Structural Estimation of Continuous Choice Models

Jørgensen, Thomas Høgholm

Published in:
Economics Letters

DOI:
10.1016/j.econlet.2013.02.027

Publication date:
2013

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Jørgensen, T. H. (2013). Structural Estimation of Continuous Choice Models: Evaluating EGM and MPEC. Economics Letters, 119(3), 287–290. https://doi.org/10.1016/j.econlet.2013.02.027

Structural Estimation of Continuous Choice Models: Evaluating EGM and MPEC

Thomas H. Jørgensen†

December 17, 2012

Abstract

In this paper, I evaluate the performance of two recently proposed approaches to solving and estimating structural models: The Endogenous Grid Method (EGM) and Mathematical Programming with Equilibrium Constraints (MPEC). Monte Carlo simulations confirm that both EGM and MPEC have advantages relative to standard methods. EGM proved particularly robust, fast and straight forward to implement. Approaches trying to avoid solving the model numerically, therefore, seem to be dominated by these approaches.

Keywords: Structural Estimation, Continuous Choice, Endogenous Grid Method (EGM), Mathematical Programming with Equilibrium Constraints (MPEC).

JEL-Codes: C61.

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen, and Centre for Applied Microeconometrics (CAM), `http://www.econ.ku.dk/phdstudent/jorgensen`. Contact: `thomas.h.jorgensen@econ.ku.dk`.

1 Introduction

One of the novelties of structural models is the ability to perform counter factual policy analysis. This requires – besides a realistic model – that researchers uncover the underlying structural parameters. Most existing approaches are notoriously slow and it is, therefore, tempting to calibrate parameters.

The Endogenous Grid Method (EGM) proposed by Carroll (2006) and Mathematical Programming with Equilibrium Constraints (MPEC) proposed by Su and Judd (2012) apply fundamentally different approaches aimed at overcoming the time consuming task of estimating structural models by, e.g., Time Iterations (TI). EGM does this by a small but efficient modification of TI while MPEC abandon the "nested fixed-point" estimation structure, NFXP, which most other approaches follow.

The aim of this paper is to discuss a concrete implementation of these two recently proposed methods and supply new Monte Carlo evidence on performance in terms of speed, accuracy and practical implementation when estimating structural continuous choice models.[1] Hopefully, this will inspire estimation of more realistic models in terms of heterogeneity and uncertainty.

The paper proceeds as follows. Section 2 presents the model used in the analysis. Section 3 briefly discuss the estimation procedures, TI, EGM and MPEC. Section 4 discuss data generation and present Monte Carlo results. Finally, Section 5 discuss and concludes the analysis.

2 The Model and DGP

I use the canonical model of Deaton (1991) where agents solve the infinite horizon problem

$$\begin{aligned}
\max_{\{c\}_{t=0}^{\infty}} \quad & \mathbb{E}_0\left[\textstyle\sum_{t=0}^{\infty}\beta^t u(c_t)\right], \\
\text{s.t.} \quad & a_{t+1} = R(a_t + y_t - c_t), \\
& a_t \geq 0\,\forall t,
\end{aligned}$$

where $0 < \beta < 1$ is the discount factor, R is the real gross interest rate, c_t is consumption in period t, a_t is assets at the beginning of period t, and $y_t \sim \mathcal{N}(\mu_y, \sigma_y^2)$ is stochastic income in beginning of period t. More complicated models could be formulated without changing the results. Preferences are assumed to be CRRA with relative risk aversion, ρ,

$$u(c_t) = \frac{c_t^{1-\rho}}{1-\rho}.$$

It is convenient to formulate the state in this model as total cash-on-hand available in the beginning of period t as $m_t = a_t + y_t$, such that the state in the model evolves as

$$m_{t+1} = R(m_t - c_t) + y_{t+1}. \tag{1}$$

[1] Su and Judd (2012) illustrate the applicability of MPEC to discrete choice models, using the bus-replacement model of Rust (1987) but do not consider explicitly continuous choice models.

3 Estimation Approaches Considered

In this section, I provide a brief introduction to the implemented approaches. The first two, TI and EGM, are based on the nested fixed point (NFXP) approach, in which the model is solved in an inner algorithm for a given set of trial values of parameters. An outer optimization algorithm estimates the structural parameters by varying these, leading to successively solving the structural model. The third approach, MPEC, abandons NFXP and formulates the solution of the model as equilibrium constraints when estimating the structural parameters.

