

Blueschke, Dmitri; Savin, Ivan

Working Paper

Solving nonlinear stochastic optimal control problems using evolutionary heuristic optimization

Jena Economic Research Papers, No. 2013-051

Provided in Cooperation with:

Max Planck Institute of Economics

Suggested Citation: Blueschke, Dmitri; Savin, Ivan (2013) : Solving nonlinear stochastic optimal control problems using evolutionary heuristic optimization, Jena Economic Research Papers, No. 2013-051, Friedrich Schiller University Jena and Max Planck Institute of Economics, Jena

This Version is available at:

<https://hdl.handle.net/10419/98449>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



JENA ECONOMIC RESEARCH PAPERS



2013 – 051

Solving nonlinear stochastic optimal control problems using evolutionary heuristic optimization

by

**Ivan Savin
Dmitri Blueschke**

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

© by the author.

Solving nonlinear stochastic optimal control problems using evolutionary heuristic optimization*

I. Savin^{†‡} and D. Blueschke[§]

Abstract

Policy makers constantly face optimal control problems: what controls allow to achieve certain targets in, e.g., GDP growth or inflation? Conventionally this is done by applying certain linear-quadratic optimization algorithms to dynamic econometric models. Several algorithms extend this baseline framework to nonlinear stochastic problems. However, those algorithms are limited in a variety of ways including, most importantly, restriction to local best solutions only and the symmetry of objective function. In Blueschke *et al.* (2013a) a new flexible optimization method based on Differential Evolution is suggested. It allows to lift these limitations and achieve better approximations of the policy targets, but is designed to deterministic problems only. This study extends the methodology by dealing with stochastic problems in two different ways: applying extreme event analysis and by minimizing the median objective value. Thus, this research is aimed to broaden the range of decision support information used by policy makers in choosing optimal strategy under much more realistic conditions.

Keywords: Differential evolution; stochastic problems; nonlinear optimization; optimal control

JEL Classification: C54, C61, E27, E61, E63.

*Ivan Savin acknowledges financial support from the German Science Foundation (DFG RTG 1411). We also thank Reinhard Neck, Marco Tucci and Peter Winker for their helpful comments and suggestions.

[†]Faculty of Economics and Business Administration, Friedrich Schiller University of Jena, Germany

[‡]Chair for Econometrics and Statistics, Graduate School of Economics and Management, Ural Federal University, Mira 19, 620002 Yekaterinburg, Russian Federation

[§]Corresponding author, Universitaetsstrasse 65-67, University of Klagenfurt, Austria, dmitri.blueschke@aau.at

1 Introduction

Over the recent few decades, a large number of studies have tried to extend or rather to replace the traditional research methods by using different computer-based simulation techniques. These methods becoming increasingly popular include, among others, those areas of research like agent-based modeling, artificial neural networks and heuristic optimization. This paper follows the latter line of research and analyses the use of an evolutionary, heuristic approach in comparison to a more 'traditional' algorithm for optimal control of nonlinear stochastic problems.

The baseline framework in the field of optimal control problems is the linear-quadratic (LQ) optimization technique which is included in nearly all existent solution algorithms. There are many extensions of those methods for more sophisticated scenarios like considering nonlinear models and stochastic problems (e.g. Chow (1975), Chow (1981) and Kendrick (1981)). One of the algorithms which deals with these types of problems is OPTCON described in Matulka and Neck (1992) and Blueschke-Nikolaeva *et al.* (2012). However, the OPTCON algorithm still relies on the LQ optimization technique and, therefore, has some limitations typical for this framework.

One very important limitation which is common for nonlinear problems is the necessity of linearization of the problem, which, especially in the case of stochastic problems, requires some simplification assumptions and causes loss of information. In a recent work by Blueschke *et al.* (2013a) a new way of handling optimal control problems is analyzed. The authors test an evolutionary (so-called heuristic) approach for this purpose, namely Differential Evolution (DE, Storn and Price (1997)), which does not rely on LQ framework. The authors apply DE to optimal control problems in nonlinear dynamic economic systems with asymmetric objective function, where the 'classical' OPTCON algorithm does not work. They choose the DE method for this purpose because of its ability to explore complex search spaces with multiple local minima thanks to cooperation and competition of individual solutions in the DE's population, and the application easiness as it needs little parameter tuning. Applying the DE method increases computational time substantially but allows to get new insights into optimal control problems. In particular, some better approximations of the targets stated by policy makers are achieved. However, the work by Blueschke *et al.* (2013a) is designed to deterministic problems only. The present study extends this methodology and analyses the application of DE for stochastic problems.

Two alternative ways of dealing with stochastic nature of the problem are considered. On the one hand, applying DE for stochastic problems allows to run an 'extreme event analysis', where the outcomes of the best and worst

scenarios (in terms of the objective function value achieved) can be investigated. On the other hand or rather in addition, DE allows to minimize the expected objective value using a different selection criterion. While in the 'classical' optimization algorithms, where policy makers have only very limited information: an optimal strategy (usually only a local optimum) and an ex-post objective value, the present paper by combining the two alternatives broadens the range of decision support information by policy makers for the choice of the optimal strategy. Here we use the notation 'ex-post objective value' to highlight the following limitation of classical stochastic optimal control algorithms. An ex-post objective value means that the problem is solved as if it were stochastic, which leads to a set of optimal controls (u_t^*). In order to evaluate this stochastic solution (i.e. to calculate the final objective value) one applies in classical algorithms the deterministic model, which means that a *certain set of parameters* is used to calculate respective states (x_t^*). The final objective value (an 'ex-post objective value', as denoted in our paper) is then calculated for this resulting solution (x_t^*, u_t^*). In contrast, using the DE algorithm allows us to calculate both the 'ex-post objective value' for the best solution and an expected objective value, which takes the stochastic nature explicitly into account.

Possibility to calculate both, ex-post objective value and an expected objective value constitutes an important advantage also because, as we demonstrate, the ex-post objective function and the actual expected objective function (represented by the least median, see Section 3.2) are not 'associated': if one function monotonously decreases on a certain sequence of solutions, the other one will not necessarily do the same on this set of solutions. As a consequence, a solution most preferable by the ex-post function does not necessarily have the lowest expected value and the other way around. This raises the question on whether the classical approaches are well specified for the stochastic nonlinear problems and whether their solutions are not merely suboptimal ones? As we show below, DE outperforms OPTCON in minimizing the expected objective function for two different problems tested supporting our criticism in this regard.

The paper proceeds as follows. In Section 2 we define the class of problems to be tackled by the algorithms. Section 3.1 briefly reviews the OPTCON algorithm and describes its limitations. Section 3.2 introduces DE as an alternative strategy for solving nonlinear stochastic optimal control problems. In Section 4 results obtained for these two approaches based on two different models are stated. Section 5 concludes.

2 Problem description

The task is to solve an optimal control problem with a quadratic objective function (a loss function to be minimized) and a nonlinear multivariate discrete-time dynamic system under additive and parameter uncertainties. The intertemporal objective function is formulated in quadratic tracking form, which is often used in applications of optimal control theory to econometric models. It can be written as

$$J = E \left[\sum_{t=1}^T L_t(x_t, u_t) \right] \quad (1)$$

with

$$L_t(x_t, u_t) = \frac{1}{2} \begin{pmatrix} x_t - \tilde{x}_t \\ u_t - \tilde{u}_t \end{pmatrix}' W_t \begin{pmatrix} x_t - \tilde{x}_t \\ u_t - \tilde{u}_t \end{pmatrix}, \quad (2)$$

where x_t is an n -dimensional vector of state variables that describes the state of the economic system at any point in time t , u_t is an m -dimensional vector of controls, $\tilde{x}_t \in R^n$ and $\tilde{u}_t \in R^m$ are given 'ideal' (desired, target) levels of the state and control variables, respectively. T denotes the terminal time period of the finite planning horizon. W_t is an $((n + m) \times (n + m))$ matrix specifying the relative weights of the state and control variables (states and controls, henceforth) in the objective function. The W_t matrix may also include a discount factor α , $W_t = \alpha^{t-1}W$. W_t (or W) is symmetric.

