ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

SOLVING, ESTIMATING, AND SELECTING NONLINEAR DYNAMIC MODELS WITHOUT THE CURSE OF DIMENSIONALITY

Author(s): Viktor Winschel and Markus Krätzig

Source: *Econometrica*, Vol. 78, No. 2 (March, 2010), pp. 803-821
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/40664494
Accessed: 25-06-2018 17:01 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://about.jstor.org/terms

The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

Econometrica, Vol. 78, No. 2 (March, 2010), 803–821

SOLVING, ESTIMATING, AND SELECTING NONLINEAR DYNAMIC MODELS WITHOUT THE CURSE OF DIMENSIONALITY

BY VIKTOR WINSCHEL AND MARKUS KRÄTZIG[1]

We present a comprehensive framework for Bayesian estimation of structural nonlinear dynamic economic models on sparse grids to overcome the curse of dimensionality for approximations. We apply sparse grids to a global polynomial approximation of the model solution, to the quadrature of integrals arising as rational expectations, and to three new nonlinear state space filters which speed up the sequential importance resampling particle filter. The posterior of the structural parameters is estimated by a new Metropolis–Hastings algorithm with mixing parallel sequences. The parallel extension improves the global maximization property of the algorithm, simplifies the parameterization for an appropriate acceptance ratio, and allows a simple implementation of the estimation on parallel computers. Finally, we provide all algorithms in the open source software JBendge for the solution and estimation of a general class of models.

KEYWORDS: Dynamic stochastic general equilibrium (DSGE) models, Bayesian structural time series econometrics, curse of dimensionality.

1. INTRODUCTION

MANY MODERN MACROECONOMIC MODELS with rational expectations are nonlinear and linearization combined with a Kalman (1960) filter has several disadvantages. Fernández-Villaverde and Rubio-Ramírez (2006) reported evidence for nonlinearities in macroeconomic data; consequently, linear estimates are likely to be biased. Theoretical nonlinearities naturally arise in the aggregate, for example, in a model with nominal rigidities and a lower bound on the nominal interest rate; see Billi and Adam (2007).

Unfortunately, a nonlinear likelihood based econometric approach is a complex numerical operation with at least four problems. We propose innovations to all of them. The first problem is to solve the model. Here we use the Smolyak operator for a polynomial approximation of the policy functions as well as for the quadrature of rational expectations. The second problem is to evaluate the likelihood where we introduce three nonlinear state space filters. The third problem is to generate parameter estimates. We do this by a new parallel Metropolis–Hastings algorithm. The fourth problem is to program all these

[1]We thank Wouter Denhaan, Paul Fackler, Jesús Fernández-Villaverde, James Heckman, Florian Heiss, Kenneth Judd, Michel Juillard, Felix Kübler, Alexander Ludwig, Thomas Mertens, Juan Rubio-Ramírez, and the participants of the Institute on Computational Economics at the University of Chicago and Argonne National Laboratory, 2005, for discussion and inspiration. We also thank the anonymous referees for their valuable comments. This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 Economic Risk.

© 2010 The Econometric Society DOI: 10.3982/ECTA6297

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

complex interacting algorithms. As a solution, we provide the open source software JBendge[2] for a general class of models.

The currently predominant approximation strategy is a local perturbation approach based on the implicit function theorem discussed by Judd and Guu (1997), Judd and Jin (2002), and Gaspar and Judd (2005).[3] A global approximation based on the projection method is presented by Judd (1992). Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) compared these approaches.

The global approximation of functions or integrands suffers from exponentially growing computational costs, called the curse of dimensionality, in case it is extended from the univariate to the multivariate case by the tensor operator. The curse effectively restricts tensor based global approximation methods to be useful only for a few (in our case, six) dimensions. We replace the tensor operator by the Smolyak (1963) operator, applicable for polynomial, finite element, or wavelet approximations. It decreases the approximation burden for sufficiently smooth functions and integrands from exponentially growing costs to a polynomial order; see Bungartz and Griebel (2004) for an extensive summary.

The current approach to the evaluation of the likelihood of state space models is dominated by the sequential importance resampling particle filter and its extensions described in Doucet, de Freitas, and Gordon (2001). We introduce three new nonlinear filters based on deterministic integration to overcome the high computational burden of the particle filter.

The first filter is the simple but very fast deterministic Smolyak Kalman filter, which is an improvement of the unscented Kalman filter by Julier and Uhlmann (1997). The unscented integration restricts the computational costs to rise linearly with the dimension of the integrand and, accordingly, suffers from a decreasing approximation quality. Moreover, the unscented method is tailored to integrands weighted by Gaussian densities and cannot be easily adapted to other cases, whereas the Smolyak quadrature is readily applicable to any existing univariate scheme.

The Smolyak as well as the unscented Kalman filters approximate some non-Gaussian densities with Gaussians during the filtering recursions. As an improvement, we propose the Smolyak particle filter, which uses the density obtained by the Smolyak Kalman filter as the proposal density for the sequential importance resampling particle filter. This is an approach similar to the unscented particle filter of van der Merwe, Doucet, de Freitas, and Wan (2000). Our third new filter deterministically approximates the non-Gaussian densities by Gaussian sums.

[2]JBendge is an acronym for Java based Bayesian estimation of nonlinear dynamic general equilibrium models. The homepage of this project is at http://jbendge.sourceforge.net/. It is licensed under the GPL.

[3]Current software packages that implement these methods are, for example, Juillard (1996) or Schmitt-Grohé and Uribe (2004).

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

For the parameter estimation, we combine the Metropolis–Hastings algorithm with the genetic algorithm of Storn and Price (1997) to receive a parallel version. This improves the global maximization properties, automatically chooses the innovation variances on the run, and simplifies the implementation on parallel computers.

One application of the Smolyak operator is presented by Heiss and Winschel (2008), where a Smolyak based mixed logit likelihood estimator systematically outperforms simulation based methods. The research on likelihood based nonlinear dynamic structural estimation has just started with a number of papers by Fernández-Villaverde and Rubio-Ramírez (2004, 2005, 2006) and a recent estimation by Amisano and Tristani (2007).

The rest of the paper is divided into four sections: economics, econometrics, results, and conclusion.[4] In the economics Section 2, we present the generic model class, the optimality conditions for the example model, the Smolyak operator, the nonlinear solution strategy based on a Chebyshev approximation, and the approximation error estimator. In the econometrics Section 3, we describe the Smolyak Kalman filter, the particle filter, the Smolyak particle filter, and the Smolyak sum filter, the parallel Metropolis–Hastings algorithm, the convergence diagnostics, and the calculation of the marginal likelihood. In the results Section 4, we summarize the performance of the solution method, the filters, and the estimators. In Section 5, we conclude.

