 2000 Monte Carlo replications of estimation of this simple DSGE model, using a single 32 core computer, in less than 72 hours. Once statistics and bandwidths have been selected (which are steps which can be performed before the sample data is available) it takes less than two minutes to perform estimation of the model. This final estimation step involves embarrassingly parallel computations (simulation and nonparametric regression), which means that ABC estimation as we have implemented it can be used for estimation of complex models in near real time.