The dynamic system of nonlinear difference equations has the form

$$x_t = f(x_{t-1}, x_t, u_t, \theta, z_t) + \varepsilon_t, \quad t = 1, \dots, T, \quad (3)$$

where θ is a p -dimensional vector of parameters that is assumed to be constant but unknown to the policy maker (parameter uncertainty), z_t denotes an l -dimensional vector of non-controlled exogenous variables, and ε_t is an n -dimensional vector of additive disturbances (system error). θ and ε_t are assumed to be independent random vectors with expectations, $\hat{\theta}$ and O_n , and covariance matrices, $\Sigma^{\theta\theta}$ and $\Sigma^{\varepsilon\varepsilon}$, respectively. f is a vector-valued function with $f^i(\dots)$ representing the i -th component of $f(\dots)$, $i = 1, \dots, n$. Solving an optimal control problem means, therefore, to find a certain set of controls $(u_1^*, u_2^*, \dots, u_T^*)$ which minimizes the objective function J , i.e. to find $u^* = \arg \min_u J$ with respect to (3).

At this point an important distinction should be made, namely the distinction between deterministic and stochastic systems. In the case of a deterministic problem, the decision-maker knows the functionality of the system for sure, i.e. without any uncertainties. In such a case the stochastic com-

ponents are set to zero: $\varepsilon_t, \Sigma^{\theta\theta}, \Sigma^{\varepsilon\varepsilon} = 0$. In the present study we explicitly consider the *parameter uncertainty* ($\Sigma^{\theta\theta} \neq 0$), which makes the optimization much more complex, but at the same time more realistic.

Later on in the present study applying the DE algorithm, the stochastic enters the system by different Monte Carlo realizations of parameter θ . For each possible set of controls different sets of states should be then calculated depending on these different θ as given by equation (3).

3 Optimization algorithms

3.1 OPTCON

The OPTCON algorithm determines approximate solutions to optimal control problems with a quadratic objective function and a nonlinear multivariate dynamic system under additive and parameter uncertainties. It relies on standard techniques of LQG framework¹ and combines elements of previous algorithms developed by Chow (1975) and Chow (1981), which incorporate nonlinear systems but no multiplicative uncertainty, and by Kendrick (1981), who deals with linear systems and all kinds of uncertainty. In our experiments we use the last version of the OPTCON algorithm, which is called OPTCON2. We skip the presentation of the OPTCON algorithm, which can be found in more detail in Matulka and Neck (1992) and Blueschke-Nikolaeva *et al.* (2012), and discuss only those issues relevant for this paper.

Due to the stochastic nature of the problems considered, optimal control becomes a challenging task. It is well known in stochastic control theory that a general analytical solution to dynamic stochastic optimization problems cannot be achieved even for very simple control problems. The main reason is the so-called dual effect of control under uncertainty, meaning that controls not only contribute directly to achieving the stated objective, but also affect future uncertainty and, hence, the possibility of indirectly improving the system performance by providing better information about the system (see, e.g., Aoki (1989) and Neck (1984)). Therefore, only approximations to the true optimum for such problems are feasible, with various schemes existing to deal with the problem of information acquisition.

As for linear problems, this field is well researched and there exist several algorithms which allow to find a reliable solution at least for problems under open-loop and open-loop feedback information patterns. A detailed description of those traditional solution concepts for the linear problems can be

¹Optimization of linear systems with Gaussian noise under a quadratic objective function.

found in Kendrick (2005). For nonlinear problems one can either use a linear approximation approach as it is done in the OPTCON algorithm or one can apply the perturbation method as shown, for example, in Cosimano (2008) and Benigno and Woodford (2012). For stochastic problems with adaptive control strategies Tucci *et al.* (2010) show that the problem is non-convex, which means that a local optimization method could be not appropriate. Some problems of the nonlinear stochastic problems are shown in Blueschke *et al.* (2013b), but there is even more to say concerning an extension to nonlinear problems. The OPTCON algorithm like all algorithms relying on LQG optimization technique solves the nonlinearity problem by a local linear approximation. This works in a stochastic case only under some simplifying assumptions. The most restrictive one among them is the assumption of the linear dependence of covariances in the linearized model from the covariances in the nonlinear model, which is required to be able to transfer the stochastic input of the nonlinear system (covariance matrix) into the linearized model. This implicates also a certain loss of information.

Another important limitation of the classical algorithms for nonlinear stochastic problems is the calculation of the expected 'cost-to-go function'. The algorithm minimizes the expected value of it, but only for the local linear approximation and, thus, there is no way to calculate the 'correct' expected objective function value of the original, nonlinear problem. As a result, only a very limited amount of information is available for the decision maker as output, namely an ex-post objective value calculated for the given (or rather expected) set of uncertain parameters.

Summarizing these limitations, there is a need for a method which can solve these problems and deliver a more informative solution of a nonlinear stochastic optimal control problem. We propose to use the Differential Evolution method for this purpose, as described below.

3.2 Differential Evolution

Thanks to the recent advances in computing technology, new nature-inspired optimization methods called heuristics have been developed. These methods are designed to provide ways of tackling complex optimization problems and detect global optima of various objective functions (eligible for certain constraints) in discrete and continuous search spaces. For an overview of these techniques see Gilli and Winker (2009).

Differential Evolution (DE), proposed by Storn and Price (1997), is a population based optimization technique for continuous objective functions. For applications of DE in finance, risk management and innovation management see Lyra *et al.* (2010), Winker *et al.* (2011) and Egbetokun and Savin

(forthcoming), respectively.

A detailed explanation on how DE deals with optimal control problem for a single parameter set θ (deterministic scenario) one can find in Blueschke *et al.* (2013a, p. 824-825). In short, starting with an initial population of random solutions (line 2 in Algorithm 1), DE updates this population by linear combination (line 7) and crossover (line 9) of four different solution vectors into one, and selects the fittest solutions among the original and the updated population. This continues until some stopping criterion is met.

Each member of the population (each candidate solution) contains all control variables for all time periods. Thus, each candidate $i = 1, \dots, p$ represents an alternative complete solution path for the whole optimum control problem, and is given as an $(m \times T)$ -matrix $P_{\dots,i}^{(1)} = (P_{j,t,i}^{(1)})_{\substack{j=1,\dots,m \\ t=1,\dots,T}} = (u_1^{(1),i}, u_2^{(1),i}, \dots, u_T^{(1),i})$, where $u_t^{(1),i}$ is an m -dimensional vector of controls and T – the size of the planning horizon.

It is important to mention that each candidate solution is also described by the time paths of corresponding state variables, which results from the dynamic system f , parameter set θ and the selected controls, i.e. $(x_{t=1,\dots,T}^{(1),i} = f(\dots, u_{t=1,\dots,T}^{(1),i}, \theta, \dots))$. For each candidate solution (for each set of control variables) and for each parameter set θ there is a unique set of state variables. These state variables are not directly included in a candidate solution but they contribute to the objective function which is to minimize.

