



*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*



## Estimating dynamic stochastic decision models: explore the generalized maximum entropy alternative

Y. Zheng; A. Gohin

INRA, SMART-LERECO, France

Corresponding author email: [yu.zheng@inra.fr](mailto:yu.zheng@inra.fr)

### **Abstract:**

*This paper proposes a generalized maximum entropy (GME) approach to estimate nonlinear dynamic stochastic decision models. For these models, the state variables are latent and a solution process is required to obtain the state space representation. To our knowledge, this method has not been used to estimate dynamic stochastic general equilibrium (DSGE) or DSGE-like models. Based on the Monte Carlo experiments with simulated data, we show that the GME approach yields precise estimation for the unknown structural parameters and the structural shocks. In particular, the preference parameter which captures the risk preference and the intertemporal preference is also relatively precisely estimated. Compare to the more widely used filtering methods, the GME approach provides a similar accuracy level but much higher computational efficiency for nonlinear models. Moreover, the proposed approach shows favorable properties for small sample size data.*

*Acknowledgment:*

**JEL Codes:** C13, Q18

#2232



# Estimating dynamic stochastic decision models: explore the generalized maximum entropy alternative

## Abstract

This paper proposes a generalized maximum entropy (GME) approach to estimate nonlinear dynamic stochastic decision models. For these models, the state variables are latent and a solution process is required to obtain the state space representation. To our knowledge, this method has not been used to estimate dynamic stochastic general equilibrium (DSGE) or DSGE-like models. Based on the Monte Carlo experiments with simulated data, we show that the GME approach yields precise estimation for the unknown structural parameters and the structural shocks. In particular, the preference parameter which captures the risk preference and the intertemporal preference is also relatively precisely estimated. Compare to the more widely used filtering methods, the GME approach provides a similar accuracy level but much higher computational efficiency for nonlinear models. Moreover, the proposed approach shows favorable properties for small sample size data.

# 1 Introduction

Dynamic stochastic decision models allow the economists to study the individuals' decision choices under uncertainty. They are widely used in macroeconomics, known as dynamics stochastic general equilibrium (DSGE) models, for optimal policy analysis under various structural shocks. There have been many advances in recent literature in solving and estimating dynamic stochastic decision models. These advances have important implications for agricultural economics because these models can be used to model farmers' dynamic decisions such as investment and consumption decisions under risks. Estimating the model parameters allows us to depict the production techniques, the farmers' preferences and other dynamic features for agricultural productions, which are essential for agricultural policy analysis.

Estimating dynamic programming models conventionally requires first solve the model numerically, second estimate the model. Solving the model with perturbation methods (Judd and Guu 1993; Schmitt-Grohé and Uribe 2004) has the highest computational efficiency. The disadvantage of this method is that it is only accurate around the steady state. Projection methods (Judd 1992) provide highly accurate solution over the whole range of state values (Aruoba et al. 2006), but the computation burden is heavy. Especially with the increase of the number of state variables and the approximation order, it suffers from the curse of dimensionality issue. Recent development in dealing with such issue includes applying Smolyak's algorithm and developing the library of sparse grid (e.g. Stoyanov 2015).

After solving the model and obtaining a state space representation, we can use the filters to estimate the latent state variables and obtain the likelihood function (An and Schorfheide 2007; Fernández-Villaverde and Rubio-Ramírez 2005, 2007). If the solution space is linear, the Kalman filter yields the optimal estimation. The Kalman filter is commonly used to estimate large DSGE models where the shocks are smooth and the solution space is considered close to linear.

However, with the increasing interests in high-order risk preferences, non-standard utility functions such as recursive utility (Epstein and Zin 1989), time-varying volatility (Caldara et al. 2012), and with the potentially large shocks in less aggregate models such as agricultural models, linear models are not sufficient to meet the research goals. For nonlinear estimation, the available methods include the extended Kalman filter (first order optimality), the unscented Kalman filter (second order optimality), and the sequential Monte Carlo filter (also called the particle filter). The nonlinear estimation with the nonlinear filters, especially the particle filter, is however, numerically more complicated and extremely time consuming. This is because it is a sampling based method, we need to apply further algorithm (e.g. Bayesian technique with Metropolis-Hasting, expectation maximization algorithm) to maximize the numerical likelihood, and we need to perform the solution and estimation steps sequentially in each loop. As a result, when involving estimating nonlinear DSGE models with the filtering techniques, projection methods are seldom used.

From the prospective of variance minimization, Ruge-Murcia (2007) and Ruge-Murcia (2012) propose the simulated method of moments (SMM) to estimate the nonlinear DSGE models. It is also a sampling based method which requires first solve the model, second generate the simulated data and evaluate the moment conditions.

With all the progress in estimating dynamic stochastic decision models, the maximum entropy method is yet little mentioned. The entropy methods origin from the information theory and are developed by Jaynes (1957) to recover the probability distribution on the basis of partial information. Golan et al. (1996) propose a generalized maximum entropy (GME) approach to estimate nonlinear state space models. They recover the unknown structural parameters and the latent state variables in an explicit nonlinear state space model (a dynamic stochastic decision model after nonlinear solution). Afterwards, Paris and Howitt (1998), Lence and Miller (1998), Lansink (1999) use the GME approach to estimate various ill-posed production problems. Bishop (2006) discusses the use of cross-entropy in machine learning. Judge and Mittelhammer (2011) refer the entropy criteria

as an empirical exponential likelihood, and the maximum entropy method is referred as the maximum empirical exponential likelihood (MEEL) method. The GME approach has several advantages in estimating dynamic stochastic decision models compare to the conventional methods. First, it evaluates directly the equilibrium conditions. As a result, the estimation is nonlinear by nature, and the nonlinear solution is only used to approximate the next period expectations in the Euler equation. Second, the filtering based approaches recover the unknown state first and evaluate the likelihood function second for each testing parameters to obtain the final parameter estimation, while the GME approach recovers the the unknown state and the unknown parameters simultaneously in one step. These two main advantages lead to much higher computational efficiency with the GME approach.

