Supplemental Material:
Robust Estimation of Finite Horizon Dynamic Economic Models

Thomas H. Jørgensen   Maxime Tô

June 30, 2017

A Example 1: Intertemporal Consumption Allocation

The buffer-stock model pioneered by Carroll (1992) and Deaton (1991) and estimated first by Gourinchas and Parker (2002) and Cagetti (2003). The model can be normalized with permanent income, P_t, and re-formulated as

$$\begin{aligned}
V_t(m_t) &= \max_{m_t \geq c_t \geq 0} \frac{c_t^{1-\rho}}{1-\rho} + \beta_t \mathbb{E}_t[(G_{t+1}\eta_{t+1})^{1-\rho} V_t(m_{t+1})] \\
&\text{s.t.} \\
m_{t+1} &= (G_{t+1}\eta_{t+1})^{-1}(1+r)a_t + \varepsilon_{t+1} \\
\log \varepsilon_t &\sim \mathcal{N}(-.5\sigma_\varepsilon^2, \sigma_\varepsilon^2) \\
\log \eta_t &\sim \mathcal{N}(-.5\sigma_\eta^2, \sigma_\eta^2)
\end{aligned}$$

where small letters denotes normalized values.

A.1 Solving the Model

To solve the model, we employ the endogenous grid method proposed by Carroll (2006) and use the normalized consumption Euler equation

$$c_t^{-\rho} = \max\left\{ m_t^{-\rho}\ ,\ \beta_t(1+r)\mathbb{E}_t\left[(c_{t+1}G_{t+1}\eta_{t+1})^{-\rho}\right]\right\}$$

such that when agents are off the credit constraint, we can invert the Euler equation to get for the jth grid point in $\vec{a} = (a^1, \ldots, a^{\#})$,

$$c_t^j = \beta_t(1+r)\sum_{k=1}^{Q}\sum_{l=1}^{Q}\omega^k\omega^l(\mathring{c}_{t+1}(m_{t+1}^{jkl})G_{t+1}\eta_{t+1}^k)^{-\rho}$$

where $\mathring{c}_{t+1}(m_{t+1}^{jkl})$ is a linear interpolation function evaluated at grid point j for the kth and lth Gauss-Hermite nodes associated with the permanent and transitory income shock, respectively, $m_{t+1}^{jkl} = (G_{t+1}\eta_{t+1}^k)^{-1}(1+r)a^j + \varepsilon_{t+1}^l$. The weights associated with these quadrature nodes are denoted ω^k and ω^l. We use 100 discrete points to approximate the consumption function ($\# = 100$) and 8 quadrature points to approximate each of the income shock integrals ($Q = 8$). The endogenous level of resources is then

$$m_t^j = c_t^j + a^j$$

At the credit constraint we simply have that $a = 0$ and including a point $(m_{t+1}^0, c_{t+1}^0) = (0, 0)$ when interpolating the future consumption function using linear interpolation automatically handles the credit constraint.

In the full solution estimator we solve the model from the terminal period T assuming that everything is consumed. This gives $c_T^j = m_T^j$, where we construct an exogenous grid over $\vec{m}$. However, we assume that the researcher insists on using a fixed discount factor, not taking into account that the true data generating process (DGP) contains time-varying preferences. Particularly, the full solution is

$$\begin{aligned}
c_{t|\delta=0}^j &= \tilde{\beta}(1+r)\sum_{k=1}^{Q}\sum_{l=1}^{Q}\omega^k\omega^l(\mathring{c}_{t+1|\delta=0}(m_{t+1}^{jkl})G_{t+1}\eta_{t+1}^k)^{-\rho},\ \forall t = 1, \ldots, \overline{T}-1 \\
c_{\overline{T}|\delta=0}^j &= m^j
\end{aligned}$$

In the robust estimator we solve the model backwards from the pseudy-terminal period $\check{T}$ where we have imposed the approximate solution $c_{\check{T}}^j = \exp(\phi_0)(m_{\check{T}})^{\exp(\phi_1)}$ where we use the same exogenous grid over $\vec{m}$. In turn, the solution is

$$\begin{aligned}
\check{c}_t^j &= \beta_t(1+r)\sum_{k=1}^{Q}\sum_{l=1}^{Q}\omega^k\omega^l(\mathring{\check{c}}_{t+1}(m_{t+1}^{jkl})G_{t+1}\eta_{t+1}^k)^{-\rho},\ \forall t = 1, \ldots, \check{T}-1 \\
\check{c}_{\check{T}}^j &= \exp(\gamma_0)(m^j)^{\exp(\gamma_1)}
\end{aligned}$$

This functional form was chosen to impose that $c(0) = 0$ and that consumption is an increasing and concave function of normalized resources.