Algorithm 1 Pseudocode for Differential Evolution

- 1: Initialize parameters m, T, p, F and CR
 - 2: Randomly initialize $P_{j,t,i}^{(1)}$, $j = 1, \dots, m$; $t = 1, \dots, T$; $i = 1, \dots, p$
 - 3: **while** the stopping criterion is not met **do**
 - 4: $P^{(0)} = P^{(1)}$
 - 5: **for** $i = 1$ to p **do**
 - 6: Generate $r_1, r_2, r_3 \in 1, \dots, p$, $r_1 \neq r_2 \neq r_3 \neq i$
 - 7: Compute $P_{\dots,i}^{(v)} = P_{\dots,r_1}^{(0)} + F \times (P_{\dots,r_2}^{(0)} - P_{\dots,r_3}^{(0)})$
 - 8: **for** $j = 1$ to m and $t = 1$ to T **do**
 - 9: **if** $u < CR$ **then** $P_{j,t,i}^{(n)} = P_{j,t,i}^{(v)}$ **else** $P_{j,t,i}^{(n)} = P_{j,t,i}^{(0)}$
 - 10: **end for**
 - 11: **if** $J(P_{\dots,i}^{(n)}) < J(P_{\dots,i}^{(0)})$ **then** $P_{\dots,i}^{(1)} = P_{\dots,i}^{(n)}$ **else** $P_{\dots,i}^{(1)} = P_{\dots,i}^{(0)}$
 - 12: **end for**
 - 13: **end while**
-

To take the stochastic nature of the problem into account some additional Monte Carlo simulations are required. In particular, a population of different θ s is generated representing alternative possible realizations of parameters.

In the following we apply two approaches on how to take these stochastic component into account: extreme event analysis and least median objective value analysis.

3.2.1 Extreme event analysis

The first alternative of solving a stochastic optimal control problem using DE one can denote as 'extreme event analysis'. The idea is to create a certain number of randomly drawn (Monte Carlo) possible realizations of uncertain parameters θ , MC_DE (default $MC_DE = 1000$).² 1000 θ realizations is considered to be, on the one hand, representative population of Monte Carlo draws, and on the other hand, manageable from the computational point of view (more details on cpu time required in Section 4). To generate those 1000 Monte Carlo draws a normal distribution is used. This choice is mainly driven by the well established properties of the distribution and its still big popularity in literature. Once there exists more evidence that for this specific problem a different distribution shall be used, it can be easily implemented in the code.³

For example, for the SLOVNL model containing 15 uncertain parameters (for more details see Appendix) we create a matrix of the size 1000×15 of possible realizations of θ 's values. For each of these 1000 sets we run an independent optimization using DE. The motivation behind is to choose the best policy for the case if the stochastic parameters occurs to have a specific realization. This alternative has the disadvantage to handle the system for each set of parameters θ as if it were deterministic. But it also allows to gain insights of what could happen, e.g., in the best, median, 20%-best quantile, or worst scenario (in terms of the objective function value achieved).

3.2.2 Least median objective value analysis

Since in the extreme event analysis only single Monte Carlo draws of θ are used, it only provides information on special cases (i.e. lowest parameter values for some of the variables in (3)) and does not allow to provide any statement on how DE can solve the model at hand with, strictly speaking, uncertain θ . To this end, we take the same set of random θ s for each member of the DE population and minimize not a single objective value (for a single θ draw), but their expected value over the whole set of θ s. Thus, in each

²To get the Monte Carlo realizations of θ , the given covariance matrix $\Sigma^{\theta\theta}$ is Cholesky decomposed to get the lower-triangular matrix. Applying this to uncorrelated random numbers produces a matrix with the covariance properties of the system being modeled.

³The Matlab code can be obtained from the authors on request.

DE restart a new population of DE candidate solutions $P_{j,t,i}$ is initialized and for the given set of 1000 θ draws DE minimizes the expected objective function value not in its mean but *median* to get a value being more robust to outliers. In short, the best DE solution is the one having the least expected median value over the 1000 Monte Carlo θ realizations. In other words, the algorithm looks for a set of controls which corresponds to the *least median objective value J (LMJ)* based on all possible realizations of θ (represented by 1000 Monte Carlo draws).

The choice of *median* instead of mean is dictated by its well known robustness to outliers, which in the view of the quadratic objective function and the complex nonlinear econometric models, where different variables have very different scaling and weights differently affecting the penalty function of the type (1) and producing those outliers, has a crucial role. Necessarily, the choice of the robust aggregated measure of expected value over the Monte Carlo draws of θ shall match the drawing distribution of the θ parameters. As long as the normal⁴, uniform or any other symmetric distribution is employed, there must be no problem with the approach. If, however, an asymmetric distribution is used, an alternative robust approach may be necessary.

Technically, the sequence of calculations proceeds as follows. For given initial controls (u) and every draw from the set of random θ s a series of initial states (x) is calculated. Afterwards, for the resulting population of solutions a median objective value *LMJ* is calculated using 1000 Monte Carlo θ realizations. Subsequently, candidates of the DE population are matched with each other creating new solutions, their new median objective value is calculated (using the same set of random θ s), the least fittest ones are discarded, while the most fittest (having lowest *LMJ*) create further new candidates. This continues until a stopping criterion is fulfilled.

3.2.3 DE calibration

Tuning of the DE parameters is always a problem specific issue. For this reason, similar to Blueschke *et al.* (2013a), we conduct a series of simulation experiments calibrating the DE parameters. In brief, we run DE for different *CR* and *F* ranging between 0.1 and 1 and construct a *phase portrait* that pictures combinations of parameter values with the lowest average number of generations required to achieve the value-to-reach (*VTR*).⁵ For the MacRae model (MacRae (1975)), the combination is highlighted if the minimum objective value obtained (J_{DE}) becomes less than or equal to *VTR* in less than

⁴Note that in the case of normal distribution mean and median coincide.

⁵*VTR* is set to either 101% or 105% of the objective value achieved by OPTCON (J_{OPTCON}). Thus, the deviations of 1% and 5% are suitable for illustrative reasons here.

or equal to 100 generations (left and middle panels in Figure 1). As a result of this exercise, we set $F = 0.8$ and $CR = 0.6$.

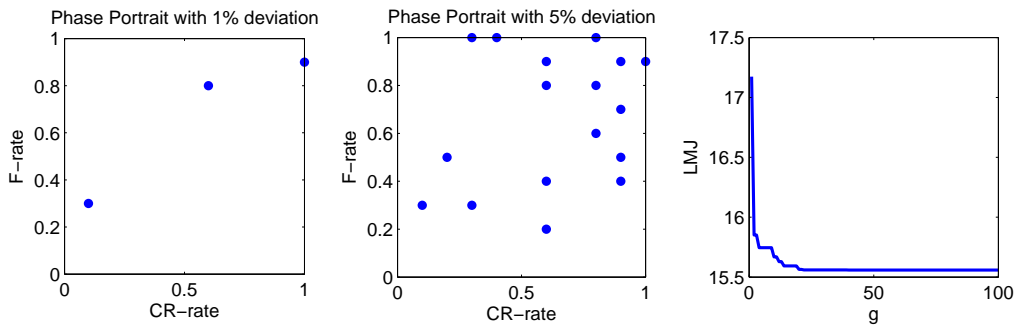


Figure 1: Phase portrait and progress plot for the MacRae model

Afterwards, to illustrate convergence of the resulting objective function values, we apply DE with 100 restarts for different g (number of generations): in the upper left plot of Figure 2 the cumulative distribution function $F(LMJ)$ for different g is given, whereas the other plots are kernel density estimations of objective function values identified.