In this paper, we use the GME approach to estimate a nonlinear dynamic stochastic model. The test model is a neoclassical growth model, which is a standard DSGE model but is also flexible to represent a farm decision model under risks. Based on the Monte-Carlo experiments with simulated data, we show that the GME approach provides optimal estimation between accuracy and efficiency. The contributions of the papers are, first, it provides an alternative to estimate the nonlinear dynamic stochastic decision models apart from the conventional methods with the filters. To our knowledge, this is a first attempt to use the GME approach to estimate a DSGE model. Depart from Golan et al. (1996), the model needs to be solved to obtain the state space equations. Second, the paper tests the GME estimator for large shocks and highly nonlinear models (5th order projection), which describes better the agricultural market where the shocks are less smooth. Third, we show that the GME estimator possesses favorable properties for small sample size data, which is useful for economic fields with limited data.

The structure of the paper is organized as follows. Section 2 contains a sketch of the model. Section 3 describes in detail the GME estimation. Section 4 presents the Monte Carlo experiment and the estimation results. Section 5 concludes.

## 2 The Model

### 2.1 Economy Model Representation

We start from a neoclassical growth model. This is the core model for the macroeconomic dynamics, but can also represent a farm decision model for investment, consumption and production decisions. This setting allows us to compare our results with a large number of papers in macroeconomics.

Consider a farm household who uses capital  $K_t$  to produce one good  $Y_t$ . The production income is used for personal consumption  $C_t$  and making investment  $I_t$  on storable capital. The agent's goal is to maximize the discounted expected utility stream of consumption,

$$\max_{C_t, I_t} E_0 \sum_{t=0}^{\infty} \beta^t u(C_t). \quad (1)$$

The utility function takes the power utility form:  $u(c_t) = c_t^{1-\gamma}/(1-\gamma)$ , where  $\gamma$  is the utility curvature that captures a mixture of risk preference and intertemporal preference. The agent's budget constraint is

$$Y_t = C_t + I_t \quad (2)$$

The agent has a Cobb-Douglas production process,

$$Y_t = A_t K_t^\alpha \quad (3)$$

where  $\alpha$  is output elasticity of capital. The total factor productivity (TFP)  $A_t$  follows a stochastic process,

$$\ln(A_{t+1}) = \rho_A \ln(A_t) + \sigma_A \tilde{\epsilon}_{A_{t+1}} \quad (4)$$

where  $\rho_A$  is the productivity persistence,  $\sigma_A$  is the standard deviation of the productivity shock,  $\epsilon_{A_t}$  is the stochastic productivity shock,  $\tilde{\epsilon}_{A_t}$  is i.i.d and  $\tilde{\epsilon}_{A_t} \sim N(0, 1)$ .

Physical capital is owned by the agent, and is quasi-fixed in each period once installed. Its level depends on the last period capital stock  $K_{t-1}$  and investment  $I_{t-1}$ . So that the law of motion for capital is,

$$K_{t+1} = (1 - \delta)K_t + I_t \quad (5)$$

where  $\delta$  is the depreciation rate.

In each period, the agent chooses strategy  $\{C_t, I_t\}_{t=0}^{t=\infty}$  to maximize the expected lifetime utility subject to the intertemporal budget constraint (Equation 2), production constraint (Equation 3), and the capital evolution function (Equation 5). The Euler condition of the dynamics is given as,

$$C_t^{-\gamma} = \beta E_t [C_{t+1}^{-\gamma} (1 - \delta + \alpha A_{t+1} K_{t+1}^{\alpha-1})] \quad (6)$$

It shows that the consumption today is decided by the expected consumption and the expected productivity in the future. Investment is implicitly decided in the Euler condition from the budget constraint.

## 2.2 State Space Representation

The model does not have an analytical solution, we need to solve it numerically. We use a high order (5th order) Chebyshev polynomials method to solve the model (Judd 1992). The nonlinear projection method provides a more accurate solution with the existence of large shocks (Aruoba et al. 2006). As our model is small, the high-order solution for this model is not severely influenced by the curse of dimensionality issue and provides accurate simulated data. We approximate the optimal policy function by interpolating the Chebyshev polynomial basis functions. The  $D$ th-degree approximation of  $C_t(A_t, K_t)$



is a complete polynomial:

$$C_t(K_t, A_t) = \sum_{d_K=0}^{D_K} \sum_{d_A=0}^{D_A} b_{d_K, d_A} \psi_{d_K}(\phi(K_t)) \psi_{d_A}(\phi(A_t)) \quad (7)$$

where  $\psi_d(.)$  are Chebyshev polynomials,  $\phi(.)$  are linear mapping of the  $K_t, A_t$  collocation points to  $[-1, 1]$ ,  $b_{d_K, d_A}$  are the Chebyshev coefficients to be estimated. The detailed solution process is described in Appendix. The solution is performed in GAMS with our own projection code.

The solution of the dynamic model gives us a state-space representation as follows. With  $\mathbf{X}_t = [Y_t, I_t, C_t]^T$  the vector of decision variables,  $\mathbf{S}_t = [K_t, A_t]^T$  the vector of state variables, and  $\boldsymbol{\Theta} = [\beta, \gamma, \alpha, \delta, \rho_A, \sigma_A]^T$  the structural parameter set. The state-space model is presented as,

$$\mathbf{X}_t = f(\mathbf{S}_t, \mathbf{V}_t; \boldsymbol{\Theta}) \quad (8)$$

$$\mathbf{S}_t = g(\mathbf{S}_{t-1}, \mathbf{W}_t; \boldsymbol{\Theta}) \quad (9)$$

where  $f$  and  $g$  are nonlinear functions with the vector of structural parameters  $\boldsymbol{\Theta}$ . Eq.(8) is the observation equation in which the observable decisions variables  $\mathbf{X}_t$  are derived from the unobservable state variables  $\mathbf{S}_t$ .  $\mathbf{V}_t$  are the exogenous shocks such as measurement errors (to avoid singularity). Eq.(9) is the state equation which describes the intertemporal evolution of state variables  $\mathbf{S}_t$ .  $\mathbf{W}_t$  are exogenous shocks such as innovations. The underlying idea is that  $\mathbf{S}_t$  is not directly observable, but we could estimate these unobservable state from the observable data  $\mathbf{X}_t$ , given the function form and the structural parameter set. Meanwhile, the structural parameters can also be estimated from the observable data  $\mathbf{X}_t$ .