The estimation framework adopted here is Maximum Likelihood. Without changing the results, a method of moments framework could be adopted where moments from the data are matched moments predicted from the model. It is assumed that panel data on consumption are observed with measurement error, such that

$$c_{it}^{data} = c(m_{it}^{data}|\rho) + \varepsilon_{it},$$

where $c(\cdot|\rho)$ is the consumption function predicted by the model and the measurement error is assumed *iid* Gaussian with mean zero and variance σ.[2] The (mean) log likelihood function can be written as

$$\mathcal{L}(\rho; c, c^{data}, m^{data}) = -\log(\sigma) - \sum_{i}^{N} \frac{1}{NT_i} \sum_{t}^{T_i} \frac{1}{2\sigma^2} \left(c_{it}^{data} - c(m_{it}^{data}|\rho) \right)^2. \tag{2}$$

Since the consumption function in the present model has no closed form solution, $c(m|\rho)$ is found numerically. TI and EGM find $c(m|\rho)$ for a given ρ and use that solution to evaluate the likelihood function. MPEC estimate $c(m|\rho)$ and ρ jointly. The solutions from each of the methods are indistinguishable, as shown in Figure 1.

I use $Q = 8$ Gauss-Hermite nodes (y^q) and weights (w^q) to approximate expectations with regard to labor market income, y. Consumption is approximated by 200 unequally spaced grid points over m_t, with more mass at the bottom of the distribution. In EGM, the grid for m_t is determined endogenously, as discussed below. Linear interpolation is applied between grid points.

All approaches are implemented in MATLAB 2012b using the KNITRO solver for optimization (see Byrd, Nocedal and Waltz, 2006) on a laptop with Intel® Core™ i5-2520M CPU @ 2.50 GHz and 4GB RAM. Code are available from authors webpage.

3.1 Time Iterations (TI)

The Euler residual from the present model is as a nonlinear equation in consumption, c_t,

$$\begin{aligned} \mathcal{E}(c_t|m_t) &\equiv R\beta\mathbb{E}[u_c(c_{t+1})|m_t] - u_c(c_t), \\ &\doteq R\beta \sum_{q=1}^{Q} w^q \check{c}_{t+1}(\underbrace{R(m_t - c_t) + y^q}_{m_{t+1}})^{-\rho} - c_t^{-\rho}, \end{aligned} \tag{3}$$

[2] Alternatively, the estimation could be framed as measurement error in the difference in log consumption or assets without changing the results.

where $\check{c}_{t+1}(m_{t+1})$ represents a linear interpolation function. A numerical procedure, such as bisection or Newton iterations, is used to find optimal consumption that puts the residual in (3)to zero,

$$\begin{aligned} c_t^*(m_t) &: \mathcal{E}(c_t^*|m_t) = 0, \\ \text{s.t.} & \quad c_t \leq m_t. \end{aligned}$$

In order to find the stationary solution to the infinite horizon model, iterate over time until $\max_m\{|c_t^*(m) - c_{t+1}^*(m)|\} < 1.0\text{E}^{-7}$.

3.2 Endogenous Grid Method (EGM)

The EGM proposed by Carroll (2006) modifies time iteration by defining the interpolation grid over end-of-period assets, a_t, instead of beginning-of-period cash-on-hand, m_t. This trick facilitates an analytical solution to optimal consumption today by inverting the Euler equation,

$$\begin{aligned} c_t^*(m_t) &= u_c^{-1}\left(R\beta\mathbb{E}[u_c(c_{t+1})|m_t]\right), \\ &\doteq \left(R\beta\sum_{q=1}^{Q} w^q \check{c}_{t+1}(Ra_t + y^q)^{-\rho}\right)^{-\frac{1}{\rho}}, \end{aligned} \tag{4}$$

where the rhs now is independent of c_t. Since no numerical methods are needed to find optimal consumption (contrary to time iteration), the method dramatically increases speed. Finding the stationary solution is done as for time iterations above.

Cash-on-hand today, m_t, consistent with end-of-period assets, a_t, and consumption, c_t^*, is determined endogenously as

$$m_t = c_t^*(m_t) + a_t.$$

EGM perfectly tracks the credit constraint. This is because the lowest point in the grid over a_t, $\underline{a} = 1.0\text{E}^{-6}$, is (very close to) the point where agents are on the curb of being credit constrained. This is illustrated in the right panel of Figure 1. Including the interpolation point $(m, c) = (0, 0)$ ensures the credit constrained level of cash-on-hand is handled correctly.