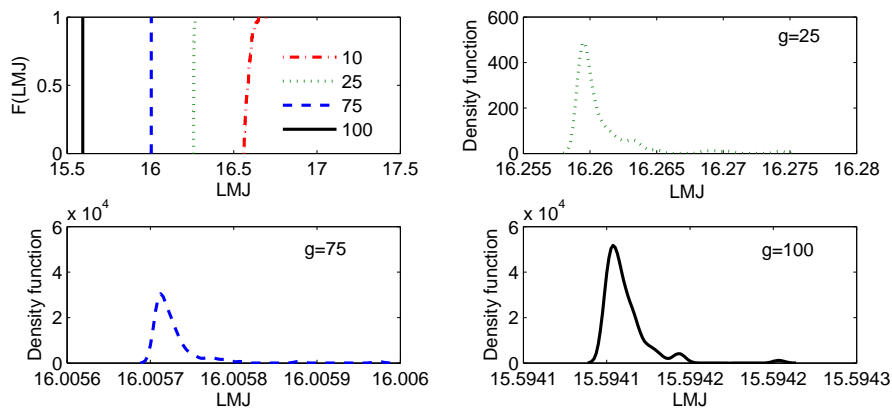


Figure 2: Empirical distribution of objective function values for different g

However, as the size of the model increases, the computational burden of the calibration exercise in the stochastic scenario described above increases tremendously. Since in the deterministic scenario considered in Blueschke *et al.* (2013a) $F = 0.4$ and $CR = 0.1$ were identified as the best parameter

combination for the SLOVNL model, we compared it with the one recommended as a default for many standard applications, namely $F = 0.8$ and $CR = 0.8$ (Storn and Price 1997) (keeping $p = 10d$ in both cases). We find that the latter leads DE population to a fast convergence into few local minima preventing further efficient exploration of the search space. In contrast, the combination applied in the deterministic approach maintains the diversity longer (mainly due to smaller crossover rate) providing a richer set of possibilities for exploration of better solutions.⁶ It is worth to mention that we also checked the DE calibration for the deterministic scenario of the MacRae model and found that the optimal calibration parameters are very close to the ones identified in the stochastic approach, namely $F = 0.7$ and $CR = 0.5$.⁷ This gives some evidence that calibration of the DE parameters conducted on the deterministic scenarios can also well fit for stochastic problems saving a lot of time.

As for the DE stopping criterion to be applied, this has to fulfill two conditions: i) DE algorithm shall ensure that population of solutions converges to an optimum (local or global); ii) DE has to stop once the convergence is observed. To combine those two conditions is particularly difficult under different θ s applied. To solve this problem, we set an upper limit on the number of DE generations to be performed within one restart (g^{max} equal to 100 and 2500 for MacRae and SLOVNL, respectively), but at the same time control for convergence within the population by looking on the candidates' objective values. Working on a continuous optimization problem, it is very unlikely that two candidates reach an exactly the same value, but a difference of 0.0001% between the fittest solution and few closest followers is realistic and, therefore, is applied in our study. Thus, DE algorithm stops if 30% of solutions in the population reach this deviation from the best solution available. In addition, if for 100 periods more than 50% of solutions in the population do not improve, the algorithm also stops. Given the stochastic nature of DE, we restart the algorithm 20 times to ensure that it converges if not to global optimum - to a good approximation of it.

Both methods, OPTCON and DE, are implemented in Matlab to simplify their comparison. The corresponding computational time for the two models tested in this study necessarily varies depending on the complexity of a particular problem. For SLOVNL around 750 generations are sufficient to obtain a solution, and require about 2-3 days (!) - for conducting both,

⁶For a detailed discussion of exploration of new solutions versus exploitation of existing ones see March (1991).

⁷In the deterministic approach DE well approximates the OPTCON results achieving the objective function value of 15.9577173 versus 15.9577151 provided by OPTCON. More details can be obtained on request.

extreme event analysis and least median optimization - using Matlab 7.11 and Pentium IV 3.3 GHz.

4 Simulation results

The present section is divided in three parts. First, we run the least median objective value analysis. To be more precise, we compare simulation results in terms of objective function values minimized: LMJ via DE and ex-post J via OPTCON and via DE. In addition to the convergence analysis of the DE results, we demonstrate that i) LMJ and ex-post J functions lead to different final solutions, while only the former one (LMJ) – explicitly tackling the stochastic problem – actually approximates a global solution for the optimal control problem; ii) DE clearly surpasses OPTCON in minimizing LMJ for two models considered: MacRae and SLOVNL. Second, in addition to the more accurate objective function being minimized, our DE algorithm provides results on extreme event analysis separately solving the optimal problem for a set of possible parameter realizations and comparing those between each other (e.g., plotting different percentiles) and to the results obtained from minimizing LMJ via DE and the corresponding ex-post J , thus, increasing amount of information for a policy maker to select an optimal solution. Third, given the simplicity of the models considered, we briefly illustrate differences between solutions in terms of states and controls obtained by DE and OPTCON.

4.1 LMJ analysis

As mentioned in Section 3.1, the classical algorithms (based on LQG framework) minimize the expected value of the objective function (expected cost-to-go), but only for the local linear approximation and, thus, there is no way to calculate the 'correct' expected objective function value of the original, nonlinear problem. As output most classical algorithms provide the objective value evaluated in a deterministic way along the best path of control and state variables. We denote this function value as *ex-post* J and provide it in addition to the least median objective value of J (LMJ), which is more appropriate to reflect the stochastic nature of the problem.

In order to calculate the corresponding ex-post J for DE, we choose the best set of controls (i.e. the one with lowest LMJ) in each generation and calculate for this set of controls u the respective states of the system using the initially given $\hat{\theta}$. While in general we observe a similar progress of the two objective values which drop substantially compared to the initial population,

for both models tested there are some fluctuations of the ex-post J (right plots in Figures 3 and 4). This points out the possibility of finding a better value and then loosing it (since it is not the objective criterion of DE and the two functions are *not associated*). At this point it is useful to specify what we mean by this expression and provide the following definition:

Definition 1 *The two functions $F_1(u_t)$ and $F_2(u_t)$ are associated iff for any two different solutions u'_t and u''_t ($\forall u'_t, u''_t$ that $u'_t \neq u''_t$) holds that*

$$F_2(u'_t) \geq F_2(u''_t) \quad \Leftrightarrow \quad F_1(u'_t) \geq F_1(u''_t).$$

In other words, if one function is monotonously decreasing on a certain sequence of solutions u_t , the other function shall also monotonously decrease on this set of solutions.

Figures 3 and 4 illustrate results of the best run (out of 20 restarts) of the DE algorithm applied to the MacRae and SLOVNL models using 1000 different sets of θ for 100 and 2500 generations (at maximum), respectively. While left part of the figures demonstrates the progress of LMJ (which is to minimize), its right part depicts the corresponding ex-post J values.

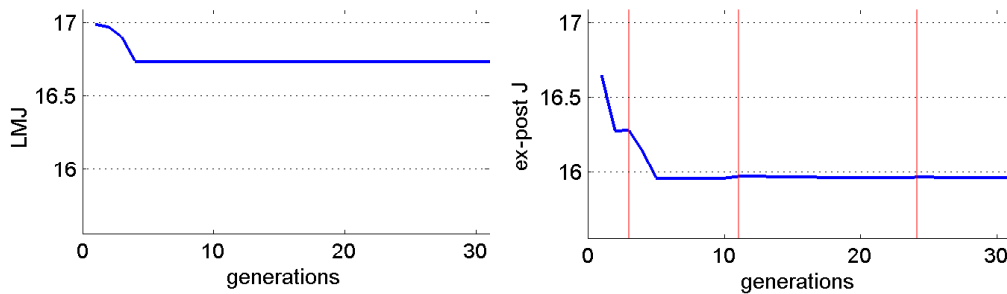


Figure 3: LMJ and ex-post J evolution over g^{max} for MacRae

Note: Transparent vertical lines on the right plot of the figure mark those generations, in which ex-post J rises towards its previous value thus illustrating the non-monotonic progress of the function.