### 3 Estimation Methods

#### Maximum Entropy

The generalized maximum entropy (GME) approach we use here is described in Golan et al. (1996). In particular, their approach is used to estimate a dynamic model with unobserved data, such as land quality, unobserved shocks, technical process, etc.. The dynamic model of Golan et al. (1996) matches explicitly the state space representation so they do not need to solve the model. The advantages of the GME approach in estimating DSGE or DSGE-like models are discussed in the introduction. In short, this approach evaluates the equilibrium conditions directly, and it recovers the unknown parameters and the unknown states simultaneously. Moreover, the consistency of the GME estimate does not depend on the validity of assumptions on the distribution of the error terms. The disadvantages of the method are, the statistical inference of this method is not well developed, and the results can be sensitivity to the choices of the prior information of the parameters and the error terms.

In a general form, Jaynes (1957) proposes finding the probability distribution that satisfies the constraints and maximizes the Shannon's entropy criterion (Shannon 1948),

$$H(\mathbf{p}) = - \sum_n p_n \ln(p_n) \quad (10)$$

where  $\mathbf{p} = (p_1, \dots, p_N)'$  is a discrete probability distribution for discrete prior information.

For our empirical estimation, we need to recover the probability distribution of the structural parameter set  $\Theta$ , and the time varying error terms, including  $\epsilon_{At}$  which represent the productivity shocks, and  $\epsilon_{euler_t}$  which represent the measurement errors. With the recovered structural parameters and structural shocks, we are able to recover the evolution process of the latent productivity and capital.

To construct the GME framework, first, we reparameterize the structural parameters  $\theta_i (i = 1, 2, \dots, I)$  and the errors  $\epsilon_{jt} (j = 1, 2, \dots, J; t = 1, 2, \dots, T)$ . Here  $i$  is the index

for the parameters,  $j$  is the index for the errors,  $t$  is the time index. Given the prior information, suppose that the value of each parameter  $\theta_i$  lies within the interval  $[z_{i1}, z_{iK}]$ . We define a set of discrete points (support values)  $z_i = [z_{i1}, z_{i2}, \dots, z_{iK}]'$ , with associated probability weights  $p_i = [p_{i1}, p_{i2}, \dots, p_{iK}]'$ . The unknown  $\Theta$ , which is a I-vector is,

$$\Theta = \mathbf{Z}\mathbf{p} = \begin{bmatrix} \mathbf{z}'_1 & 0 & \dots & 0 \\ 0 & \mathbf{z}'_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{z}'_I \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_I \end{bmatrix} \quad (11)$$

where  $\mathbf{Z}$  is an  $I \times IK$  matrix and  $\mathbf{p}$  is an  $IK$  vector. For each parameter  $\theta_i$ ,

$$\mathbf{z}'_i \mathbf{p}_i = \sum_k z_{ik} p_{ik} = \theta_i \text{ for } i = 1, 2, \dots, I \quad (12)$$

Similarly, suppose the error terms  $\epsilon_{jt}$  lies within the interval  $[v_{jt1}, v_{jtD}]$ . Note here we have one more dimension, time  $t$ . We define a set of discrete points  $v_{jt} = [v_{jt1}, v_{jt2}, \dots, v_{jtD}]'$ , with associated probability weights  $w_{jt} = [w_{jt1}, w_{jt2}, \dots, w_{jtD}]'$ . The unknown shocks  $\epsilon_t$  at time  $t$ , which is a J-vector, is,

$$\epsilon_t = \mathbf{V}_t \mathbf{w}_t = \begin{bmatrix} \mathbf{v}'_{t1} & 0 & \dots & 0 \\ 0 & \mathbf{v}'_{t2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{v}'_{tD} \end{bmatrix} \begin{bmatrix} \mathbf{w}_{t1} \\ \mathbf{w}_{t2} \\ \vdots \\ \mathbf{w}_{tD} \end{bmatrix} \quad (13)$$

where  $\mathbf{V}_t$  is an  $J \times JD$  matrix and  $\mathbf{w}$  is an  $JD$  vector. For each shock at time  $t$   $\epsilon_{jt}$ ,

$$\mathbf{v}'_{jt} \mathbf{w}_{jt} = \sum_d v_{jtd} w_{jtd} = \epsilon_{jt} \text{ for } j = 1, 2, \dots, J; \ t = 1, 2, \dots, T \quad (14)$$

Given the reparameterization, our objective is to find the optimal probability distribution  $(\mathbf{p}, \mathbf{w})$  of the corresponding support values, which maximize the objective entropy.

The empirical program is,

$$\max_{\mathbf{p}, \mathbf{w}} - \sum_i \sum_k p_{ik} \ln(p_{ik}) - \sum_j \sum_t \sum_d w_{jtd} \ln(w_{jtd}) \quad (15)$$

Subject to the equilibrium conditions of the dynamic decision program, and the adding up constraints,

$$C_t^{-\gamma} = \beta E_t [C_{t+1}^{-\gamma} (1 - \delta + \alpha A_{t+1} K_{t+1}^{\alpha-1})] \quad (16)$$

$$K_{t+1} = (1 - \delta)K_t + Y_t - C_t \quad (17)$$

$$Y_t = A_t K_t^\alpha \quad (18)$$

$$\ln(A_{t+1}) = \rho_A \ln(A_t) + \sigma_A \tilde{\epsilon}_{A_{t+1}}, \quad \tilde{\epsilon}_{A_t} \sim N(0, 1) \quad (19)$$

$$\sum_k p_{ik} = 1 \quad \text{for } i = 1, 2, \dots, I \quad (20)$$

$$\sum_d w_{jtd} = 1 \quad \text{for } j = 1, J; \quad t = 1, 2, \dots, T-1 \quad (21)$$

$$p_{ik} > 0, w_{jtd} > 0 \quad \text{for } \forall i, j, t, k, d \quad (22)$$

where  $p_{ik}$  and  $w_{jtd}$  are the probability weights of the supporting values which we have specified in the reparameterization part, and Eq.(16) - (19) correspond to the equilibrium conditions in Eq.(2) - (6).