3.3 Mathematical Programming with Equilibrium Constraints (MPEC)

Su and Judd (2012) propose formulating the solution and estimation problem as a joint constrained maximization problem. The intuition is that NFXP spent most of the time solving models with high accuracy for "wrong" parameters. The behavior only needs to be optimal at the true parameters. Formalized as a nonlinear constrained optimization problem,

$$\begin{array}{rcll}
\max\limits_{c,\rho} & & \mathcal{L}(\rho; c, c^{data}, m^{data}) & \\
\text{s.t.} & & & \\
1 & < & \rho, & (5)\\
0 & \leq & c \quad \leq m - c, & (6)\\
0 & \geq & \beta R \mathbb{E}\left[u'\left(c(R(m-c)+y)\right)\right] - u'(c), & (7)\\
0 & = & (m-c)\left(\beta R \mathbb{E}\left[u'\left(c(\cdot)\right)\right] - u'(c)\right), & (8)
\end{array}$$

where $\mathcal{L}(\cdot)$ is the likelihood function in (2), (5) is a lower bound on the risk aversion parameter, (6) are lower and upper bounds on the consumption parameters, (7) is the Euler residual formulated as a nonlinear inequality constraint, and (8) is a complementarity constraint, stating that if the credit constraint is not binding, the Euler equation must hold.

The consumption function is estimated along with the structural parameters. Hence, the number of parameters is the number of grid points used to approximate consumption in addition to the structural parameters. Here, that amounts to 201 parameters.

Convergence problems due to loose inner-loop stopping criteria are avoided completely. Inner loop iterations are simply not performed in MPEC. In practice, however, supplying good starting values for consumption parameters was necessary to obtain convergence to the right optimum.

4 Monte Carlo Comparison

To asses the performance of the approaches described in Section 3, synthetic data (5000 individuals in 10 time periods) are generated for value of $\beta \in \{.70, .95, .99\}$. To mitigate the influence of stochastic draws, I perform 50 Monte Carlo runs for each β.

Table 1 reports the Root Mean Squared Error (RMSE) and Monte Carlo Standard Error (MCSTD) along with average time used, the standard deviation of time use across MC runs, and the number of iterations used by each method. Iterations at level 1 refers to the outermost optimization, level 2 refers to iterations until convergence to the infinite horizon stationary solution, and level 3 refers to the innermost numerical procedure, finding the optimal consumption. The three methods differ in the levels of iteration. EGM circumvents the inner most procedure while MPEC only operates on the outer level. All approaches are initialized using the same starting value for ρ.

As expected, TI is slowest overall and both TI and EGM (which both rely on NFXP) is slowed by higher values of β. EGM does, however, seem to be less sensitive to β relative to TI. MPEC should be roughly invariant to the level of the discount rate and the variation across β-values reflect the difficulties in supplying good starting values for consumption paramters rather than the effect of changing β. This instability is also reflected in the relatively large dispersion in time to convergence across MC runs (column 4) for MPEC. The large RMSE of 0.049 when $\beta = 0.7$ stems from MPEC not converging to right optimum in five of the MC runs.

EGM and TI use the same number of level 1 and 2 iterations. The great speed gain from EGM is clearly stemming from the elimination of the inner most searches for optimal consumption

Table 1 – Monte Carlo Comparison.

β		RMSE	MCSTD	Time (secs)	Std. time	Iterations level 1	Iterations level 2	Iterations level 3
	TI	0.002	0.002	26.0	0.48	5	142	147,900
.70	EGM	0.002	0.002	0.1	0.03	5	174	–
	MPEC	0.049	0.046	112.4	269.97	123	–	–
	TI	0.009	0.006	650.7	6.80	5	3,473	3,621,124
.95	EGM	0.006	0.006	1.9	0.05	5	3,636	–
	MPEC	0.009	0.006	93.7	37.00	94	–	–
	TI	0.000	0.000	1,682.6	15.74	6	9,215	8,475,336
.99	EGM	0.000	0.000	5.0	0.08	6	9,247	–
	MPEC	0.000	0.000	30.9	6.26	23	–	–

Notes: Based on 50 MC runs with $N \cdot T = 5000 \cdot 10$ simulated observations each run. Columns 3, 5, 6 and 7 are Monte Carlo averages. only ρ is estimated. $R = 1.05$, $\mu_y = 10$, $\sigma_y^2 = 100$ and 200 grid points are used to approximate consumption.

(level 3), that TI suffers from. MPEC use significantly more level 1 iterations due to the fact that 201 parameters are estimated in MPEC. Since MPEC only operates on the outer level, the approach is considerably faster than TI.

EGM outperforms MPEC on both speed and RMSE. EGM is able to uncover the structural parameter in less than ten seconds while MPEC uses around 50 seconds and TI use 30 minutes to complete the same task. Due to EGMs relatively straight forward reformulation of time iterations, this result is very encouraging.