A.2 Estimation Details

We assume that consumption is contaminated with additive mean zero measurement error

$$C_{i,t} = C_t^\star(M_{i,t}, P_{i,t}|\theta) + \xi_{i,t},\ \mathbb{E}[\xi|M,P] = 0$$

such that we can estimate model parameters using non-linear least squares

$$\begin{aligned}
\hat{\rho} &= \arg\min_{\rho} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (C_{i,t} - C_t^\star(M_{i,t}, P_{i,t}|\rho, \delta = 0))^2 \\
(\check{\rho}, \check{\gamma}) &= \arg\hat{o}\min_{\rho,\gamma} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (C_{i,t} - \check{C}_t^\star(M_{i,t}, P_{i,t}|\rho, \gamma))^2
\end{aligned}$$

where $C_t^\star(M,P)$ denotes the model-implied optimal consumption. To get this, we simply have to multiply the found solution with permanent income, $C_t^\star(M,P) = c_t^\star(m)P$. We assume for simplicity that we observe all state variables, including permanent income. When simulating data, we draw measurement error from a normal distribution with mean zero and variance $\sigma_\xi^2 = 0.1$.

Table 1 reports the mean and standard deviation of the estimated nuisance parameters $\gamma = (\gamma_0, \gamma_1)$ across the Monte Carlo runs.

Table 1: MC results, Nuisance Parameters. Buffer-Stock model.

	$\delta = 0.000$		$\delta = 0.001$		$\delta = 0.005$		$\delta = 0.010$	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
γ_0	-0.030	0.011	-0.029	0.011	-0.027	0.012	-0.026	0.014
γ_1	-2.019	0.163	-2.003	0.167	-1.939	0.182	-1.859	0.199

Notes: The table reports the mean and variance of γ for $\delta \in \{0, 0.001, 0.005, 0.01\}$ across $S = 200$ replications. The remaining model parameters are $\rho = 2$, $\tilde{\beta} = 0.98$, $R = 1.03$, $\sigma_\varepsilon^2 = \sigma_\eta^2 = 0.01$, $G_t = 1.02$, $\sigma_\xi^2 = 0.1$.

B Example 2: Discrete Labor Market Model

Here, we present details about the dynamic discrete choice model of education and labor supply. As mentioned in the paper, the Bellman equation of the model is:

$$\begin{aligned}
V_t(\mathcal{S}_t) &= \max_k U(k, \mathcal{S}_t) + \beta \mathbb{E}\left[V_{t+1}(\mathcal{S}_{t+1}) | \mathcal{S}_t, d_{kt} = 1\right] \\
&\text{s.t.} \\
X_{t+1} &= X_t + d_{1t} \\
S_{t+1} &= S_t + d_{2t} \\
\varepsilon_{kt} &\sim \mathcal{N}(0, \sigma_k^2), \ \forall k = 1, 2, 3
\end{aligned}$$

We first present how the solution of the dynamic programming was obtained. The model is estimated through maximum likelihood estimation, we thus provide elements for the derivation of the likelihood.

B.1 Dynamic programming solution

To describe the solution of the program, let $V_{kt}(\mathcal{S}_t)$, the alternative-specific value of decision k:

$$V_{kt}(\mathcal{S}_t) = U(k, \mathcal{S}_t) + \beta \mathbb{E}\left[V_{t+1}(\mathcal{S}_{t+1}) | \mathcal{S}_t, d_{kt} = 1\right]$$

Given a set of parameters, and knowing the next period value $V_{t+1}(\mathcal{S}_{t+1})$, we need to solve $\mathbb{E}\left[\max_k V_{k,t}(\mathcal{S}_t)\right]$ .To do so, several techniques can be used. Keane and Wolpin (1994) propose to get the expected value from simulations, but the computational cost of such estimate for a three dimensional integral may be substantial. Given the simple form of our model, we propose to take advantage of the knowledge of the cdf of each of the alternative specific value functions, and of the independence between random terms. Indeed, we can show that $M(\mathcal{S}_t) = \max_k V_{k,t}(\mathcal{S}_t)$ is a positive random variable which cdf is equal to the product of the cdfs of $V_{k,t}(\mathcal{S}_t)$, and for any positive random variable it is a well known result that:

$$\mathbb{E}(M(\mathcal{S}_t)) = \int_0^{+\infty} (1 - P(M(\mathcal{S}_t) < m)) dm$$