To bring the above program, especially the Euler equation Eq.(16) to the data, one important assumption is rational expectation. At time  $t$ , the agent observes two time series, consumption  $C_t^{obs}$  and production  $Y_t^{obs}$ . The agent cannot precisely predict the point value of the next period productivity shocks  $\epsilon_{A_{t+1}}$ , but he or she knows the distribution of the shocks. We model the anticipated shocks  $\epsilon_{A_{t+1}}^e$  by Gaussian Quadrature according to the distribution of the real shocks<sup>1</sup> (and the real shocks are to be retrieved in the GME estimation). Consequently, the anticipated next period state  $A_{t+1}^e$  are decided by the anticipated shock  $\epsilon_{A_{t+1}}^e$ , and the anticipated consumption  $C_{t+1}^e$  is decided by the

---

<sup>1</sup>To save space, the Gaussian Quadrature details are in Appendix

anticipated state from the policy function Eq.(8) with the Chebyshev polynomials. This setting adds two more constraints to the GME program,

$$\ln(A_{t+1}^e) = \rho_A \ln(A_t) + \sigma_A \epsilon_{A_{t+1}}^e \quad (23)$$

$$C_{t+1}^e = \sum_{d_K=0}^{D_K} \sum_{d_A=0}^{D_A} b_{d_K, d_A} \psi_{d_K}(\phi(K_{t+1})) \psi_{d_A}(\phi(A_{t+1}^e)) \quad (24)$$

The Chebyshev coefficients are jointly estimated in the GME program by interpolating the basis functions of the state variables  $K_t$  and  $A_t$  into the observed consumption data series,

$$C_t^{obs} = \sum_{d_K=0}^{D_K} \sum_{d_A=0}^{D_A} b_{d_K, d_A} \psi_{d_K}(\phi(K_t)) \psi_{d_A}(\phi(A_t)) + \epsilon_{euler_t} \quad (25)$$

where  $\epsilon_{euler_t}$  is a mixture of measurement errors and approximation errors of the consumption data series. Adding one measurement errors series  $\epsilon_{euler_t}$  is to avoid the singularity problem because we need to have the same number of shocks as the number of observable data series (detailed singularity problem is discussed in Fernández-Villaverde and Rubio-Ramírez 2005, Ruge-Murcia 2007). The empirical Euler condition is rewritten as,

$$(C_t^{obs})^{-\gamma} = \beta E_t [(C_{t+1}^e)^{-\gamma} (1 - \delta + \alpha A_{t+1}^e (K_{t+1})^{\alpha-1})] \quad (26)$$

Finally, the objective entropy (Eq.(15)) is maximized subjective to the constraints Eq.(17) - (26). The time-constant parameters dimension  $I = 6 + 25 = 31$  (with 6 the number of the structural parameters,  $25 = 5^2$  the number of Chebyshev coefficients), and the time-varying shocks dimension  $J = 2$ . By forming Lagrange, the first order conditions provide the basis for the solution  $p_{ik}$  and  $w_{jtd}$ . By the reparameterization definition, the estimated parameter and shocks are,

$$\sum_k \hat{p}_{ik} z_{ik} = \hat{\theta}_i \quad (27)$$

$$\sum_d \hat{w}_{jtd} v_{ktj}^\epsilon = \hat{\epsilon}_{jt} \quad (28)$$

Given the recovered shocks in Eq(20), the estimates of the TFP evolution process are determined by Eq.(4).

## 4 Sampling experiments

### 4.1 Experiment design

In a general case, the GME estimator cannot be expressed in a closed form and its finite sample properties cannot be derived from direct evaluation. In order to test the performance of the GME approach and compare it with the more widely used filtering methods, we perform the Monte Carlo sampling experiments on simulated data. Given the parameter calibration in Table 1, we use a 5th order Chebyshev polynomial approximation to generate production and consumption data series. The data are generated in GAMS version2017 using our own projection code.

The first parameter calibration is a benchmark setting in previous literature (e.g. Fernández-Villaverde and Rubio-Ramírez 2005 and Ruge-Murcia 2012). This setting is realistic for the macro data series and allows us to compare our results with previous literature. The inverse elasticity of intertemporal substitution parameter, or equivalently, the preference parameter ( $\gamma$ ), is set to 2. The output elasticity of capital ( $\alpha$ ) is 0.36 and the depreciation rate ( $\delta$ ) is 0.025. Regarding the TFP evolution process, the standard deviation of the TFP shocks is low (0.04), but higher than 0.007 in Fernández-Villaverde and Rubio-Ramírez (2005).

Table 1: Calibrated Parameters

Parameters	Description	Case 1 (macro)	Case 2 (agr)
$\beta$	discount factor	<i>0.95</i>	<i>0.95</i>
$\gamma$	preference parameter	<i>2</i>	<i>0.75</i>
$\alpha$	output elasticity of capital	<i>0.36</i>	<i>0.36</i>
$\delta$	depreciation rate	<i>0.025</i>	<i>0.05</i>
$\rho_A$	productivity persistence	<i>0.85</i>	<i>0.70</i>
$\sigma_A$	standard deviation of productivity shocks	<i>0.04</i>	<i>0.10</i>

For the second parameter calibration, we introduce realistic calibrations for agricultural models. The depreciation rate is higher (0.05) considering the use of machines, the level of risk aversion is lower (0.75) considering the farmers have long been protected from policies, productivity persistence is lower (0.70), and productivity shocks are higher ( $\sigma_A = 0.10$ ). In this way we introduce the nonlinearity to the economy.