2. ECONOMICS

Our general model class

$$
\begin{aligned}
0 &= f(s_t, x_t, z_t; \theta), \quad \text{where} \quad f : \mathbb{R}^{d_s+d_x+d_z} \to \mathbb{R}^{d_x}, \\
z_t &= \mathrm{E}_e h(s_t, x_t, e_{t+1}, s_{t+1}, x_{t+1}; \theta), \quad \text{where} \quad h : \mathbb{R}^{d_s+d_x+d_e+d_s+d_x} \to \mathbb{R}^{d_z}, \\
s_{t+1} &= g(s_t, x_t, e_{t+1}; \theta), \quad \text{where} \quad g : \mathbb{R}^{d_s+d_x+d_e} \to \mathbb{R}^{d_s},
\end{aligned}
$$

is formulated in terms of dynamic first-order optimality conditions f, expected functions h, and state transitions g. The variables are states $s_t \in \mathbb{R}^{d_s}$, policies $x_t \in \mathbb{R}^{d_x}$, expected variables $z_t \in \mathbb{R}^{d_z}$, and random shocks $e_{t+1} \in \mathbb{R}^{d_e}$, which are usually assumed to follow a normal distribution $e_{t+1} \sim \mathcal{N}(0, \Sigma_e)$ with a diagonal Σ_e. The vector θ contains the structural parameters.

[4]Additional material is available in the supplemental material to this paper (Winschel and Krätzig (2010)), which has details on the linearization of the general model, the new filters, the Metropolis–Hastings algorithm, an example of a two-dimensional Smolyak polynomial approximation, and detailed simulation results.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

2.1. *Model*

In our example, the allocation problem is solved by the dynamic optimization

$$\max_{\{\{c_{n,t},l_{n,t},i_{n,t}\}_{n=1}^{N}\}_{t=0}^{\infty}} U = \mathrm{E}_0 \sum_{t=0}^{\infty} \sum_{n=1}^{N} \beta^t U_{n,t}$$

for $n = 1, \ldots, N$ countries and all future periods $t \geq 0$. The welfare function U is a discounted sum of country utilities $U_{n,t} = (c_{n,t}^{\theta_n}(1 - l_{n,t})^{1-\theta_n})^{1-\tau_n}/(1 - \tau_n)$ with a common discount factor β for all countries, an elasticity of intertemporal substitution τ_n, and consumption and leisure substitution rate θ_n, for each country. The world budget constraint $\sum_{n=1}^{N}(y_{n,t} - c_{n,t} - i_{n,t}) = 0$ restricts the world output to be either consumed or invested. The policy variables are consumption $c_{n,t}$, labor $l_{n,t}$, and investment $i_{n,t}$ for each country. The production technology $y_{n,t} = e^{a_{n,t}} k_{n,t}^{\alpha_n} l_{n,t}^{1-\alpha_n}$ depends on productivity $a_{n,t}$, capital $k_{n,t}$, labor $l_{n,t}$, and the technical substitution rate α_n. The capital and productivity transitions are

$$k_{n,t+1} = i_{n,t} + (1 - \delta_n)k_{n,t} - 0.5\kappa_n i_{n,t}^2, \tag{1}$$

$$a_{n,t+1} = \rho_n a_{n,t} + e_{n,t+1}, \tag{2}$$

where δ_n are the depreciation rates and ρ_n are the autocorrelation coefficients for the productivity processes with normally distributed shocks $e_{n,t} \sim \mathcal{N}(0, \sigma_{e_n})$ which are independent across countries and time. In the capital transition equation (1), we include capital adjustment costs parameterized by κ_n. They assure that the state of the system is not simply the aggregate capital stock, but its distribution across the countries.

The variables and parameters of the example model correspond to the variables of the general model class as follows: $s_t = \{k_{n,t}, a_{n,t}\}_{n=1}^{N}$, $x_t = \{c_{n,t}, l_{n,t}, i_{n,t}\}_{n=1}^{N}$, $e_t = \{e_{n,t}\}_{n=1}^{N}$, $\theta = \{\tau_n, \theta_n, \alpha_n, \delta_n, \rho_n, \kappa_n, \sigma_{e_n}\}_{n=1}^{N} \cup \{\beta\}$.

2.2. *Optimality*

The Bellman equation for $V_t \equiv V(k_{1,t}, \ldots, k_{N,t}, a_{1,t}, \ldots, a_{N,t}; \theta)$ is

$$\begin{aligned} V_t = \max_{\{c_{n,t},l_{n,t},i_{n,t},k_{n,t+1}\}_{n=1}^{N}} & \sum_{n=1}^{N} U_{n,t} + \beta \mathrm{E}_t V_{t+1} + \lambda_B \sum_{n=1}^{N} (y_{n,t} - c_{n,t} - i_{n,t}) \\ & + \sum_{n=1}^{N} \lambda_n (k_{n,t+1} - (1 - \delta_n)k_{n,t} - i_{n,t} + 0.5\kappa_n i_{n,t}^2). \end{aligned}$$

As a characterization of the solution, we obtain N Euler equations

$$\frac{1}{1 - \kappa_n i_{n,t}} \frac{\partial U_{n,t}}{\partial c_{n,t}} - \beta \mathrm{E}_t \left(\frac{\partial U_{n,t+1}}{\partial c_{n,t+1}} \left(\frac{\partial y_{n,t+1}}{\partial k_{n,t+1}} + \frac{1 - \delta_n}{1 - \kappa_n i_{n,t+1}} \right) \right) = 0, \tag{3}$$

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

N intratemporal trade-offs between consumption and labor supply

$$\frac{\partial U_{n,t}}{\partial l_{n,t}} + \frac{\partial U_{n,t}}{\partial c_{n,t}} \frac{\partial y_{n,t}}{\partial l_{n,t}} = 0, \tag{4}$$

and $N-1$ cross-country optimality conditions for $n = 2, \ldots, N$,

$$\frac{\partial U_{1,t}}{\partial c_{1,t}} = \frac{\partial U_{n,t}}{\partial c_{n,t}}. \tag{5}$$

The $4N$ equations for the variables $c_{n,t}, l_{n,t}, i_{n,t}$, and $k_{n,t+1}$ are the N Euler conditions (3), N intratemporal trade-offs between consumption and labor (4), $N-1$ equalities of marginal utilities (5), and the budget constraint and N capital transitions (1).