So for any state S_t, we can obtain the expected value $V_{kt}(\mathcal{S}_t)$ at the cost of this unique integral, which we compute by adaptive quadrature.[1]

[1] To realize this, note first that since the labor income is always positive, the maximum over this and the remaining alternatives must also be positive. Further, note that since $(1 - P(M(\mathcal{S}_t) < m)) = P(M(\mathcal{S}_t) \geq m) =$

B.2 Likelihood function

Data consists in observed decisions (d_{1t}, d_{2t}, d_{3t}) and wages $w_t$if individuals are working. The derivation of the likelihood is obtained from the normality assumption of the ε_t random variables. At a given period given the individual state $\mathcal{S}_t$, we derive the likelihood contribution of each of the three decisions:

1. In the case of the labor market participation, the econometrician observes both the decision and the wage w_t. So the contribution is:

$$\ell_1(w_t|\mathcal{S}_t) = \frac{1}{\sigma_1}\phi\left(\frac{\log w_t - f_1(\mathcal{S}_t)}{\sigma_1}\right) P_1\left(d_{1t} = 1|w_t, \mathcal{S}_t\right)$$

   the probability of choosing to work conditional on the wage residual $\varepsilon_{1t} = \log w_t - f_1(\mathcal{S}_t)$ is

$$\begin{aligned} P_1(d_{1t} = 1|\varepsilon_{1t}, \mathcal{S}_t) &= P(V_{1t}(\mathcal{S}_t) > V_{2t}(\mathcal{S}_t), V_{1t}(\mathcal{S}_t) > V_{3t}(\mathcal{S}_t)|\varepsilon_1) \\ &= P(V_{1t}(\mathcal{S}_t) > V_{2t}(\mathcal{S}_t)|\varepsilon_1)P(V_{1t}(\mathcal{S}_t) > V_{3t}(\mathcal{S}_t)|\varepsilon_1) \end{aligned}$$

   where the second step obtained from the independence between unobserved heterogeneity terms. ε_2 and ε_3 appear separately in the two conditional probabilities, so these probabilities can be obtained from the functional form of the utility function and the normal distribution of ε_{2t} and ε_{3t}.

2. If the individual chooses to go to school, none of the three unobserved heterogeneity terms are observed. As in the previous case, one can use the independence conditional on ε_2 to

---

$\int_m^{+\infty} \frac{\partial}{\partial M(\mathcal{S}_t)} P(M(\mathcal{S}_t) \geq k)dk$

$$\begin{aligned} \int_0^{+\infty}(1 - P(M(\mathcal{S}_t) < m))dm &= \int_0^{+\infty}\int_m^{+\infty}\frac{\partial P(M(\mathcal{S}_t) \geq k)}{\partial M(\mathcal{S}_t)}dkdm \\ &= \int_0^{+\infty}\int_0^{k}\frac{\partial P(M(\mathcal{S}_t) \geq m)}{\partial M(\mathcal{S}_t)}dmdk \\ &= \int_0^{+\infty}\left[\frac{\partial P(M(\mathcal{S}_t) \geq m)}{\partial M(\mathcal{S}_t)}\right]_0^k dk \\ &= \int_0^{+\infty} k\frac{\partial P(M(\mathcal{S}_t) \geq k)}{\partial M(\mathcal{S}_t)}dk \\ &= \mathbb{E}(M(\mathcal{S}_t)) \end{aligned}$$

where the second equality sign follows from changing the order of integration. It is also evident from the fourth equality that this only holds if $M(\mathcal{S}_t)$ is non-negative.

derive a simple expression for the likelihood:

$$\begin{aligned}P_2(d_{2t}=1|\mathcal{S}_t) &= P(V_{2t}(\mathcal{S}_t) > V_{1t}(\mathcal{S}_t), V_{2t}(\mathcal{S}_t) > V_{3t}(\mathcal{S}_t)) \\ &= \int P(V_{2t}(\mathcal{S}_t) > V_{1t}(\mathcal{S}_t), V_{2t}(\mathcal{S}_t) > V_{3t}(\mathcal{S}_t)|\varepsilon_{2t})dF(\varepsilon_{2t}) \\ &= \int P(V_{2t}(\mathcal{S}_t) > V_{1t}(\mathcal{S}_t)|\varepsilon_{2t})P(V_{2t}(\mathcal{S}_t) > V_{3t}(\mathcal{S}_t)|\varepsilon_{2t})dF(\varepsilon_{2t})\end{aligned}$$

Both terms in the integral can be obtained from functional forms. It is still necessary to compute the integral. In practice, we evaluate this integral using 20 Gauss-Hermite quadrature nodes.