The objective parameter set for the estimation is  $\Theta = (\gamma, \alpha, \beta, \delta, \rho_A, \sigma_A)^T$ . It is worth mentioning that different from the SMM sampling experiments in Ruge-Murcia (2012), where two parameters  $(\alpha, \delta)$  are fixed and the rest four parameters  $(\beta, \gamma, \rho_A, \sigma_A)$  are to be estimated, we do not fix any parameter and we estimate the entire parameter set.

We generate small sample size with 50 observations to reflect limited availability of agricultural data. Our sample size is small compare to a sample size of 100 in Fernández-Villaverde and Rubio-Ramírez (2005), and a sample size 200 in Ruge-Murcia (2012). We also test for sample sizes of 30 and 100.

For the GME estimation, the support values given for the parameters and the error terms are very important. On the one hand, the support values allow the economists give prior informations on the parameters. On the other hand, it permits a possibility of manipulation, and the estimation can be sensitive to the support values. In our experiments, we choose loose priors to avoid manipulation. The discount factor  $\beta$  is generally known to be larger than 0.9, we set the prior between 0.9 and 0.99. The utility curvature  $\gamma$ , has a reference value in previous literature between 0.1 and 3. We set the prior accordingly. The output elasticity of capital is set between 0 and 1. The depreciation rate in agricultural is generally smaller than 10 percent, and we set a prior between 0.01 and 0.15. The persistence of a stationary TFP process should be smaller than 1, so our prior for the TFP persistence is set between 0.01 and 0.99. To allow a large variation in volatility, the standard deviation of productivity is set between 0.001 and 0.15. The intervals of the shocks are set between -1 and 1. Table 2 lists the detailed support values. Furthermore, in previous literature, the optimization routine always starts from the true values (e.g. Ruge-Murcia 2007). Indeed, good starting values can largely facilitate the

Table 2: GME estimation: prior information for the parameters

Parameters	Support values		
	Low	Centre	High
$\beta$	0.9	0.95	0.999
$\gamma$	0.01	1	3
$\alpha$	0	0.5	1
$\delta$	0.01	0.10	0.15
$\rho_A$	0.01	0.5	0.99
$\sigma_A$	0.001	0.1	0.15
$\epsilon_{At}$	-1	0	1
$\epsilon_{euler_t}$	-0.001	0	0.001

optimization and make the estimates more accurate. However, for real data, it is not realistic to initialize from the true values. As a result, in order to test the real feasibility of the method, we start the optimization routine near the centre of the support values  $((\beta_0, \gamma_0, \alpha_0, \delta_0, \rho_{A0}, \sigma_{A0})' = (0.9, 0.9, 0.5, 0.05, 0.5, 0.1)')$ , and these values are not necessarily the true values.

The GME estimation is replicated 100 times for Case 1 and Case 2 to test the accuracy and robustness of the estimator. The empirical properties of the estimator are measured using the root mean square error (RMSE) criteria. In particular, for the estimated parameter  $\hat{\theta}$ , RMSE is the root of the sum of the variance and the squared bias:  $RMSE = \sqrt{(\theta - \hat{\theta}) + Var(\hat{\theta})}$ . The estimation is performed in GAMS with the Conopt solver.

For the Bayesian estimation with the particle filter, we adopt similar prior specification and starting values as in the GME setting. The detailed prior information is reported together with the results in table 4. We use 60,000 particles to get 50,000 draws from the posterior distribution. Since one replication takes more than 10 hours, we are not able to do the replication for 100 times. The estimation is performed in Matlab R2016a.

## 4.2 Results

Table 3 presents the GME estimation results. For the first case with small shocks, which represents the macro calibration, the GME estimator yields precise estimates of the entire parameter set. The estimation bias are small, with the lowest 0.42% and the highest



Table 3: GME estimation: Monte Carlo experiment results

Parameters		Mean	S.D.	Bias		RMSE
$\Theta$	True	Case 1 (macro): Small shocks				
$\beta$	0.95	0.9460	0.0154	-0.0040	(0.42%)	0.0159
$\gamma$	2	2.1490	0.5477	0.1490	(7.45%)	0.5676
$\alpha$	0.36	0.3665	0.0223	0.0065	(1.81%)	0.0233
$\delta$	0.025	0.0261	0.0050	0.0011	(4.28%)	0.0051
$\rho_A$	0.85	0.8346	0.1151	-0.0154	(1.81%)	0.1162
$\sigma_A$	0.04	0.0404	0.0074	-0.0004	(1.08%)	0.0074
Case 2 (agr): Large shocks						
$\beta$	0.95	0.9490	0.0096	-0.0010	(0.11%)	0.0100
$\gamma$	0.75	0.7876	0.4997	0.0376	(5.02%)	0.5011
$\alpha$	0.36	0.3512	0.0227	-0.0088	(2.45%)	0.0244
$\delta$	0.05	0.0499	0.0047	-0.0001	(0.26%)	0.0047
$\rho_A$	0.70	0.7046	0.1161	0.0046	(0.66%)	0.1162
$\sigma_A$	0.10	0.0979	0.0113	-0.0021	(2.11%)	0.0115

Note: Estimation based on 100 replications of 50 period random samples generated from a 5th order Chebyshev approximation. Each replication takes in average 59.27 seconds in GAMS2017.

