Preference estimation via MSM

We attempt to estimate risk and time preferences subject to taste disturbance via method of simulated moments (MSM). We structure the analysis around Ross' recovery theorem where we observe state prices (or, equivalently, the full range of state-contingent payoffs and nominal prices) and state-contingent consumption.

First, we discuss the recovery theorem and the pricing kernel. Then, we discuss estimation of time and risk preference parameters along with some asynchronous data limitations.

Recovery theorem

Ross' [2015][1] recovery theorem says that in a complete, pure exchange market setting, linear no arbitrage equilibrium state prices convey a representative investor's state probability assignments. That is, state prices convey Markovian state transition probabilities (and preferences regarding timing of consumption and risk) for a representative investor.

The key is the pricing kernel which says the (state) price, p_{ij}, per unit probability, f_{ij}, is equal to a personal discount factor, δ, times the ratio of marginal utilities for consumption in the future state, c_j, to current, c_0, where j refers to the future state and i refers to the initial state.

$$\frac{p_{ij}}{f_{ij}} = \delta \frac{U^{'}\left(c_j\right)}{U^{'}\left(c_0\right)}$$

In other words, a representative investor with wealth or endowment, W_0, solves for optimal (expected utility of) consumption subject to a budget or wealth constraint.

$$\max_{c_0, c_j \geq 0} U\left(c_0\right) + \delta \sum_{j=1}^{n} f_{ij} U\left(c_j\right)$$

$$s.t. \quad c_0 + \textstyle\sum_{j=1}^{n} p_{ij} c_j \leq W_0$$

The first order conditions for the Lagrangian representation of the above constrained optimization problem yield the pricing kernel.

$$\lambda = U^{'}\left(c_0\right)$$

$$\delta f_{ij} U^{'}\left(c_j\right) = p_{ij} U^{'}\left(c_0\right)$$

For Markovian transition probabilities assigned as $F = \frac{1}{\delta} D P D^{-1}$ where D is a diagonal matrix with elements $U^{'}\left(c_1\right), \ldots, U^{'}\left(c_n\right)$ and with $U^{'}\left(c_0\right) = U^{'}\left(c_i\right)$, then the pricing kernel for the representative investor is

$$\frac{p_{ij}}{f_{ij}} = \delta \frac{U^{'}\left(c_j\right)}{U^{'}\left(c_i\right)}$$

[1] Ross, S. 2015, "The recovery theorem," *Journal of Finance*.

Suppose preferences exhibit constant relative risk aversion

$$U(c) = \frac{c^{1-r}}{1-r}$$

where $r \to 1$ leads to $U(c) = \ln c$.[2] For constant relative risk aversion, relative marginal utility is

$$\frac{U'(c_j)}{U'(c_i)} = \left(\frac{c_j}{c_i}\right)^{-r}$$

and logarithmic relative marginal utility is

$$\frac{U'(c_j)}{U'(c_i)} = \left(\frac{c_j}{c_i}\right)^{-1}$$

-

State transition probability assignment follows from eigensystem decomposition of the dynamic system of state prices P along with the requirement the rows of F sum to one.

$$P\zeta = \delta\zeta$$

where, by the Perron-Frobenius theorem, ζ is the positive-valued eigenvector associated with the largest eigenvalue δ. The Perron-Frobenius theorem says for a nonnegative matrix the largest eigenvalue and its associated eigenvector are nonnegative. Since P is a matrix of state prices, P is a nonnegative matrix (otherwise, there exist arbitrage opportunities).

Let ι be a vector of ones and recall eigenvectors are scale-free, $P(\alpha\zeta) = \delta(\alpha\zeta)$ implies $P\zeta = \delta\zeta$. Then, we can write

$$D^{-1}\iota = \zeta$$

with ζ scaled appropriately. Notice, the pricing kernel is also scale-free as only ratios of marginal utilities enter. Collecting terms, we have

$$\begin{aligned} P\zeta &= \delta\zeta \\ PD^{-1}\iota &= \delta D^{-1}\iota \\ \frac{1}{\delta}DPD^{-1}\iota &= \iota \\ F\iota &= \iota \end{aligned}$$

which confirms that F is a proper probability assignment as the terms are nonnegative and sum to one.[3]

[2]Constant relative risk aversion is attractive as a change in wealth leads to no change in the relative composition of an individual's portfolio (fraction of wealth invested in various assets).