For the MacRae model we see that the DE result in terms of ex-post J surpasses the OPTCON stochastic solution (≈ 16.6471) almost reaching the deterministic solution for that model (15.9577): deviation is of the order 10^{-3} (Table 1). The LMJ value obtained by DE is respectively lower than the one calculated for the OPTCON solution (using the same set of randomly chosen θ s as for DE). For the SLOVNL model, in case of LMJ the decrease is from 1.26×10^7 to 9.17×10^6 and in case of the ex-post J it is from 2.76×10^6 to 1.23×10^6 . Thus, for both models considered instead of converging to

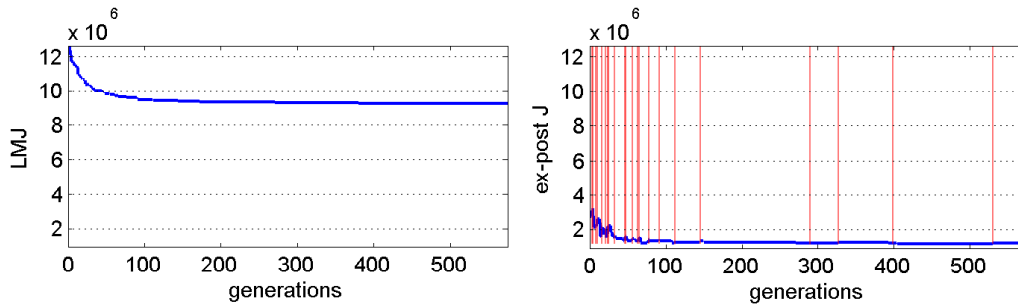


Figure 4: LMJ and ex-post J evolution over g^{max} for SLOVNL

Note: Transparent vertical lines on the right plot of the figure mark those generations, in which ex-post J rises towards its previous value thus illustrating the non-monotonic progress of the function.

the results obtained by OPTCON, DE arrives at different optima having a superior LMJ value, while ex-post J may be better or worse.⁸

Table 1: Results for both optimization algorithms with different settings

		OPTCON			Differential Evolution		
		deterministic	ex-post J	LMJ	deterministic	ex-post J	LMJ
MacRae	min	15.9577	16.6471	17.2721	15.9577	15.9638	16.7295
	std	–	–	–	(.0000)	(.0003)	(.0000)
	cpu	.01s	.02s	5.2s	23s	4169s	
SLOVNL	min	904,649.68	918,574.53	9,505,241.36	904,649.98	1,226,557.91	9,175,091.31
	std	–	–	–	(0.62)	(11,335.89)	(41,017.66)
	cpu	.5s	6s	4.9s	58s	217930s	

Note: Results on deterministic DE solution are obtained by the algorithm explained in detail in Blueschke *et al.* (2013a). Ex-post J for OPTCON stands for the standard stochastic result reported (for consistent comparison). LMJ for OPTCON is calculated for the resulting OPTCON solution using the same set of randomly chosen θ s as for DE. Results on minimum and standard deviation are stated for 20 restarts, while cpu time corresponds to one restart only.

The clear difference between LMJ and ex-post J obtained via DE for both models is due to the different θ parameters involved in computation: while in the first case those are the Monte Carlo θ draws used in each DE generation to compute the median of all possible objective value realizations,

⁸In particular, comparing results obtained by DE through 20 different restarts one could see that solutions with lower LMJ have higher ex-post values than those with slightly higher least median values. Looking closer on the series depicted in Figures 3 and 4 one could see that, e.g., ex-post J value rises to a value higher than the preceding one (to 15.9691 from 15.9583 on 11th generation for MacRae and to 2.91×10^6 for SLOVNL just after beginning) but than also loses a value lower than the one corresponding to the final solution (15.9583 for MacRae and 1.21×10^6 for SLOVNL).

in the latter case a single set of uncertain parameters (as if they were deterministic) is used. Recalling equation (3), we calculate the optimal states in case of LMJ function from the dynamic system f , parameter set θ_i and the optimal controls, i.e. $(x_i^* = f(\dots, u^*, \theta_i, \dots))$ ($i = 1, \dots, 1000$) and use these possible stochastic solutions (u^*, x_i^*) to calculate the LMJ value. Usually, the system function f is very sensitive to the changes in θ .⁹ As a result of many 'unexpected' realizations of θ many states can exhibit a higher deviation from the ideal path and, respectively, higher objective values J , which leads to a high LMJ value. For calculating ex-post J the expected values of parameters θ are used, namely $\hat{\theta}$, which are estimated and available to policy maker, and which clearly reduce a possibility of outliers and other more 'unexpected' results.

The difference between LMJ and ex-post J does not always have to be in favour of one of the approaches and may vary depending on the specific set of θ s employed. It is more important to highlight the fact that the ex-post J is not necessarily monotonically decreasing while DE minimizes LMJ . Hence, the solutions identified through explicit minimization of stochastic θ – with minimum LMJ – and the ones reported by traditional algorithms do not necessarily coincide. This finding can be considered as a direct consequence of LMJ and ex-post J not being associated according to Definition 1.

Since in the traditional optimal control algorithms the ex-post J is the only information available to the decision-maker (except the optimal set of controls), the present DE algorithm with no doubt broadens the range of information available and can be a helpful complementary tool for anyone facing an optimal control problem. And what is more, the fact that DE surpasses OPTCON based on the actual least median function evaluated for a random set of 1000 θ for both models illustrates that OPTCON only reaches suboptimal solutions, which can be further improved by the more flexible (heuristic) approach.

To test the robustness of the result (ex-post J and LMJ not being associated) we run a sensitivity analysis with respect to the variance of the parameters θ applied to SLOVNL model. To this end, we apply a similar procedure as it was performed by Blueschke *et al.* (2013b) and multiply the covariance matrix $\Sigma^{\theta\theta}$ (defined in Section 2) by a parameter ρ :

$$\Sigma_{SA}^{\theta\theta} = \rho \Sigma^{\theta\theta}. \quad (4)$$

Afterwards, Monte Carlo realizations of θ_{SA} drawn from the resulting covari-

⁹For the MacRae model this is not the case, because the model has just one state variable, one control variable and one parameter. Thus, ex-post J may be larger than LMJ in absolute terms for MacRae

ance matrix $\Sigma_{SA}^{\theta\theta}$ are used by DE to calculate LMJ and the corresponding ex-post J . The sensitivity analysis is conducted for three different parameter values $\rho=0.5, 1$ and 2 with 20 restarts for each value. Given the type of problem specified (equations (1-3)), one would expect the resulting LMJ solution to rise in ρ and variance between the results (different DE restarts) obtained to increase. The latter shall illustrate the transmission of a larger parameter variance into a larger variance in results, while the former is explained by the following relations: higher variance in parameters θ leads to higher variance of resulting state variables (because of equation (3), or rather due to high sensitivity of system function f to changes in θ), which leads to higher objective values and higher LMJ value. Regarding the ex-post J function, there is no clear theoretical foundation on which impact of higher variance is to be expected.