Table 4: Bayesian estimation with the particle filter results

Parameters		Prior distribution	Posterior distribution		
			Mean	90% HPD interval	
$\Theta$	True		Case 1 (macro): Small shocks		
$\beta$	0.95	uni(0.8,1)	0.9495	0.9444	0.9544
$\gamma$	2	uni(0,10)	2.7757	2.7348	2.8142
$\alpha$	0.36	uni(0,1)	0.3665	0.3568	0.3769
$\delta$	0.025	uni(0,0.1)	0.0250	0.0231	0.0269
$\rho_A$	0.85	uni(0,1)	0.8127	0.7929	0.8346
$\sigma_A$	0.04	invg(0.1,inf)	0.0224	0.0199	0.0250
			Case 2 (agr): Large shocks		
$\beta$	0.95	uni(0.8,1)	0.9556	0.9505	0.9615
$\gamma$	0.75	uni(0,10)	0.7961	0.7550	0.8404
$\alpha$	0.36	uni(0,1)	0.3440	0.3313	0.3548
$\delta$	0.05	uni(0,0.1)	0.0487	0.0448	0.0520
$\rho_A$	0.70	uni(0,1)	0.6737	0.6435	0.7040
$\sigma_A$	0.10	invg(0.1,inf)	0.0499	0.0445	0.0544

Note: Estimation based on 50 period random samples generated from a 5th order Chebyshev approximation. The computing time for Case 1 is 16h13m03s, and for Case 2 is 15h34m18s in Matlab2016a.

7.45%. In particular, the accurate estimation of the depreciation rate, the persistence and the standard deviation of the shocks indicates the the latent capital evolution process, and the latent TFP evolution process are retrieved.

The second case is more realistic for the agricultural models, but the nonlinearity brought by the large shocks may impose more difficulty for the estimation. Again, the GME estimation recovers all the parameters with precision. The bias are small and most of the RMSE is explained by the standard deviation of the estimations. For both case, according to the experiments,  $\beta$  and  $\alpha$  are sharply identified regardless the setting of parameter bounds and error bounds. The parameters describing the characters of the state variables ( $\delta$ ,  $\rho_A$ ,  $\sigma_A$ ) have in average slightly larger bias and RMSE, and are theoretically more difficult to retrieve, but they are also relatively accurately recovered. In all, the preference parameter  $\gamma$  is the most difficult one to be estimated in the experiments. It is correspondingly shown in table 3 that the bias and RMSE of  $\gamma$  are relatively larger than other parameters. This is because the risk preference and the consumption smoothing preference are not easy to capture in the data - the objective function is relatively flat with the change in preference values. However, we are still able to estimate  $\gamma$  accurately when we allow for the Euler errors at a low level (lies within the range  $[-0.001, 0.001]$  as the support values). This is in accordance with Attanasio and Low (2004) who state that when bringing the Euler equation to the data, the presence of measurement error can have large effects on the consistent estimates of the relative risk aversion parameter.

As a comparison, we present the results from the Bayesian estimation with the particle filter. As expected, and as is proved in previous literature (Fernández-Villaverde and Rubio-Ramírez 2005), the Bayesian estimation delivers relatively accurate estimation for the parameters. The estimation outcomes of the two methods are alike. Similarly, the preference parameter  $\gamma$  is the most difficult one to be precisely retrieved. The Bayesian estimation tends to over-estimate  $\gamma$  a little more than the GME estimation. Besides, the GME estimation slightly outperforms the Bayesian estimation in term of the unobserved TFP shocks - the bias are lower for  $\rho_A$  and  $\sigma_A$  under the GME estimation. Overall, both

estimation methods provide relatively good estimates of the parameters. However, in term of the computation burden, the GME estimation is much faster than the Bayesian estimation, which is around 60 seconds compare to more than 10 hours for one replication. We suppose the Bayesian methods with the filters is preferred in macroeconomics may because they have been proved to be robust also for large macroeconomic models with numbers of smooth shocks and many sectors, for which we have not tested for the GME approach. In all, our experiments show that the GME approach provides a more efficient, yet solid, estimates for the small and highly nonlinear dynamic stochastic models.

**Sensitivity to the support values and the initial values** It is mentioned in different literature that one drawback of the GME approach is that it is sensitive to the support values (Lansink 1999, Lansink and Carpentier 2001). For our estimation, it is not addressed as a serious problem because we choose large bounds for support values, as long as the economic meanings of the parameters are satisfied (see Table 2). Our results show that the estimations are not manipulated by the support values. It is in accordance with Lence and Miller (1998) who demonstrate that the GME estimates are not especially sensitive to the choice of parameter bounds. Regarding the initial values (the starting values), not only the GME estimation, but all the optimization problems depend more or less on them. Good initial values, sometime the true ones, yield good estimates, while bad initial values, results in very different results. The Monte Carlo experiments show that the GME estimation retrieves the parameters when the initial values are away from the true ones. We also test different initial values to ensure that the parameter estimates always return to the true values.

**Small sample properties** Table 5 shows the small sample property of the GME estimation. As expected, the estimation on the large sample size ( $T = 100$ ) outperforms the estimation on the small sample size ( $T = 30$ ), especially for the preference parameter  $\gamma$ . The estimation bias and RMSE of  $\gamma$  decrease steadily with the increase of sample size. Most parameters ( $\beta, \alpha, \delta, \rho_A, \sigma_a$ ) can be retrieved under a smaller sample size of 30.