The mapping into the general model class is the following: the $2N-1$ equations (3) and (5) describe the policy functions for $l_{1,t}, \ldots, l_{N,t}$ and $i_{2,t}, \ldots, i_{N,t}$ in the general first-order conditions f. The general model functions h for N forward looking variables are given by the arguments of the expected values in the Euler equations (3). The N capital transitions (1) and the N productivity transitions (2) form the state transition functions g of the general model.

2.3. *Solution*

Our solution approach is to iterate on the implicitly defined policy functions $x^*: \mathbb{R}^{d_s} \to \mathbb{R}^{d_x}$. The policy values $x^{(k)}$ at the grid of the states in each iteration k are obtained as the solutions to $f(s, x^{(k)}, z^{(k-1)}) = 0$ for given expected variables $z^{(k-1)} = \mathrm{E}_e h(x^{(k-1)})$ based on the policies $x^{(k-1)}$ from the previous iteration. We generate the start values for the iteration from a linear approximation.

The evaluation of $z^{(k)}$ involves a numerical integration which can be thought as a polynomial approximation of the integrand with a subsequent trivial integration. This amounts to evaluating $\mathrm{E}_e h(e) = \int h(e)p(e)\,de \approx \sum_j w_j h(e_j)$, where the continuous random variable e and its density $p(e)$ are essentially discretized into realizations e_j with mass w_j.

An alternative approach would solve for x in $f(s, x, \mathrm{E}_e h(x)) = 0$ in one large system of $d_g \times d_x$ equations, where d_g denotes the number of grid points of the approximation. Our method has the advantage that we can solve d_g separate systems with d_x equations during each iteration. Moreover, the Jacobians for these small systems are available.[5]

A restriction of the function iteration scheme is that all policy variables have to appear in the contemporary policy vector x in $f(s, x, z) = 0$.

[5] JBendge uses a symbolic engine for an automatic calculation of these Jacobians.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

2.3.1. *Approximation*

The approximation of the policy function $x^*(s;c) = \sum_j c_j b_j(s)$ is characterized by the coefficient vector $c = \{c_1, \ldots, c_J\}$ of Chebyshev polynomials with basis functions $b_j(s)$. These coefficients are obtained from the constraint that the true function values are equal to the approximation at the Gauss–Lobatto grid of the states whose number drives the solution costs.

The *tensor operator* extends a univariate approximation to the multivariate case by combining each element of the univariate grids and basis functions with each other. The univariate approximation of function x is given by

$$U^i(x) = \sum_{j=1}^{m_i} a_j^i x(s_j^i)$$

with approximation level $i \in \mathbb{N}$ and grid points s_j^i. The a_j^i are functions in a function approximation and weights in a numerical integration.

The multidimensional ($d > 1$) tensor product operator $\otimes$ is defined as

$$(U^{i_1} \otimes \cdots \otimes U^{i_d})(x) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_d=1}^{m_{i_d}} \left(a_{j_1}^{i_1} \otimes \cdots \otimes a_{j_d}^{i_d}\right) x\left(s_{j_1}^{i_1}, \ldots, s_{j_d}^{i_d}\right).$$

The number of grid points $\prod_{j=1}^{d} m_{i_j}$ and, thus, the number of function evaluations needed to identify the coefficients of the approximation c grows exponentially with d.

The *Smolyak operator* is given by

$$A_{q,d}(x) = \sum_{q-d+1 \le |i| \le q} (-1)^{q-|i|} \binom{d-1}{q-|i|} (U^{i_1} \otimes \cdots \otimes U^{i_d}),$$

where d is the dimensionality of the function to be approximated, q is the level of approximation, and $i = \{i_1, \ldots, i_d\}$ is a multiindex with $|i| = \sum_j i_j$. This formula highlights the fact that the Smolyak operator is a simple linear combination of some lower level tensor products.

The Smolyak operator constructs the multivariate grid as a combination of some lower level Cartesian products $\times$ of the univariate grids s^{i_j}:

$$H_{q,d} = \bigcup_{q-d+1 \le |i| \le q} (s^{i_1} \times \cdots \times s^{i_d}). \tag{6}$$

We illustrate the Smolyak operator step by step in the Supplemental Material (Winschel and Krätzig (2010)). For alternative presentations, see Kübler and Krüger (2004), Heiss and Winschel (2008), Bungartz and Griebel (2004), or Barthelmann, Novak, and Ritter (1999). The last paper is the closest to our

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

application and contains the following result: for a smooth policy function $x(s)$ (i.e., $x \in F_d^k = \{f : [a, b]^d \to \mathbb{R} | D^\alpha f$ is continuous if $\alpha_i \leq k \; \forall i\}$), the accuracy of a polynomial Smolyak approximation is $O(n^{-k}(\log n)^{(k+1)(d-1)})$, where n is the number of grid points.

The error of the approximation is controlled by the Euler error proposed by Judd (1992), which is a normalization of the residuals of the first-order conditions. Dividing the residual by $(1 - l_n(s_t))^{(1-\theta_n)(1-\tau_n)}/(\kappa_n i_{n,t} - 1)$ and taking it to the power of $1/(\theta_n(1 - \tau_n) - 1)$ gives the Euler equation in terms of consumption:

$$r^c(s) = c_n(s) - \left(\frac{\beta \mathrm{E}_t h(s, x^*(s), e, s', x^*(s'); \theta)}{(1 - l_n(s))^{(1-\theta_n)(1-\tau_n)}/(\kappa_n i_{n,t} - 1)} \right)^{1/(\theta(1-\tau)-1)},$$

$$s' = g(s, x^*(s), e').$$

The Euler error is finally given by $r^E = |r^c(s)/c_n(s)|$. A $\log_{10}$ error of -3 means that the utility loss due to the approximation is less than one per 1000 dollars.