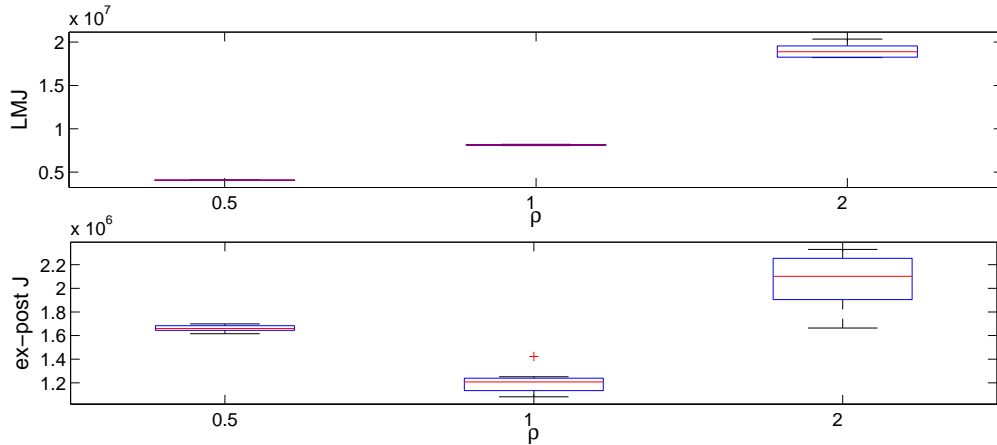


Figure 5: Sensitivity results for different ρ applied to SLOVNL

The results of this sensitivity analysis are presented in Figure 5 and support our expectations. One can see that variance in results rises in ρ for both objective functions (with increasing ρ the boxplots are larger). LMJ values also increase with higher values of ρ . A very interesting detail is that LMJ values increase not only monotonously but very much linearly in ρ . For ex-post J a non-linear (U-shape dynamics) is observed. The latter result is fairly unexpected (and requires further research) but indicates again that the LMJ and ex-post J functions are not associated according to Definition 1.

4.2 Extreme event analysis

Figures 6 and 7 are meant to illustrate the variety of the θ draws (*extreme event analysis*). The lower panel contains boxplots of the 1000 Monte Carlo draws for each model parameter rescaled between 0 and 1.¹⁰ The more equally the draws are dispersed between zero and one, the more representative they are. The 'worst' and 'best' draws illustrate specific θ realizations with the largest and the lowest objective value J obtained via DE. Thus, these draws can be attributed to the 'best' and 'worst' scenarios possible. One can see that though those draws are distinct from each other, they do not necessarily locate on the opposite sides of the boxplots.

In the upper panel of the Figures 6 and 7 one can see how the objective function values identified by DE relate to each other. In particular, how the LMJ and the corresponding ex-post J values are related to the objective function values obtained for single parameter draws: to certain percentiles of the entire distribution of 1000 θ s (upper left plot) and percentiles of J outcomes (upper right plot). Thus, for example, one can evaluate how wide is the range of outcomes between different parameter realizations (percentiles in θ) and between 'best' (5th percentile in J) and 'worst' (95th percentile in J) possible realizations and how the objective value (LMJ and ex-post J) of the optimal result obtained via DE relates to them.

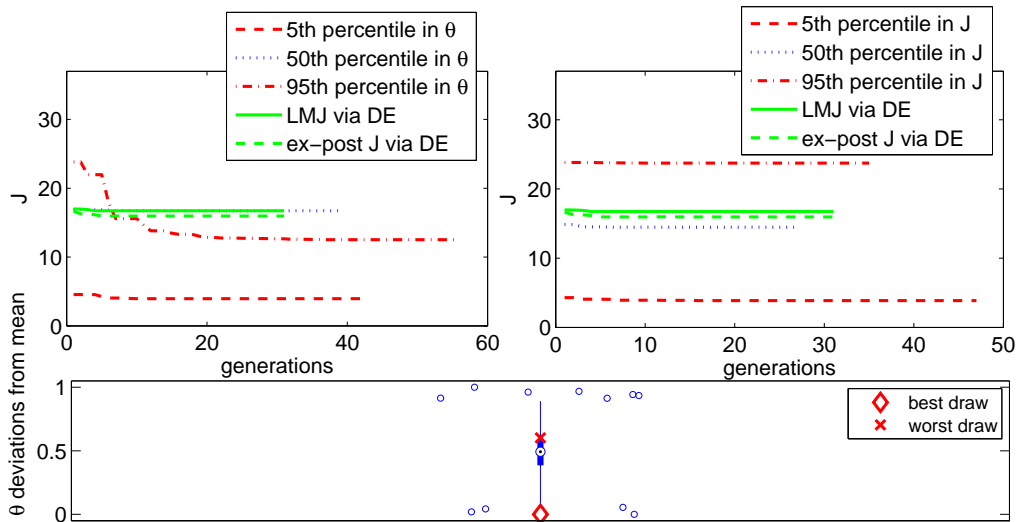


Figure 6: Different θ s and corresponding objective values for MacRae

¹⁰For this, we deduct from each θ draw the minimum of the draws for the respective parameter and divide it by the difference between the maximum and minimum: $\frac{\theta_i - \min(\theta)}{\max(\theta) - \min(\theta)}$.

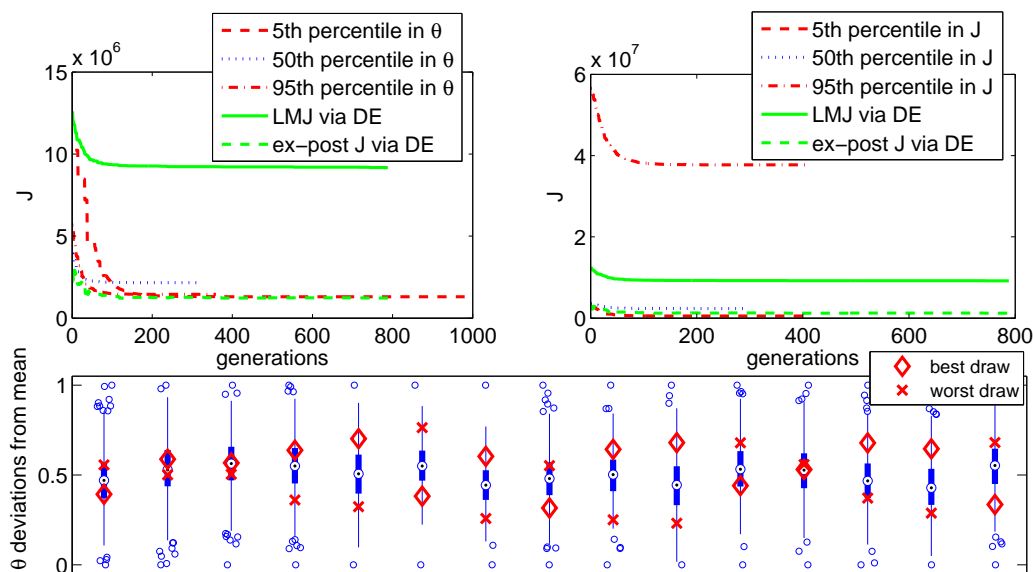


Figure 7: Different θ s and corresponding objective values for SLOVNL

4.3 Comparison of states and controls

On Figures 8 (left plot) and 9 we plot percentage deviations of the controls obtained by DE corresponding to single Monte Carlo θ draws over the respective planning horizons. These figures also illustrate the difference between the OPTCON solution (marked with a cross) and least median solution (*LMJ*) obtained by DE (marked with a box) and their relative position to the controls obtained in the 'best' and 'worst' scenarios. For the MacRae model, e.g., one can see that DE and OPTCON solutions lie around the distributions mean, while OPTCON is closer to the θ draw of the 'best draw' scenario. For SLOVNL this generally also holds being most evident for the net tax rate (TaxRate). The 'worst draw' scenario (i.e. best solution in worst case of θ realizations) for the control variable $M3N$ (nominal money stock) can be interpreted, for example, as a recommendation to run a more active monetary policy in most of the time periods as a best response to the pessimistic scenario.