Table 5: Comparing Monte Carlo experiment results with different sample size

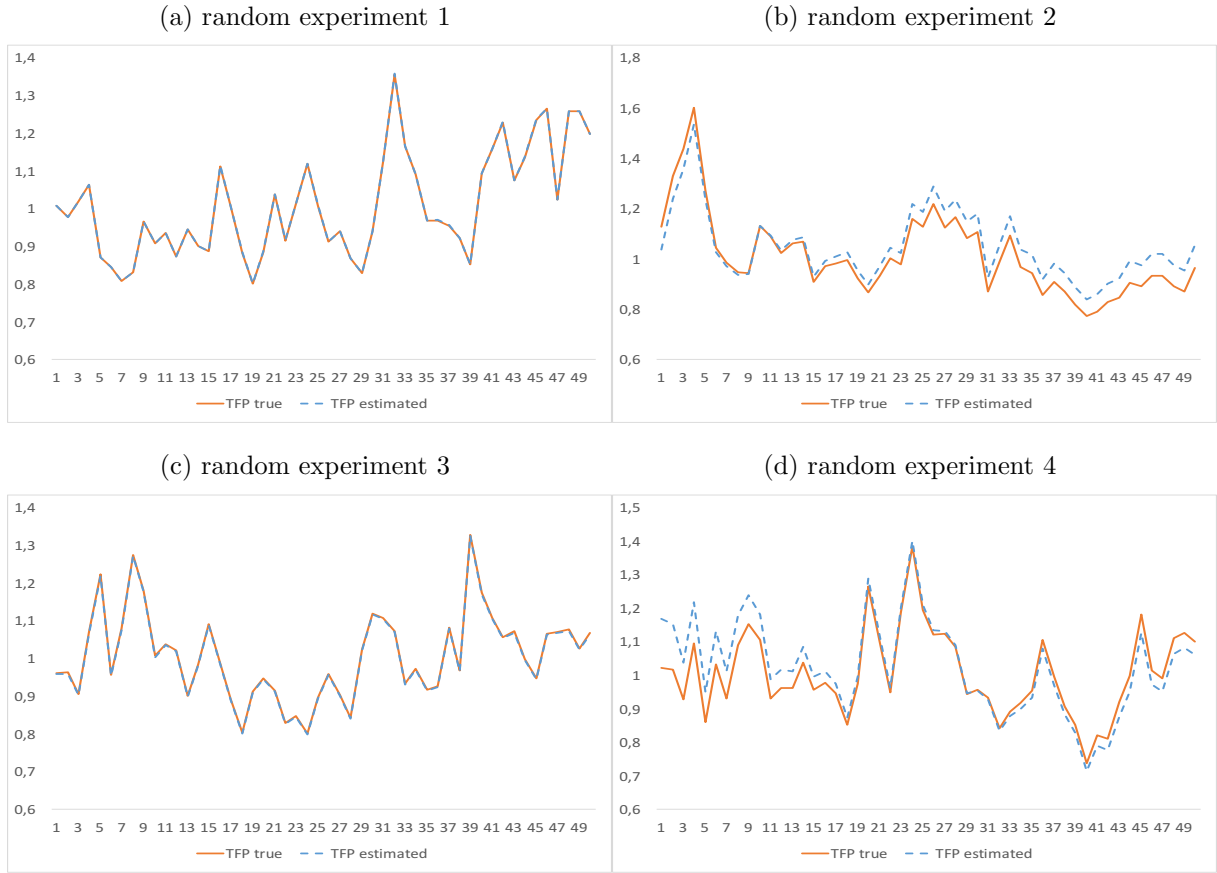
		T=30			T=50			T=100		
Parameters		Mean	S.D.	RMSE	Mean	S.D.	RMSE	Mean	S.D.	RMSE
$\Theta$	True									
$\beta$	0.95	0.949	0.008	0.008	0.949	0.010	0.010	0.947	0.010	0.011
$\gamma$	0.75	0.894	0.601	0.618	0.788	0.040	0.500	0.730	0.283	0.284
$\alpha$	0.36	0.355	0.016	0.017	0.351	0.023	0.024	0.358	0.019	0.019
$\delta$	0.050	0.053	0.010	0.011	0.050	0.005	0.005	0.050	0.003	0.003
$\rho_A$	0.70	0.677	0.083	0.087	0.705	0.116	0.116	0.692	0.117	0.118
$\sigma_A$	0.10	0.097	0.013	0.013	0.098	0.011	0.012	0.100	0.008	0.008

Note: Estimation based on 100 replications of 30, 50, 100 period random samples generated from a 5th order Chebyshev approximation.

Furthermore, with the increase in sample size, the accuracy level of  $\delta$  and  $\rho_A$  have largely improved, while rest parameters ( $\beta$ ,  $\alpha$ ,  $\sigma_A$ ) remain a high accuracy level regardless the sample size.

**Recovering the unknown state** We pick up randomly 4 estimations of the unknown state, total factor productivity (TFP), from the 100 experiments, to show if the unknown state can be recovered from the GME estimation. The state variables are unobserved in real data, but in the experiment, we simulate the random shocks, so that we can access the true values of the shocks and the state. Figure 1 depicts the estimated TFP evolution comparing to the true TFP evolution. We can see that in good cases (Figure 1a, 1c), the estimated TFP matches exactly the evolution of true TFP; In less good (but still satisfactory) cases (Figure 1b, 1d), the estimated TFP moves at the same direction with the true TFP, and the evolution of the TFP process is approximately recovered by the GME estimation.

Figure 1: Comparing the recovered state with the true state from the GME estimation



## 5 Conclusion

In this paper, we propose a GME method that is feasible to estimate nonlinear dynamic stochastic decision models. For these models, the state variables such as productivity are unobserved, and a solution process is needed to obtain an explicit state space representation. To our knowledge, this method has not been used to estimate DSGE or DSGE-like models, and the filtering methods are more widely applied in this field. Based on the Monte Carlo experiments, we show that the GME method recovers estimates of all the unknown structural parameters and the stochastic shocks. In particular, the preference parameter which captures the risk preference and the intertemporal preference is also relatively precisely estimated. Compare it with the Bayesian estimation with the particle filter, we show that the GME approach provides a similar estimation accuracy level, but much higher computational efficiency for nonlinear models. This is because the GME

method does not sequentially solve and estimate the model for each testing parameters, but solve and estimate the model simultaneously given the discrete support values as the prior information. Moreover, the GME estimator shows favorable properties for small sample size data. This is useful for agricultural economics research since the agricultural data series are usually at an annual basis and are not sufficiently long.

Estimating dynamic stochastic decision models has numerous empirical applications for agricultural economics. It allows the economists to structural model the farmers' dynamic decisions, in particular, investment decisions and consumption decisions, under risks, and depict the real values of the structural parameters from estimation instead of calibration. This has important meanings for agricultural policy analysis in response to unknown shocks. Moreover, Griliches and Mairesse (1998) have discussed the problems in estimating production functions, in particular, the data measurement problem for the capital data series and the endogeneity problem to estimate the TFP as a residual. The proposed GME approach may provide feasible solution to these problems, because it can deal with the missing data (latent state variables), and estimating the structural equations is free from the endogeneity problem. Depart from the growth model, the future works involve fitting this classical model to agricultural decision models by adding more agricultural production factors, and estimating the structural parameters and stochastic shocks using real agricultural data.