[3]For the a simple two-state economy we have $F = \frac{1}{\delta}DPD^{-1}$ or $\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} = \begin{bmatrix} \frac{p_{11}}{\delta} & \frac{U'(c_1)p_{12}}{\delta U'(c_2)} \\ \frac{U'(c_2)p_{21}}{\delta U'(c_1)} & \frac{p_{22}}{\delta} \end{bmatrix}$ as indicated by the pricing kernel.

Example

Since state-contingent consumption is likely observed asynchronously (at different times)[4] we allow for taste disturbance ν_t so that $U(c_t) = \frac{c_t^{1-r}}{1-r}\nu_t$ and the pricing kernel is

$$\frac{p_{ij}}{f_{ij}} = \delta \frac{U'(c_j)}{U'(c_i)} = \delta \left(\frac{c_j}{c_i}\right)^{-r} \frac{\nu_j}{\nu_i}$$

DGP

We generate a sample based on a simple two-state economy where state prices in initial state one are $p_{1j} = \begin{bmatrix} 0.619 & 0.333 \end{bmatrix}$ and in initial state are $p_{2j} = \begin{bmatrix} 0.500 & 0.471 \end{bmatrix}$. The recovery theorem provides

$$F = \frac{1}{\delta} DPD^{-1}$$

$$\begin{bmatrix} 0.645 & 0.355 \\ 0.509 & 0.491 \end{bmatrix} = \frac{1}{0.9599} \begin{bmatrix} 1.430 & 0 \\ 0 & 1.399 \end{bmatrix} \begin{bmatrix} 0.619 & 0.333 \\ 0.500 & 0.471 \end{bmatrix} \begin{bmatrix} 0.699 & 0 \\ 0 & 0.715 \end{bmatrix}$$

With observed or proxy consumption data c_t subject to taste disturbance ν_t the pricing kernel is

$$\frac{p_{ij}}{f_{ij}} - \delta \left(\frac{c_j}{c_i}\right)^{-r} = 0$$

The proposed DGP involves $\delta = 0.9599$, $r = 0.9$, $c_1 = 0.9756$, $c_2 = 1.0250$ (when $\nu_j/\nu_i = 1$), and $\nu_t \sim N(1, \sigma)$ with sample size $n = 30$. Asynchronicity is increasing in σ. In our simulations below, we vary σ as $0.01, 0.02, 0.03$.

MSM estimation

Since the analyst cannot observe ν_t simulation of the sample moments to estimate the risk and time preference parameters, r and δ, presents an alternative to GMM with proxy data. The simulated sample is

$$\frac{p_{ij}}{f_{ij}} - \delta \left(\frac{\widetilde{c_j}}{\widetilde{c_i}}\right)^{-r} = 0$$

where $\widetilde{c_t}$ are the $T = 100$ simulated draws for c_t. That is,

$$\left(\frac{\widetilde{c_j}}{\widetilde{c_i}}\right)^{-r} = \left(\frac{c_j}{c_i}\right)^{-r} \frac{\nu_j}{\nu_i}$$

[4] If prices are also asynchronous then observed state prices are not fully deterministic of state transition probabilities. In this case, probability assignment might employ maximum entropy subject to observed state prices (Jaynes [2003]). We proceed with asynchronous consumption only.

The MSM sample moment is

$$m(i,j;\delta,r) \equiv \frac{1}{n}\sum\left\{\frac{p_{ij}}{f_{ij}} - \frac{1}{T}\sum\delta\left(\frac{\widetilde{c_j}}{\widetilde{c_i}}\right)^{-r}\right\}$$

Estimation minimizes the quadratic

$$\left[\begin{array}{cc} m(1,2;\delta,r) & m(1,1;\delta,r)\end{array}\right]\left[\begin{array}{c} m(1,2;\delta,r) \\ m(1,1;\delta,r)\end{array}\right]$$

We compare estimation of δ and r from the DGP utilizing observed, proxy consumption c_t (GMM)[5] to MSM estimation for $S = 1,000$ simulations.