On Figures 8 (right plot) and 10 we plot percentage deviations of the obtained states over the respective planning horizons. Similarly to the controls, these figures show the relative position of the DE and OPTCON solutions to the mean, worst and best scenarios for different state variables. For MacRae, in contrast to the controls, DE solution is closer to the θ draw of the 'best draw' scenario. For SLOVNL the resulting states seem to be very close ex-

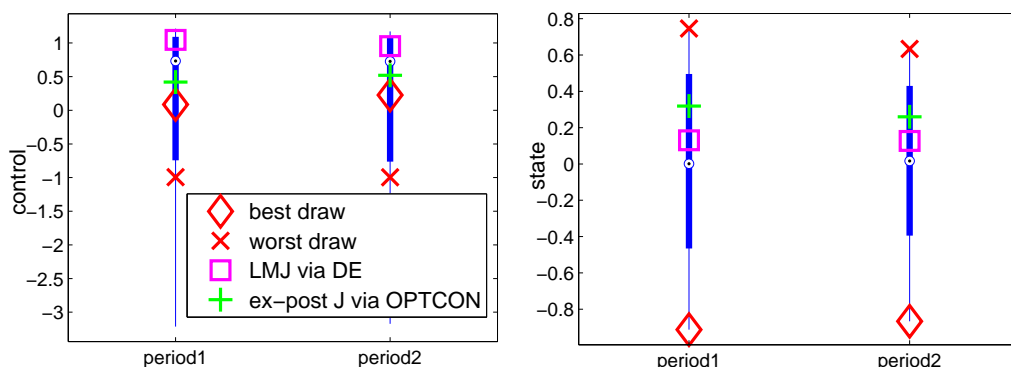


Figure 8: Percentage deviation in controls and states for MacRae.

cept the short term interest rate (STIRLN), where DE starting from the forth period systematically prefers some larger values than OPTCON.

Interestingly, one can see that for SLOVNL basically for almost all states (except of inflation) the volatility of the results obtained in the extreme event analysis rises consistently over time (in every next period relative deviation of solutions corresponding to single θ draws from the mean over 1000 draws increases). This may be explained by the stochastic nature of the problem itself and its increasing uncertainty for every next period involved.

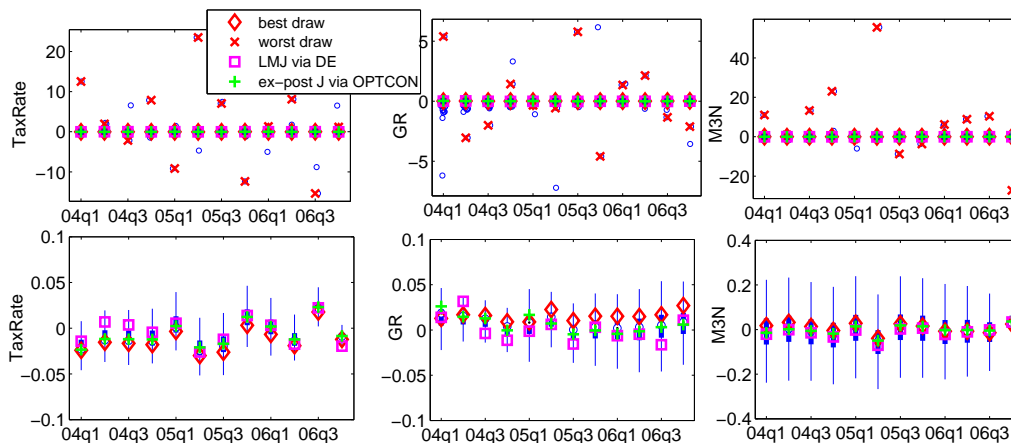


Figure 9: Percentage deviation in controls for SLOVNL

Note: While the upper panel illustrates the positioning of the worst draws relative to other results, the lower one provides boxplots on the majority of the results depicting how their variation changes over time.

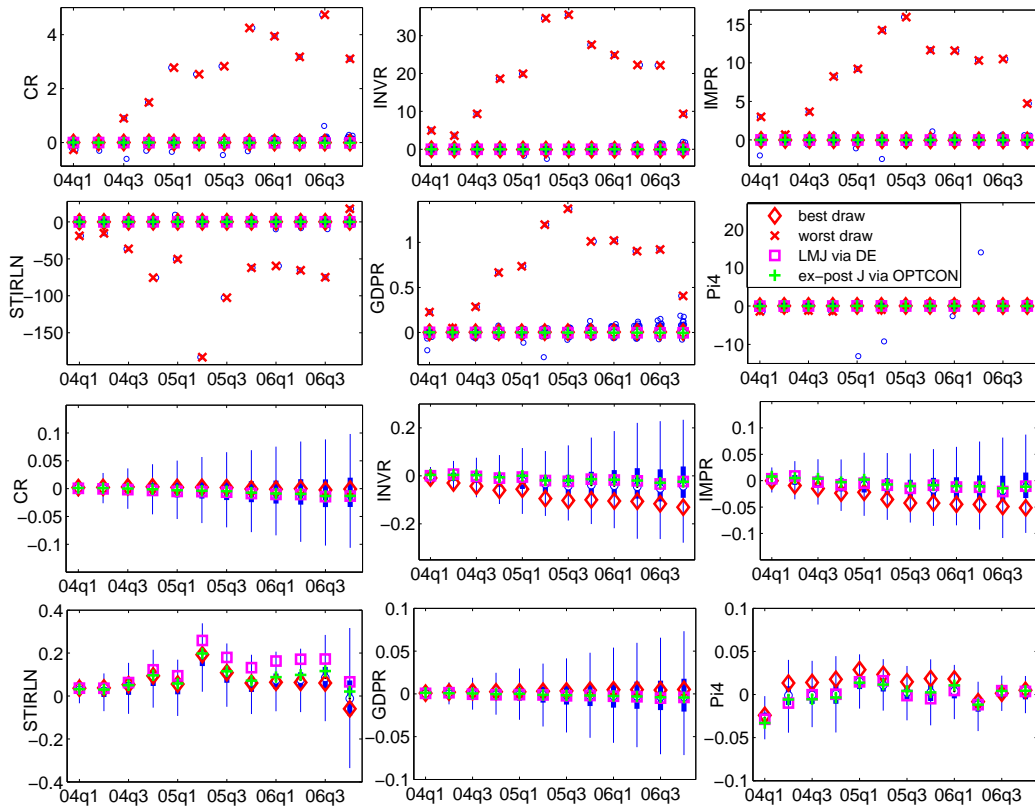


Figure 10: Percentage deviation in states for SLOVNL

Note: While the upper two panels illustrate the positioning of the worst draws relative to other results, the lower ones provide boxplots on the majority of the results depicting how their variation changes over time.

5 Conclusions and Outlook

In this paper we apply a heuristic approach (Differential Evolution) to solve stochastic nonlinear optimal control problems. The main reason to do that is the DE's flexibility allowing to deliver solutions in the specific situations, where the classical methods exhibit some limitations. In particular, we apply extreme event analysis and minimize the median objective value obtained over a set of randomly drawn (stochastic) parameter realizations aiming by this to broaden the range of decision support information used by policy makers in choosing the optimal strategy. To test the quality and performance of DE we compare its results with the ones obtained by the OPTCON algorithm, which uses the classical techniques of linear-quadratic optimization, on the MacRae and SLOVNL models.