## References

- An, Sungbae and Schorfheide, Frank. Bayesian analysis of dsge models. *Econometric reviews*, 26(2-4):113–172, 2007.
- Aruoba, S Borağan, Fernandez-Villaverde, Jesus, and Rubio-Ramirez, Juan F. Comparing solution methods for dynamic equilibrium economies. *Journal of Economic dynamics and Control*, 30(12):2477–2508, 2006.
- Attanasio, Orazio P and Low, Hamish. Estimating euler equations. *Review of Economic Dynamics*, 7(2):406–435, 2004.
- Bishop, Christopher M. *Pattern recognition and machine learning*. springer, 2006.
- Caldara, Dario, Fernandez-Villaverde, Jesus, Rubio-Ramirez, Juan F, and Yao, Wen. Computing dsge models with recursive preferences and stochastic volatility. *Review of Economic Dynamics*, 15(2):188–206, 2012.
- Epstein, Larry G and Zin, Stanley E. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969, 1989.
- Fernández-Villaverde, Jesús and Rubio-Ramírez, Juan F. Estimating dynamic equilibrium economies: linear versus nonlinear likelihood. *Journal of Applied Econometrics*, 20(7):891–910, 2005.
- Golan, Amos, Judge, George, and Karp, Larry. A maximum entropy approach to estimation and inference in dynamic models or counting fish in the sea using maximum entropy. *Journal of Economic Dynamics and Control*, 20(4):559–582, 1996.
- Griliches, Zvi and Mairesse, Jacques. Production functions: The search for identification. In *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, number 31, page 169. Cambridge University Press, 1998.

- Jaynes, Edwin T. Information theory and statistical mechanics. *Physical review*, 106(4): 620, 1957.
- Judd, Kenneth L. Projection methods for solving aggregate growth models. *Journal of Economic Theory*, 58(2):410–452, 1992.
- Judd, Kenneth L and Guu, Sy-Ming. Perturbation solution methods for economic growth models. In *Economic and Financial Modeling with Mathematica®*, pages 80–103. Springer, 1993.
- Judge, George G and Mittelhammer, Ron C. *An information theoretic approach to econometrics*. Cambridge University Press, 2011.
- Lansink, Alfons Oude. Generalised maximum entropy estimation and heterogeneous technologies. *European review of agricultural economics*, 26(1):101–115, 1999.
- Lansink, Alfons Oude and Carpentier, Alain. Damage control productivity: An input damage abatement approach. *Journal of Agricultural Economics*, 52(3):11–22, 2001.
- Lence, Sergio H and Miller, Douglas J. Estimation of multi-output production functions with incomplete data: A generalised maximum entropy approach. *European Review of Agricultural Economics*, 25(2):188–209, 1998.
- Paris, Quirino and Howitt, Richard E. An analysis of ill-posed production problems using maximum entropy. *American journal of agricultural economics*, 80(1):124–138, 1998.
- Ruge-Murcia, Francisco. Estimating nonlinear dsge models by the simulated method of moments: With an application to business cycles. *Journal of Economic Dynamics and Control*, 36(6):914–938, 2012.
- Ruge-Murcia, Francisco J. Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics and Control*, 31(8):2599–2636, 2007.



- Schmitt-Grohé, Stephanie and Uribe, Martin. Solving dynamic general equilibrium models using a second-order approximation to the policy function. *Journal of economic dynamics and control*, 28(4):755–775, 2004.
- Shannon, CE. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.
- Stoyanov, Miroslav. User manual: Tasmanian sparse grids v3.0. *Oak Ridge National Laboratory*, 2015.

# Appendix

## Description of the Solution Method

We approximate the optimal consumption policy function  $C_t(A_t, K_t)$  by interpolating the Chebyshev polynomial basis functions. After the labor rule is approximated, production and investment can also be recovered from the equilibrium conditions. We drop the time index  $t$  from now on for the reason of simplicity. The  $D$ th-degree approximation of  $C(A, K)$  is a complete polynomial:

$$C_t(K_t, A_t) = \sum_{d_K=0}^{D_K} \sum_{d_A=0}^{D_A} b_{d_K, d_A} \psi_{d_K}(\phi(K_t)) \psi_{d_A}(\phi(A_t)) \quad (29)$$

where  $\psi_d(\cdot)$  are Chebyshev polynomials,  $\phi(\cdot)$  are linear mapping of  $K, A$  to  $[-1, 1]$ ,  $b_{d_K, d_A, d_p}$  are coefficients to be estimated.

To find these coefficients, the Euler condition Eq.6 is used as an estimating equation.

Trial points

$$\{A_i, K_i\}_{i=0}^{(D_A)(D_K)}$$

are generated from the nodes of the  $D_A, D_K$  degree polynomials, by taking all possible combinations of the node points. Each trial point can be individually applied to the Euler condition to minimize the Euler residual, for a total of  $(D_A D_K)$  equations. This identified system can be solved for the coefficients by a nonlinear root-finding algorithm.

## Gaussian Quadrature to simulate the anticipated shocks

To evaluate the conditional expectation in the Euler equation, we simulate the productivity shocks  $\epsilon_{A_t}$  using the Gaussian quadrature. This involves modeling the error terms as a random variable with Gaussian Quadrature nodes and the corresponding Gaussian Quadrature weights.

In our paper, for productivity shocks that follow a normal distribution with zero mean

and standard deviation  $\sigma$ ,  $\epsilon \sim N(0, \sigma)$ , we use a 5-point Gaussian Quadrature grid with the nodes and weights specified in Table 6 to describe the anticipated shocks:

Table 6: Gauss-Hermite approximation

i	$\epsilon_i$	$w_i$
1	$-2.8570\sigma$	0.0113
2	$-1.3556\sigma$	0.2221
3	0	0.5333
4	$1.3556\sigma$	0.2221
5	$2.8570\sigma$	0.0113