3. ECONOMETRICS

The Bayesian estimation given a model M_i is based on the joint density

$$p(\omega, y, \theta_{M_i} | M_i) = p(\omega | y, \theta_{M_i}, M_i) p(y | \theta_{M_i}, M_i) p(\theta_{M_i} | M_i)$$

of observables y, variables of interest ω, and unobservable parameters and states $\theta_{M_i} = \{\theta\} \cup \{s\}$. The factorization describes functional forms and shock distributions for the variables of interest $p(\omega | y, \theta_{M_i}, M_i)$, the likelihood $p(y | \theta_{M_i}, M_i)$, and the prior density of unobservables $p(\theta_{M_i} | M_i)$. The likelihood transforms the prior into the posterior of unobservables

$$p(\theta_{M_i} | y, M_i) = \frac{p(y | \theta_{M_i}, M_i) p(\theta_{M_i} | M_i)}{p(y | M_i)}$$

$$= \frac{p(y | \theta_{M_i}, M_i) p(\theta_{M_i} | M_i)}{\int p(y | \theta_{M_i}, M_i) p(\theta_{M_i} | M_i) \, d\theta_{M_i}}.$$

The marginal likelihood $p(y | M_i)$ integrates out the parameters and transforms a model prior $p(M_i)$ into its posterior $p(M_i | y)$ whose ratio, the Bayes factor

$$\frac{p(M_i | y)}{p(M_j | y)} = \frac{p(M_i)}{p(M_j)} \frac{p(y | M_i)}{p(y | M_j)},$$

selects the model with the best in-sample forecast quality, that is, the one which allows a higher compression of the explained data. This criterion is valid for nonnested models with different functional forms or shock densities.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

3.1. *Filtering*

The model solution $x^*(s)$, the state transition equation $s_t = g^*(s_{t-1}, x^*(s_{t-1}), e_t) = g(s_{t-1}, e_t) \Leftrightarrow p(s_t|s_{t-1})$, and a measurement equation $y_t = m^*(s_t, x^*(s_t)) + \varepsilon_t = m(s_t) + \varepsilon_t \Leftrightarrow p(y_t|s_t)$ form a nonlinear state space model.

For a given parameter vector and prior of the unobserved state $p(s_0) = p(s_0|y_0)$, the data of each period y_t transform the previous posterior density of unobserved states $p(s_{t-1}|y_{1:t-1})$ in a prediction and filtering step into the next posterior $p(s_t|y_{1:t})$.[6]

In the prediction step, the prior is formed according to the Chapman–Kolmogorov equation

$$p(s_t|y_{1:t-1}) = \int p(s_t, s_{t-1}|y_{1:t-1})\, ds_{t-1} = \int p(s_t|s_{t-1})\, p(s_{t-1}|y_{1:t-1})\, ds_{t-1}$$

and is updated in the filtering step to the new posterior

$$p(s_t|y_{1:t}) = \frac{p(s_t, y_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} = \frac{p(y_t|s_t)\, p(s_t|y_{1:t-1})}{\int p(y_t|s_t)\, p(s_t|y_{1:t-1})\, ds_t}$$

with respect to y_t, which becomes the prior with respect to new data y_{t+1} in the next iteration. The normalizing constant $l_t = \int p(y_t|s_t)\, p(s_t|y_{1:t-1})\, ds_t = p(y_t|y_{1:t-1})$ is the period's contribution to the likelihood of the complete sample

$$\mathcal{L}(\theta; y_{1:T}) = p(y_{1:T}|\theta) = \prod_{t=1}^{T} p(y_t|y_{1:t-1}, \theta) = \prod_{t=1}^{T} l_t.$$

The likelihood transforms the prior parameter density into the posterior

$$p(\theta|y_{1:T}) = \frac{p(y_{1:T}|\theta)\, p(\theta)}{p(y_{1:T})}.$$

An overview of various approaches to this problem can be found in Arulampalam, Maskell, Gordon, and Clapp (2002). The following subsections summarize the main ideas for the filters. Detailed formulas can be found in the Supplemental Material.

3.1.1. *Smolyak Kalman Filter*

The extended Kalman filter linearizes the state space equations and then applies the Kalman filter. A widely used improvement is the deterministic unscented filter by Julier and Uhlmann (1997), where the prediction $p(s_t|y_{1:t-1})$ and posterior densities $p(s_t|y_{1:t})$ are approximated by Gaussians whose first

[6]The notation $y_{1:t}$ is a shorthand for $\{y_1, \ldots, y_t\}$.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

two moments are updated by the Kalman method. The unscented integration of these two moments addresses the curse of dimensionality by restricting the number of grid points to rise linearly with the dimension and there by shifts the curse to a rising inaccuracy of the approximation. Our Smolyak Kalman filter uses the theoretically sound Smolyak Gaussian quadrature instead. Its advantage is that the approximation level can be freely chosen and that any shock density with an existing univariate quadrature scheme can be incorporated. Our approach reduces the moment approximation error, but also retains the error from assuming the densities to be Gaussian.

3.1.2. *Particle Filter*

The particle filter updates the posterior density of the unobserved states by sequential importance sampling. It is easy to implement, very general, and arbitrarily accurate, but it needs very large samples since it does not use the smoothness of the integrand and the latest observation in its proposal density $p(s_t|s_{t-1})$, which should also condition on y_t.

3.1.3. *Smolyak Particle Filter*

Our second filter is the Smolyak particle filter which generates the particle filter's proposal density by the densities obtained from the Smolyak Kalman filter. By that it incorporates the latest observation and combines the advantages of both filters—the accuracy of the slow particle filter and the speed of the potentially inaccurate Smolyak Kalman filter.

This idea is similar to that used by Amisano and Tristani (2007), who generated the proposal density by some kind of extended Kalman filter, or to that used by van der Merwe et al. (2000), who used the unscented filter.

In contrast to the particle filter, these filters do not break down in state space models with very small measurement errors or without any errors at all.

3.1.4. *Smolyak Sum Filter*

The basic idea for the Smolyak sum filter is that any density of practical concern can be approximated as a weighted sum of normal densities $p(x) \approx \sum_i \omega_i \mathcal{N}(x; \mu^i, \Sigma^i)$. This filter approximates the prediction and filtering densities and effectively runs several Smolyak Kalman filters in parallel. The weights are updated according to the likelihood contribution of each summand.

The idea can be traced back to Alsbach and Sorenson (1972). More recently, Kotecha and Djurić (2003) revived this approach, but they used costly importance sampling to approximate the involved integrals. Instead, we propose to use Smolyak Gaussian quadrature again.

Anderson and Moore (1979) presented this approach for a model with additive noise in the measurement and state equations. In our setting of nonadditive state shocks, we therefore implicitly assume that the weights are preserved during the prediction step. This simplifying assumption deserves further elaboration.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

3.2. *Posterior Density*

The Metropolis–Hastings algorithm generates draws from a density if this density can be evaluated at any point of its domain. Once we can evaluate the likelihood in $p(y|\theta)\,p(\theta)$, we obtain a histogram of draws of θ that approximate its posterior $p(\theta|y)$. For details, see the introdution in Chib and Greenberg (1995).