$\sigma = 0.01$	GMM	MSM	GMM	MSM
	δ	δ	r	r
mean	0.960	0.960	0.914	0.919
std dev	0.000003	0.000005	0.112	0.113
1% quantile	0.960	0.960	0.707	0.713
5% quantile	0.960	0.960	0.757	0.759
10% quantile	0.960	0.960	0.781	0.785
25% quantile	0.960	0.960	0.840	0.842
50% quantile	0.960	0.960	0.900	0.904
75% quantile	0.960	0.960	0.977	0.981
90% quantile	0.960	0.960	1.060	1.068
95% quantile	0.960	0.960	1.131	1.129
99% quantile	0.960	0.960	1.213	1.212

[5] GMM sample moments and the quadratic to be minimized are analogous to those for MSM except c_t replaces $\tilde{c}_t$.

$\sigma = 0.02$	GMM	MSM	GMM	MSM
	δ	δ	r	r
mean	0.960	0.960	0.990	1.009
std dev	0.0002	0.0002	0.354	0.363
1% quantile	0.960	0.960	0.593	0.601
5% quantile	0.960	0.960	0.662	0.676
10% quantile	0.960	0.960	0.697	0.712
25% quantile	0.960	0.960	0.780	0.800
50% quantile	0.960	0.960	0.927	0.940
75% quantile	0.960	0.960	1.092	1.117
90% quantile	0.960	0.960	1.308	1.351
95% quantile	0.960	0.960	1.518	1.539
99% quantile	0.960	0.960	2.097	2.156

$\sigma = 0.03$	GMM	MSM	GMM	MSM
	δ	δ	r	r
mean	0.960	0.960	1.137	1.184
std dev	0.0127	0.0128	0.689	0.713
1% quantile	0.955	0.954	0.516	0.535
5% quantile	0.960	0.960	0.581	0.605
10% quantile	0.960	0.960	0.635	0.662
25% quantile	0.960	0.960	0.747	0.780
50% quantile	0.960	0.960	0.923	0.965
75% quantile	0.960	0.960	1.248	1.293
90% quantile	0.960	0.960	1.805	1.894
95% quantile	0.960	0.960	2.443	2.568
99% quantile	0.960	0.960	4.045	4.424

Not surprisingly, estimation of δ is more precise than for r. Estimation results are slightly better for proxy consumption data (GMM) than for simulated consumption data (MSM). As expected, the performance of both declines with asynchronicity (increasing σ). Median estimates are a more effective measure of central tendency than means for both GMM and MSM. Results are likely to improve with additional moments and/or adjustment for differential precision.

Three-state economy example

Next, we undertake another experiment. First, since estimation of δ is not stochastic given our assumptions regarding the data, we focus estimation on r alone. Further, we compare estimation of r with one moment based on a two-state economy with three moments based on a three-state economy.

The sample moment based on proxy consumption data for the two-state economy is

$$m_p\left(i, j; \delta, r\right) \equiv \frac{1}{n} \sum \left\{ \frac{p_{ij}}{f_{ij}} - \delta \left(\frac{c_j}{c_i} \right)^{-r} \right\}$$

and the simulated sample moment for the two-state economy is

$$m_s\left(i,j;\delta,r\right) \equiv \frac{1}{n}\sum\left\{\frac{p_{ij}}{f_{ij}} - \frac{1}{T}\sum\delta\left(\frac{\widetilde{c_j}}{\widetilde{c_i}}\right)^{-r}\right\}$$

where $i = 1, j = 2$. The risk parameter r is estimated by setting the sample moments equal to zero. For this experiment we set $n = 30, T = 300$ and again we vary $\sigma = 0.01, 0.02, 0.03$.