5 Discussion

Through this analysis, two recent proposed approaches to structural estimation, EGM and MPEC, have been evaluated. The theoretically appealing constraint optimization approach, MPEC, proved to be somewhat disappointing. Even if researchers apply state of the art solvers to problems supplied with (correct) gradients, hessian and sparsity pattern, the size limitation on the solvable problems is a significant constraint. Problems that are not sparse with large state space dimensions would require an intimidating amount of memory. This limitation is also recognized by Su and Judd (2012, p. 2215).

The size limitations of MPEC effectively rules out (realistic) finite horizon models since the number of parameters and constraints are the number of time periods multiplied the number of grid points in addition to the structural parameters, $T \cdot n + k$. Furthermore, using simulation based estimation methods, such as indirect inference or simulated method of moments are generally not feasible in the MPEC framework. A small perturbation in a consumption parameter requires (costly) re-simulation of synthetic data.

EGM proved very robust and fast. The small change to time iteration is very straight forward to implement. Furthermore, EGM includes the exact point where agents are on the curb of being liquidity constrained, increasing accuracy. EGM (as well as TI and MPEC) can also handle continuous-discrete choice models, see, e.g., Iskhakov, Rust and Schjerning (2012) who generalize EGM to handle discrete choices.

The fact that structural parameters can be estimated in a fraction of the time conventional methods require has widespread implications. Heterogeneous parameters and correlated uncertainty could be some of many "new" improvements in structural models. These features have often not been feasible to implement in structural estimation. This also means that several approaches trying to avoid solving the model numerically, such as non-linear GMM estimation (Alan, Attanasio and Browning, 2009) or Synthetic Residual Estimation (Alan and Browning, 2010), are dominated by MPEC and EGM.

Acknowledgments

I wish to thank Kenneth Judd and participants at the Initiative for Computational Economics, especially Vera Molitor, Cormac O'Dea, and Kathrin Schlafmann. I have also recieved valuable comments from Anders Munk-Nielsen, Bo Honoré, and Bertel Schjerning.

References

Alan, S., O. Attanasio and M. Browning (2009): "Estimating Euler equations with noisy data: two exact GMM estimators," *Journal of Applied Econometrics*, 24(2), 309–324.

Alan, S. and M. Browning (2010): "Estimating Intertemporal Allocation Parameters using Synthetic Residual Estimation," *The Review of Economic Studies*, 77(4), 1231–1261.

Byrd, R. H., J. Nocedal and R. A. Waltz (2006): "KNITRO: An Integrated Package for Nonlinear Optimization," in *Large-Scale Nonlinear Optimization*, ed. by G. di Pillo and M. Roma, pp. 35–59. Springer: New York.

Carroll, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91(3), 312–320.

Deaton, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59(5), 1221–1248.

Iskhakov, F., J. Rust and B. Schjerning (2012): "A generalized endogenous grid method for discrete-continuous choice," Working paper.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

Su, C.-L. and K. L. Judd (2012): "Constrainted Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80(5), 2213–2230.

A Figures

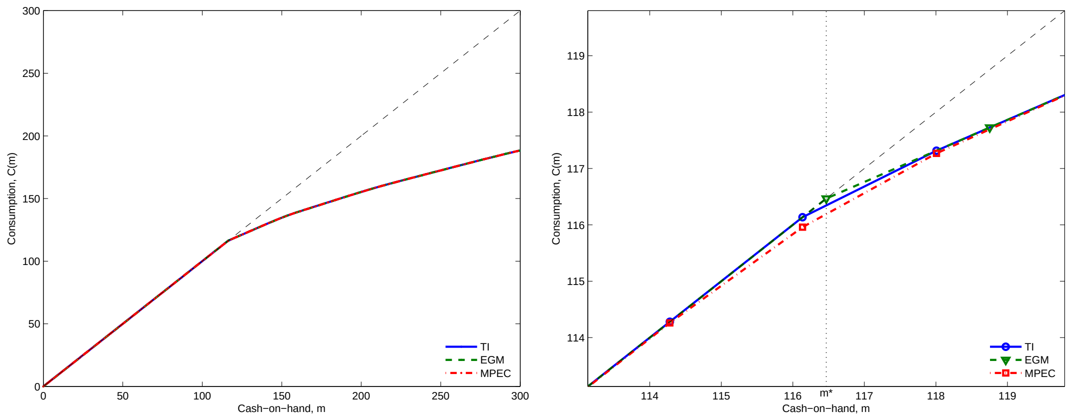

Figure 1 – The Consumption Function, $c(m|\rho)$, from TI, EGM and MPEC.