The parameter space is traversed by a random walk. A generated candidate parameter $\hat{\theta}_n$ is accepted with the probability given by the ratio between the candidate's posterior and the posterior of the previously accepted parameter $\hat{\theta}_{n-1}$. This qualifies the algorithm as a global maximizer. The survival according to the parameter's fitness, measured by the posterior, qualifies it as a genetic algorithm.

The innovation variance influences the density region covered by the random walk. After convergence, sampling around the mode with a large variance generates candidates far from the current value and results in a low acceptance ratio. Smaller variances increase the acceptance ratio, but decrease the region being covered so that low probability regions may be undersampled. If the innovation variance is tuned to balance this trade-off at an acceptance ratio of around 30%, the algorithm converges to a representative sample drawn from the target density.

The critical choices of the algorithm are the starting value $\hat{\theta}_0$, the convergence detection, and the innovation variance of the random walk.

3.2.1. *Convergence Test*

The starting value $\hat{\theta}_0$ drives the number of burn-in draws to be discarded before a representative region of the target density is reached. Formal convergence tests along with eyeballing make up an important part of the estimation process.

The convergence test we use can be based on one subdivided sequence or on several parallel sequences as in our parallel algorithm. The idea is to diagnose convergence if the sequences appear to be drawn from the same distribution according to a distance measure between the within- and across-sequence variances. Examining only one sequence will result in overly optimistic diagnostic tests. Gelman and Rubin (1992) pointed out that lack of convergence, in many problems, can easily be detected from many sequences but not from one.

3.2.2. *Parallel Mixing Metropolis–Hastings*

The variance for the innovation shock is the major problem of the algorithm, since the optimal variance can be derived only from the variance of the target density. The usual approach is, therefore, to estimate the variance repeatedly from training sequences which are also useful to assure robustness with respect to the start value.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

We propose to run multiple sequences simultaneously and not sequentially. By doing so, we can assure robustness with respect to the start values, unbias the convergence diagnostic test, and, most importantly, estimate the innovation variance on the fly. The idea is to run several sequences $m = 1, \ldots, M$ in parallel and generate the candidate vector in each sequence according to

$$\hat{\theta}^*_m = \hat{\theta}_{m,n} + \gamma^{\mathrm{GE}}(\hat{\theta}_{m_1,n} - \hat{\theta}_{m_2,n}) + \varepsilon, \quad \text{where} \quad \varepsilon \sim \mathcal{N}(0, bI).$$

The innovation from the random shock is common to the standard algorithm but we use only one variance common to all shocks. The additional source of innovation is the difference between the parameter vectors from two randomly chosen sequences m_1 and m_2. The scalar γ^{GE} and the shock variance b determine the relative weight of the mixing and random walk innovations, and are the only two values that need to be tuned for an appropriate acceptance ratio instead of one variance for each estimated parameter.

The intuition of our approach is that the variance of the difference between two randomly drawn parameter vectors is the optimal variance given that the sequence has converged. Our idea originated from the diagnosis test where the within- and between-sequence variances are examined and convergence is detected when they are of a similar size. Since we know that the optimal variance is the scaled variance of the target density, the between-sequence variance is a good estimator of the optimal variance. ter Braak (2006) derived the same candidate by analogy to the global evolutionary optimization algorithm, called differential evolution, by Storn and Price (1997). This evolutionary feature substantially improves the search for the modus of the posterior in the beginning of the sampling process.

3.3. *Marginal Likelihood*

The marginal likelihood $p(y|M_i)$ can be approximated by a simple calculation based on the Metropolis–Hastings draws and their associated posterior values. Gelfand and Dey (1994) showed that for any density $h(\theta_{M_i}|M_i)$, the expected value

$$\mathrm{E}_{p(\theta_{M_i}|y,M_i)}\left(\frac{h(\theta_{M_i}|M_i)}{p(y|\theta_{M_i}, M_i)\,p(\theta_{M_i}|M_i)}\right)$$

approximates the inverse of the marginal likelihood $p(y|M_i)^{-1}$.

4. RESULTS

All results are calculated by JBendge, which is programmed in Java. The hardware we use is a shared memory cluster with 16 CPUs. The parallel Metropolis–Hastings algorithm allows for a simple parallelization.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

The model is parameterized equally for all countries by $\alpha_n = 0.4, \beta = 0.99, \delta_n = 0.02, \rho_n = 0.95, \theta_n = 0.357, \tau_n = 2.0$, and $\sigma_{n,a} = 0.007$. The solutions for the multicountry models are calculated for $\kappa_n = 0.01$. The estimations are done only for the one country model with the parameterization $\tau = 50$ and $\sigma_a = 0.035$. To simplify the estimation, we set capital adjustment costs to zero ($\kappa = 0$) and obtain analytical expressions for the steady state values, $\bar{a} = 0.0, \bar{k} = 23.27, \bar{c} = 1.29, \bar{i} = 0.47, \bar{l} = 0.31$, and $\bar{y} = 1.75$, which are independent of τ and σ_a. The missing investment costs allow us to compare our estimates with those obtained by Fernández-Villaverde and Rubio-Ramírez (2005).

4.1. *Solution*

The multicountry model allows us to vary the number of countries and thus the dimensionality of the approximated policy functions.

Table I documents the results. The tensor approximations are constructed from a minimal 3 point univariate approximation so as to obtain a nonlinear solution. This results in 3^d grid points for the simplest tensor approximation. For the Smolyak approximation, we use levels 2, 3, and 4 for the models with four and six states, and afterward use only levels 2 and 3.

The quality of the approximations is measured by the number of grid points the running time for a solution in seconds, and the maximal absolute Euler error at 10,000 random points in the approximation space.

The Smolyak operator is already superior to the tensor operator for a small model with four states, where the Euler error on a 41 point Smolyak grid is smaller than the error on a tensor grid with 81 points, although the solution time is the same. For the next level with a similar error, the Smolyak operator is more than three times faster and uses about five times less points (137 vs. 625).

The efficiency gain for the model with six states is even more dramatic. Here the Smolyak operator needs only 85 compared to 729 points of the tensor operator for similar approximation accuracy. The running times are accordingly about four times lower for the Smolyak operator.

The tensor operator breaks down for models beyond six states, while the Smolyak operator is still doing fine. The biggest model we were able to solve has 22 states and it took around 1 hour for an approximation error of $1.7E{-}5$.