3. The probability to choose to stay out of labor force $P_2(d_{2t}=1|\mathcal{S}_t)$ is derived analogously to the schooling case and does not need further details.

The complete log-likelihood is then

$$\mathcal{L} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} d_{1,i,t}\log \ell_1(w_{i,t}\mathcal{S}_{it}) + d_{2,i,t}\log P_2(d_{2,i,t}=1|\mathcal{S}_{i,t}) + d_{3,i,t}\log P_3(d_{3,i,t}=1,\mathcal{S}_{i,t})$$

B.3 Pseudo-Terminal Value Function

In the case of the robust estimate, the backward induction start from period $\check{T}$, instead of the terminal period $\bar{T}$. The pseudo-terminal value is set as a function of the additional set of parameters ϕ, which are estimated jointly with the rest of the parameters of the model. Table 2 reports the mean and standard deviation of the estimated nuisance parameters $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)$ across the Monte Carlo runs.

Table 2: MC results. Labor Supply Model.

	$\delta=0$		$\delta=0.01$		$\delta=0.05$		$\delta=0.1$	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
γ_0	11.4076	0.0020	11.4030	0.0021	11.3849	0.0021	11.3629	0.0024
γ_1	0.8424	0.0011	0.8421	0.0012	0.8405	0.0011	0.8379	0.0013
γ_2	0.2972	0.0007	0.2963	0.0007	0.2921	0.0007	0.2862	0.0008
γ_3	0.0010	0.0004	0.0013	0.0004	0.0025	0.0004	0.0042	0.0005
γ_4	0.0029	0.0004	0.0031	0.0004	0.0039	0.0004	0.0050	0.0005
γ_5	0.0106	0.0067	0.0106	0.0068	0.0095	0.0078	0.0074	0.0111

Notes: The table reports the average absolute bias (BIAS) and the root mean square error (RMSE) under different values of the true model, $\delta \in \{0, 0.1, 0.25, 0.5\}$ across $S=200$ replications. The remaining model parameters are $\alpha_0 = 8.9$, $\alpha_2 = 0.031$, $\alpha_3 = 5000$, $\alpha_4 = -8000$, $\alpha_5 = -8000$, $\alpha_6 = 20000$, $\sigma_1^2 = 0.4$, $\sigma_2^2 = 1600$, and $\sigma_3^2 = 2000$.

† The estimate refers to the full solution approach where $\delta = 0.0$ has been imposed under estimation, irrespective of the true value of this parameter.

B.4 Estimating the whole wage equation

Table 3 shows results for the model if we were estimating the whole wage equation (i.e. estimating the wage intercept as well as estimating the returns to experience).

Table 3: MC results. Labor Supply Model.

	$\delta = 0$			$\delta = 0.01$			$\delta = 0.05$			$\delta = 0.1$		
	BIAS	STD	RMSE	BIAS	STD	RMSE	BIAS	STD	RMSE	BIAS	STD	RMSE
Full Solution												
β_1	-0.002	0.143	0.143	0.065	0.157	0.169	0.431	0.231	0.489	0.958	0.294	1.002
β_2	0.000	0.006	0.006	-0.006	0.007	0.009	-0.036	0.010	0.037	-0.077	0.013	0.078
β_3	0.001	0.005	0.005	-0.003	0.006	0.006	-0.018	0.007	0.019	-0.038	0.009	0.039
Robust Solution												
β_1	0.032	0.089	0.095	0.023	0.090	0.093	0.029	0.089	0.093	0.031	0.103	0.107
β_2	-0.002	0.004	0.005	-0.002	0.004	0.005	-0.002	0.005	0.005	-0.002	0.006	0.007
β_3	-0.002	0.012	0.012	-0.002	0.011	0.011	-0.002	0.011	0.011	-0.001	0.012	0.012

Notes: See notes for Table 2.

References

Cagetti, M. (2003): "Wealth Accumulation Over the Life Cycle and Precautionary Savings," *Journal of Business and Economic Statistics*, 21(3), 339–353.

Carroll, C. D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 2, 61–135.

——— (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.

Gourinchas, P.-O. and J. A. Parker (2002): "Consumption Over the Life Cycle," *Econometrica*, 70(1), 47–89.

Keane, M. P. and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76(4), 648–672.