First, we demonstrate that DE indeed allows to obtain a broader infor-

mation set about the stochastic solution. In addition to the ex-post J , which is the only information available in traditional algorithms, DE calculates the expected objective value and the best responses for different draws of parameters. Second, we find that the best solution (the optimal set of controls) in terms of ex-post J does not necessarily have the minimum expected objective value, and hence, may represent a suboptimal solution. This shall be considered as a specific finding of a more general one, namely that the two objective functions are not associated, which questions the correctness of the objective function specification by classical approaches based on local-linear approximation. Finally, it is shown that DE surpasses classical methods in finding better solutions in terms of expected objective value (LMJ) for both models tested, which supports the claim that heuristic methods are better suited in identifying good approximations of global optima in complex search spaces.

Clearly, these results require further analysis both, from the side of additional models to be tested and (potentially) alternative approaches to sample stochastic parameters to be examined. Furthermore, usefulness of Differential Evolution in handling further limitations of the classical methods (as, e.g., symmetry of the objective function demonstrated in Blueschke *et al.* (2013a)) shall be addressed in future research. Additionally, particular attention shall be given to the computational costs of applying the suggested DE algorithm to stochastic problems. As the cpu time corresponding to the small and very small models like SLOVNL and MacRae already varies between few hours and days, it may become prohibitively large for more complex models. Thus, more effort shall be devoted to shrink the computational cost by i) outsourcing certain parts of the code into a lower-level language, like C++, which is more efficient for the large loops (e.g., 1000 θ s) and ii) by parallelizing the computational burden of the extreme event analysis (where individual parameter sets are independently analyzed) between several computers.

References

- Aoki, M. (1989). *Optimization of Stochastic Systems. Topics in Discrete-Time Dynamics*. 2 ed.. Academic Press. New York.
- Benigno, P. and M. Woodford (2012). Linear-quadratic approximation of optimal policy problems. *Journal of Economic Theory* **147**, 1 – 42.
- Blueschke, D., V. Blueschke-Nikolaeva and I. Savin (2013a). New insights into optimal control of nonlinear dynamic econometric models: Application

- of a heuristic approach. *Journal of Economic Dynamics and Control* **37**(4), 821 – 837.
- Blueschke, D., V. Blueschke-Nikolaeva and R. Neck (2013b). Stochastic control of linear and nonlinear econometric models: Some computational aspects. *Computational Economics* **42**(1), 107 – 118.
- Blueschke-Nikolaeva, V., D. Blueschke and R. Neck (2012). Optimal control of nonlinear dynamic econometric models: An algorithm and an application. *Computational Statistics and Data Analysis* **56**(11), 3230–3240.
- Chow, G. C. (1975). *Analysis and Control of Dynamic Economic Systems*. John Wiley & Sons. New York.
- Chow, G. C. (1981). *Econometric Analysis by Control Methods*. John Wiley & Sons. New York.
- Cosimano, T.F. (2008). Optimal experimentation and the perturbation method in the neighborhood of the augmented linear regulation problem. *Journal of Economic Dynamics and Control* **32**(6), 1857–1894.
- Egbetokun, A. and I. Savin (forthcoming). Absorptive capacity and innovation: When is it better to cooperate?. *Journal of Evolutionary Economics*.
- Gilli, M. and P. Winker (2009). Heuristic optimization methods in econometrics. In: *Handbook of Computational Econometrics* (D.A. Belsley and E. Kontoghiorghes, Eds.). pp. 81–119. Wiley. Chichester.
- Kendrick, D. A. (1981). *Stochastic Control for Economic Models*. McGraw-Hill, New York, second edition, [http : //www.utexas.edu/cola/_db/app/file.php?id = 495395](http://www.utexas.edu/cola/_db/app/file.php?id=495395).
- Kendrick, D. A. (2005). Stochastic control for economic models: past, present and the paths ahead. *Journal of Economic Dynamics and Control* **29**, 3–30.
- Lyra, M., J. Paha, S. Paterlini and P. Winker (2010). Optimization heuristics for determining internal rating grading scales. *Computational Statistics & Data Analysis* **54**(11), 2693–2706.
- MacRae, E. C. (1975). An adaptive learning rule for multiperiod decision problems. *Econometrica* **43**(5-6), 893–906.

- March, J.G. (1991). Exploration and exploitation in organizational learning. *Organization Science* **2**(1), 71 – 87.
- Matulka, J. and R. Neck (1992). OPTCON: An algorithm for the optimal control of nonlinear stochastic models. *Annals of Operations Research* **37**, 375–401.
- Neck, R. (1984). Stochastic control theory and operational research. *European Journal of Operations Research* **17**, 283–301.
- Storn, R. and K. Price (1997). Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *J. Global Optimization* **11**, 341–359.
- Tucci, M. P., D. A. Kendrick and H. M. Amman (2010). The parameter set in an adaptive control Monte Carlo experiment: Some considerations. *Journal of Economic Dynamics and Control* **34**(9), 1531 – 1549.
- Winker, P., M. Lyra and C. Sharpe (2011). Least median of squares estimation by optimization heuristics with an application to the CAPM and a multi-factor model. *Computational Management Science* **8**(1-2), 103–123.

6 Appendix

In the following we briefly describe the models used in our study. The MacRae model is a very simple one-state, one-control, one unknown parameter linear-quadratic tracking model (MacRae (1975)). The SLOVNL model is based on the Slovenian economy using quarterly data from databases and publications by the Slovenian Statistical Office, the Bank of Slovenia and Eurostat.

The SLOVNL model

The small nonlinear macroeconomic model of the Slovenian economy (SLOVNL) consists of 8 equations (4 behavioral and 4 identities). The model includes 8 state variables, 4 exogenous non-controlled variables, 3 control variables and 15 unknown (estimated) parameters. The quarterly data for the period 1995:1 to 2006:4 yield 48 observations and admit a full-information maximum likelihood (FIML) estimation of the expected values and the covariance matrices for the parameters and system errors. The start period for the optimization is 2004:1 and the end period is 2006:4 (12 periods).

Endogenous (state) variables:

$x[1]$	<i>CR</i>	real private consumption
$x[2]$	<i>INVR</i>	real investment
$x[3]$	<i>IMPR</i>	real imports of goods and services
$x[4]$	<i>STIRLN</i>	short term interest rate
$x[5]$	<i>GDPR</i>	real gross domestic product
$x[6]$	<i>VR</i>	real total aggregate demand
$x[7]$	<i>PV</i>	general price level
$x[8]$	<i>Pi4</i>	rate of inflation

Control variables:

$u[1]$	<i>TaxRate</i>	net tax rate
$u[2]$	<i>GR</i>	real public consumption
$u[3]$	<i>M3N</i>	money stock, nominal

Table 2: Weights of the variables in the SLOVNL model

a: 'raw' weights		b: 'correct' weights	
variable	weight	variable	weight
<i>CR</i>	1	<i>CR</i>	3.457677
<i>INVR</i>	1	<i>INVR</i>	12.16323
<i>IMPR</i>	1	<i>IMPR</i>	1.869532
<i>STIRLN</i>	1	<i>STIRLN</i>	216403.9
<i>GDPR</i>	2	<i>GDPR</i>	2
<i>VR</i>	1	<i>VR</i>	0.333598
<i>PV</i>	1	<i>PV</i>	423.9907
<i>Pi4</i>	0	<i>Pi4</i>	0
<i>TaxRate</i>	2	<i>TaxRate</i>	37770.76
<i>GR</i>	2	<i>GR</i>	63.77052
<i>M3N</i>	2	<i>M3N</i>	0.090549