4.2. *Estimation*

We simulated data sets of 100 observations, starting from the deterministic steady states generated by very accurate nonlinear solutions of the one country model with the state's productivity a, capital k, and one labor decision l. The consumption policy can be solved analytically from the intratemporal trade-off in terms of the labor policy. The observables in the measurement model are investment i, labor l, and output y. The measurement shocks are assumed to be

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

TABLE I

SMOLYAK (S) AND TENSOR (T) BASED SOLUTIONS

States	Op.	Points	Error	Sec.
4	S	9	6.6*E*–4	0.3
4	S	41	8.1*E*–6	2.5
4	S	137	9.3*E*–7	24.0
4	T	81	4.9*E*–5	2.5
4	T	625	1.8*E*–7	88.5
6	S	13	6.2*E*–4	0.7
6	S	85	5.1*E*–5	12.5
6	S	389	9.3*E*–7	201.5
6	T	729	6.5*E*–5	54.1
8	S	17	5.9*E*–4	1.3
8	S	145	3.5*E*–5	29.9
10	S	21	7.5*E*–4	2.3
10	S	221	4.0*E*–5	69.2
12	S	25	4.4*E*–4	3.8
12	S	313	4.8*E*–5	157.8
14	S	29	4.3*E*–4	5.7
14	S	421	3.7*E*–5	339.1
16	S	33	4.5*E*–4	8.5
16	S	545	4.0*E*–5	724.1
18	S	37	3.7*E*–4	12.2
18	S	685	2.6*E*–5	1819.4
20	S	41	3.3*E*–4	17.1
20	S	841	1.9*E*–5	2107.4
22	S	45	3.3*E*–4	23.3
22	S	1013	1.7*E*–5	4087.4

additive. We present the estimates of two variants of the models: one with small $\{\sigma_i, \sigma_l, \sigma_y\} = \{8.66E\text{–}4, 1.1E\text{–}3, 1.58E\text{–}4\}$ and one with large measurement errors $\{\sigma_i, \sigma_l, \sigma_y\} = \{4.65E\text{–}4, 3.12E\text{–}3, 1.75E\text{–}2\}$. The large standard deviations are set to 1% of the steady state values and the small ones are set to the values in Fernández-Villaverde and Rubio-Ramírez (2005).

4.2.1. *Likelihood*

The particle filter is run with 40,000 particles, the Smolyak Kalman filter with integration level 3, the Smolyak particle filter with integration level 2 and 500 particles, and the Smolyak sum filter with integration level 3 and 20 Gaussian summands. All solutions are calculated with Smolyak level 3 for the policy approximation and the rational expectation integrals.

Figures 1 and 2 show slices through the multidimensional likelihood. The left plots show likelihood values from data with small measurement errors and

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

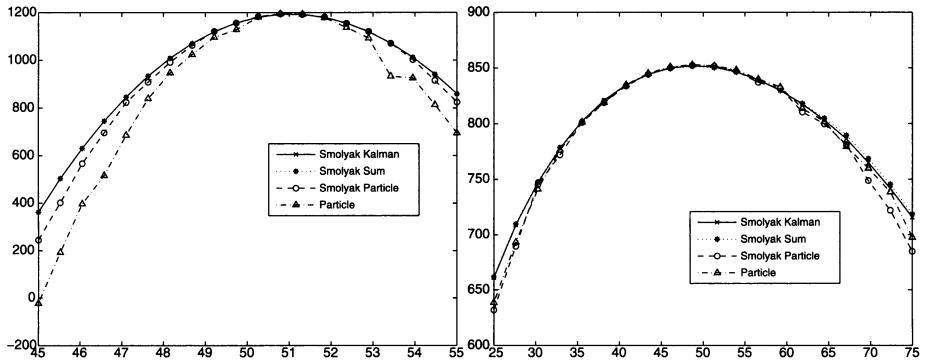

FIGURE 1.—Likelihood at true parameters for $\tau = 50$ and $\sigma_a = 0.035$.

the right plots show values from data with large measurement errors. We set all parameters to their true values and vary τ at the abscissas, and plot them against the likelihood values at the ordinates.

The results are rather encouraging for all our filters as the values are very similar. As expected, the particle filter gets into trouble for small measurement errors in the model with $\tau = 50$ and $\sigma_a = 0.035$ in the left plot of Figure 1.

The running times for one likelihood evaluation are very different: 120 seconds for the particle filter, 0.2 seconds for the Smolyak Kalman filter, 6 seconds for the Smolyak particle filter, 0.5 seconds for the Smolyak sum filter, 0.015 seconds for the linear Kalman filter, and 0.2 seconds for the extended Kalman filter.

The particle filter is hardly useful in combination with a Chebyshev approximation where the policy function interpolation at the particles is the main bottleneck. A possible remedy is to use a finite element solution with its fast interpolation and to accept a larger grid for the same approximation quality.

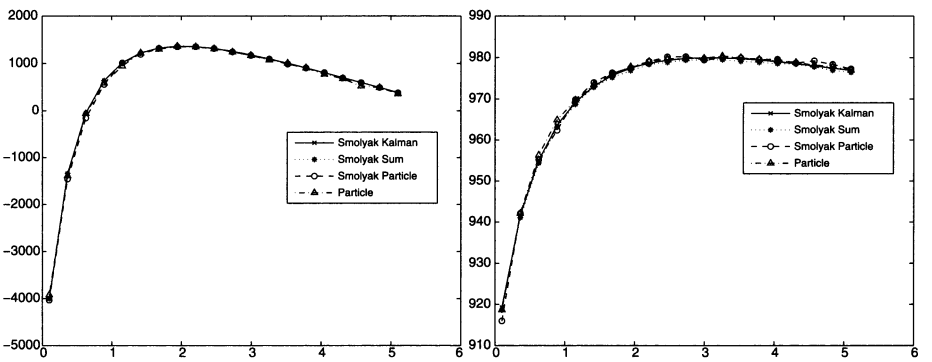

FIGURE 2.—Likelihood at true parameters for $\tau = 2.0$ and $\sigma_a = 0.007$.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

4.2.2. *Parameters*

We have implemented an interactive sampling environment in JBendge where the Metropolis–Hastings parameters γ^{GE} and b can be changed, and individual sequences can be restarted while sampling. This greatly simplifies the maximization and acceptance ratio tuning.

The usual procedure within our framework has three stages. The first one searches for the posterior mode. At this stage the mixing parameter γ^{GE} can be rather large between 0.8 and 2.0 according to the prescription for the differential evolution algorithm of Storn and Price (1997). During this stage the acceptance ratio is usually very low. During the second stage, we sample until the diagnostic tests signal convergence, and tune the parameters γ^{GE} and b to obtain an acceptance ratio around 0.3. For all estimations, parameter b is set to $1E$–6 and γ^{GE} is set between 0.1 and 0.4. Both parameters are remarkably stable across the linear and nonlinear estimations. This helps us to find the appropriate values for the nonlinear estimation by fast linear estimation runs. Once the convergence of the sampler is detected, we get into the third stage, where we sampled 50,000 draws.

The start values of the sequences are random draws from the uniform parameter priors: $[0, 1]$ for α, ρ, θ; $[0.75, 1]$ for β; $[0, 0.05]$ for δ; $[0, 100]$ for τ; $[0, 0.1]$ for all shock variances.

We run the estimation for 13 parameters with 32 parallel sequences. The synchronization necessary for the 16 CPUs results in an average utilization of 50% for the nonlinear estimation, while the linear estimation, scales almost linearly.

It takes about 1500 draws for each of the 32 sequences to find the mode of the posterior density and another 1000–4000 draws to converge. The complete estimation process takes around 5 minutes for the linear estimation, about 2 hours for the nonlinear estimation with the Smolyak Kalman filter, 4 hours with the Smolyak sum filter, and 20 hours with the Smolyak particle filter. The particle filter is much too slow to be of practical use.

In Figure 3 we report representative Smolyak Kalman filter estimates of the model with $\tau = 50$ and $\sigma_a = 0.035$. The vertical bars indicate the true parameter values. The data are generated with a very accurate solution with integration and solution approximation level 5, an Euler error of $1E$–10, and small measurement errors.

The summary of our findings is the following. All nonlinear filters clearly outperform the linearization plus Kalman filter approach. All nonlinear filters have problems accurately estimating the parameter τ, which is biased and has a large standard deviation of the posterior. The standard deviation of the posteriors for the measurement errors of output and investment are large and are of the same magnitude as the estimates themselves. The results are similar with respect to the various filters. Large measurement errors do not deteriorate the mean of the estimates, but consistently increase the standard deviations of the posteriors.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

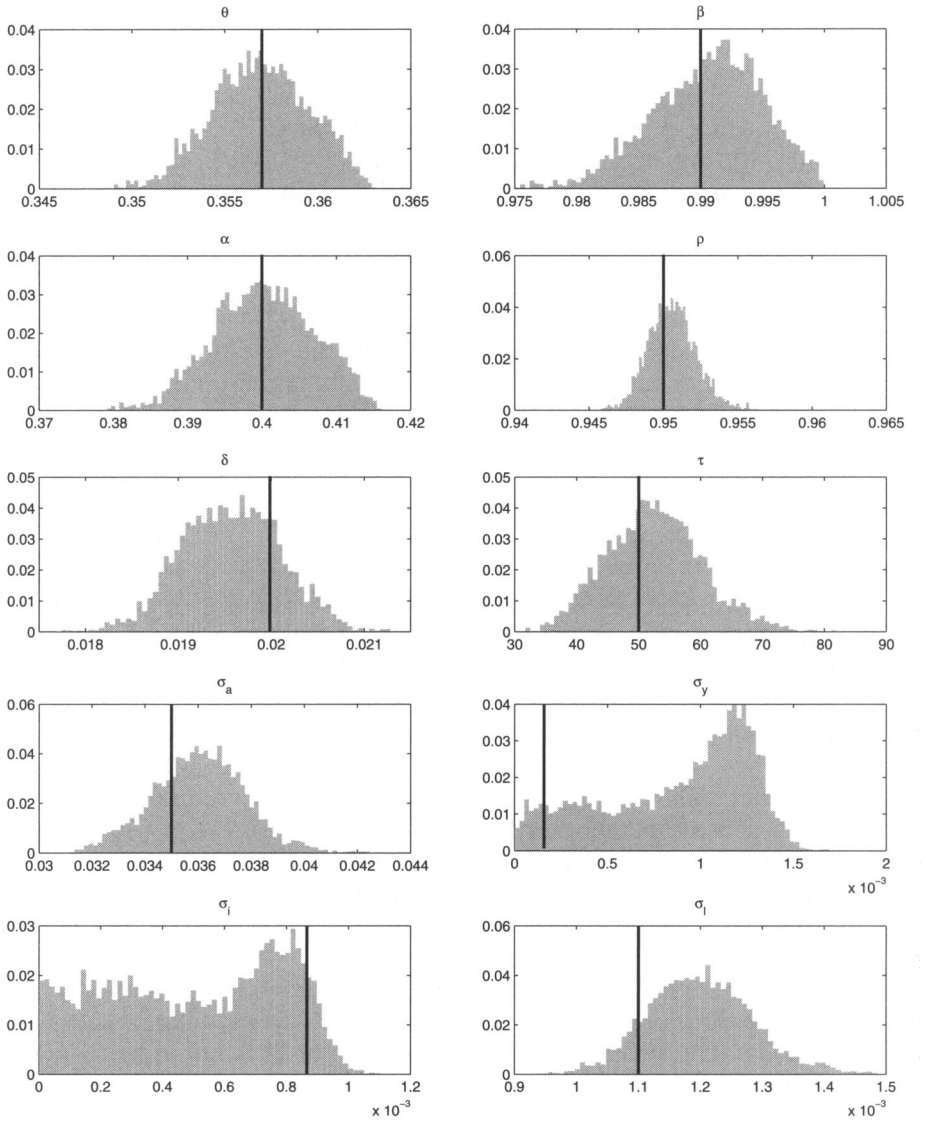

FIGURE 3.—Parameter posterior with the Smolyak Kalman filter.

We investigated the estimation problems for the parameter τ in some detail. Our first suspicion was that this result is due to poor performance of our filters. Since we did not run a full estimation with an accurate particle filter, we had no benchmark. What we did was to compare the likelihood values calculated by the particle filter at several parameter vectors visited by other filters during their sampling process. The finding was that τ is not properly identified

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

in this model as the particle filter also clearly indicates a flat likelihood along the τ dimension. The detailed simulations can be found in the Supplemental Material.

We conclude that the poor estimates of τ are not a problem of our nonlinear filters, but are a feature of the model, and thanks to good global search properties of our parallel Metropolis–Hastings algorithm, we were able to find these parameters.

The last calculation is the marginal likelihood which gives similar results for large and small measurement errors. The Smolyak Kalman filter clearly outperforms the estimation based on the linear Kalman filter. The estimations with the Smolyak sum and Smolyak particle filter are better than that with the Smolyak Kalman filter.

5. CONCLUSION

The Smolyak operator is highly effective for a global solution of models with about 20 states. The operator is also useful for numerical integration essential in many econometric applications.

The Smolyak Kalman filter is very fast and useful if the posterior and prediction densities are reasonably approximated by a Gaussian density. If not, they can be approximated by Gaussian sums in the Smolyak sum filter. The particle filter is of little use in our implementation in combination with a Chebyshev approximation where the interpolation is costly and it results in very costly likelihood evaluations. The Smolyak particle filter is slower than both deterministic filters, but much faster than the particle filter.

The parallelized Metropolis–Hastings algorithm improves the global maximization properties of the serial algorithm, avoids extensive training sequences and robustness checks, simplifies the choice of the innovation variance, provides a less biased convergence diagnostic test, and allows implementation of the algorithms on parallel computers.

A major practical improvement for handling the Metropolis–Hastings algorithm is the interactive graphical user interface of JBendge, which allows a very comfortable estimation and algorithm tuning process.

REFERENCES

ALSBACH, D., AND H. SORENSON (1972): "Nonlinear Estimation Using Gaussian Sum Approximation," *IEEE Transactions on Automated Control*, 17, 439–448. [811]

AMISANO, G., AND O. TRISTANI (2007): "Euro Area Inflation Persistence in an Estimated Nonlinear DSGE Model," Discussion Paper 6373, CEPR. [805,811]

ANDERSON, B. D. O., AND J. B. MOORE (1979): *Optimal Filtering*. Mineola, NY: Dover Publications. [811]

ARULAMPALAM, S., S. MASKELL, N. GORDON, AND T. CLAPP (2002): "A Tutorial on Particle Filters for On-Line Non-Linear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, 50, 174–188. [810]

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

ARUOBA, B. S., J. FERNÁNDEZ-VILLAVERDE, AND J. F. RUBIO-RAMÍREZ (2006): "Comparing Solution Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control*, 30, 2477–2508. [804]

BARTHELMANN, V., E. NOVAK, AND K. RITTER (1999): "High Dimensional Polynomial Interpolation on Sparse Grids," *Advances in Computational Mathematics*, 12, 273–288. [808]

BILLI, R., AND K. ADAM (2007): "Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates," *Journal of Monatary Economics*, 54, 728–752. [803]

BUNGARTZ, H.-J., AND M. GRIEBEL (2004): "Sparse Grids," *Acta Numerica*, 13, 147–269. [804, 808]

CHIB, S., AND E. GREENBERG (1995): "Understanding the Metropolis–Hastings Algorithm," *The American Statistician*, 49, 327–335. [812]

DOUCET, A., N. DE FREITAS, AND N. GORDON (eds.) (2001): *Sequential Monte Carlo Methods in Practice*. New York: Springer. [804]

FERNÁNDEZ-VILLAVERDE, J., AND J. F. RUBIO-RAMÍREZ (2004): "Comparing Dynamic Equilibrium Models to Data: A Bayesian Approach," *Journal of Econometrics*, 123, 153–187. [805]

——— (2005): "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood," *Journal of Applied Econometrics*, 20, 891–910. [805,814,815]

——— (2006): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087. [803,805]

GASPAR, J., AND K. JUDD (2005): "Solving Large-Scale Rational-Expectations Models," *Macroeconomic Dynamics*, 1, 45–75. [804]

GELFAND, A., AND D. DEY (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society, Ser. B*, 56, 501–514. [813]

GELMAN, A., AND D. B. RUBIN (1992): "Inference From Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7, 457–472. [812]

HEISS, F., AND V. WINSCHEL (2008): "Likelihood Approximation by Numerical Integration on Sparse Grids," *Journal of Econometrics*, 144, 62–80. [805,808]

JUDD, K. L. (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58, 410–452. [804,809]

JUDD, K. L., AND S.-M. GUU (1997): "Asymptotic Methods for Aggregate Growth Models," *Journal of Economic Dynamics and Control*, 21, 1025–1042. [804]

JUDD, K. L., AND H.-H. JIN (2002): "Perturbation Methods for General Dynamic Stochastic Models," Discussion Paper, Hoover Institution, Stanford. [804]

JUILLARD, M. (1996): "Dynare: A Program for the Resolution and Simulation of Dynamic Models With Forward Variables Through the Use of a Relaxation Algorithm," Discussion Paper, CEPREMAP. [804]

JULIER, S. J., AND J. K. UHLMANN (1997): "A New Extension of the Kalman Filter to Nonlinear Systems," in *Proceedings of AeroSense: 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, Orlando, FL. [804,810]

KALMAN, R. E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME—Journal of Basic Engineering*, 82, 35–45. [803]

KOTECHA, J., AND P. DJURIĆ (2003): "Gaussian Sum Particle Filter," *IEEE Transactions on Signal Processing*, 51, 2602–2612. [811]

KÜBLER, F., AND D. KRÜGER (2004): "Computing Equilibrium in OLG Models With Stochastic Production," *Journal of Economic Dynamics and Control*, 28, 1411–1436. [808]

SCHMITT-GROHÉ, S., AND M. URIBE (2004): "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, 28, 755–775. [804]

SMOLYAK, S. (1963): "Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions," *Soviet Mathematics—Doklady*, 4, 240–243. [804]

STORN, R., AND K. PRICE (1997): "Differential Evolution—A Simple and Efficient Heuristic for Global Optimization Over Continuous Spaces," *Journal of Global Optimisation*, 11, 341–359. [805,813,817]

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms

TER BRAAK, C. J. (2006): "A Markov Chain Monte Carlo Version of the Genetic Algorithm Differential Evolution: Easy Bayesian Computing for Real Parameter Spaces," *Statistics and Computing*, 16, 239–249. [813]

VAN DER MERWE, R., A. DOUCET, N. DE FREITAS, AND E. WAN (2000): "The Unscented Particle Filter," *Advances in Neural Information Processing Systems*, 13, 584–590. [804,811]

WINSCHEL, V., AND M. KRÄTZIG (2010): "Supplement to 'Solving, Estimating, and Selecting Nonlinear Dynamic Models Without the Curse of Dimensionality'," *Econometrica Supplemental Material*, 78, http://www.econometricsociety.org/ecta/Supmat/6297_Extensions.pdf; http://www.econometricsociety.org/ecta/Supmat/6297_data and programs.pdf. [805,808]

Dept. of Economics, L7, 3-5, University of Mannheim, 68131 Mannheim, Germany; winschel@rumms.uni-mannheim.de

and

School of Business and Economics, Institute of Statistics and Econometrics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany; mk@mk-home.de.

Manuscript received February, 2006; final revision received January, 2009.

This content downloaded from 144.32.240.90 on Mon, 25 Jun 2018 17:01:20 UTC
All use subject to http://about.jstor.org/terms