The sample moments for the three-state economy are defined as above except the (i,j) pairs are $(1,2), (1,3), (2,3)$. Estimation of the risk aversion parameter minimizes the quadratic (with identity weight matrix and δ, r implied).

$$m_k m_k^T = \begin{bmatrix} m_k\left(1,2\right) & m_k\left(1,3\right) & m_k\left(2,3\right) \end{bmatrix} \begin{bmatrix} m_k\left(1,2\right) & m_k\left(1,3\right) & m_k\left(2,3\right) \end{bmatrix}^T$$

for $k = p, s$.

Results for the two-state economy are reported below.

	$\sigma = 0.01$		$\sigma = 0.02$		$\sigma = 0.03$	
	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$
mean	0.919	0.923	0.988	1.007	1.098	1.144
std dev	0.112	0.112	0.352	0.363	0.590	0.607
1% quantile	0.718	0.721	0.584	0.595	0.524	0.541
5% quantile	0.764	0.769	0.643	0.655	0.577	0.603
10% quantile	0.789	0.793	0.697	0.712	0.626	0.651
25% quantile	0.839	0.841	0.786	0.800	0.734	0.764
50% quantile	0.904	0.910	0.916	0.931	0.927	0.972
75% quantile	0.983	0.988	1.104	1.130	1.254	1.319
90% quantile	1.073	1.078	1.321	1.347	1.722	1.813
95% quantile	1.128	1.130	1.512	1.539	2.250	2.323
99% quantile	1.234	1.254	2.142	2.206	3.465	3.479

Results for the (equally-weighted moments) three-state economy are reported below. As expected, three moment conditions produce smaller bias and variation in the estimation of r than the single moment condition, two-state economy.

.

	$\sigma = 0.01$		$\sigma = 0.02$		$\sigma = 0.03$	
	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$
mean	0.898	0.903	0.898	0.916	0.900	0.936
std dev	0.041	0.036	0.089	0.076	0.140	0.123
1% quantile	0.809	0.829	0.709	0.767	0.629	0.710
5% quantile	0.833	0.847	0.762	0.806	0.695	0.766
10% quantile	0.845	0.857	0.791	0.824	0.734	0.795
25% quantile	0.870	0.877	0.838	0.851	0.810	0.851
50% quantile	0.897	0.901	0.894	0.911	0.888	0.919
75% quantile	0.924	0.927	0.952	0.965	0.978	0.999
90% quantile	0.951	0.949	1.011	1.016	1.077	1.091
95% quantile	0.965	0.965	1.050	1.044	1.147	1.162
99% quantile	1.006	0.998	1.134	1.130	1.343	1.309

Weighted moments

Experiments with weighting based on the inverse of the covariance of the "residuals" from the first stage estimation of r produce greater bias and greater variance (results reported below) than those without weighting.

$$m_k \Omega m_k^T$$

$$\Omega = \left(\frac{1}{n} e e^T \right)^{-1}$$

where the $n = 30$ "residuals" $e_{ij} = \frac{p_{ij}}{f_{ij}} - \delta \left(\frac{c_j}{c_i} \right)^{-r}$ and $e^T = \begin{bmatrix} e_{12} & e_{13} & e_{23} \end{bmatrix}$ have similar variances and are highly correlated (0.5 to 0.7 in absolute value).

The small sample size likely contributes to weak estimation of the covariance and subsequent poor results. Results for the three-state economy based on minimization of weighted quadratic moments are reported below.

	$o = 0.01$		$\sigma = 0.02$		$\sigma = 0.03$	
	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$	$r\,(GMM)$	$r\,(MSM)$
mean	0.902	0.867	0.916	0.795	0.943	0.707
std dev	0.042	0.055	0.093	0.091	0.155	0.124
1% quantile	0.814	0.752	0.736	0.610	0.687	0.486
5% quantile	0.835	0.779	0.785	0.660	0.741	0.532
10% quantile	0.849	0.799	0.804	0.682	0.770	0.565
25% quantile	0.874	0.829	0.851	0.733	0.840	0.624
50% quantile	0.901	0.885	0.907	0.789	0.916	0.697
75% quantile	0.928	0.902	0.969	0.851	1.016	0.774
90% quantile	0.957	0.936	1.036	0.915	1.136	0.859
95% quantile	0.973	0.960	1.086	0.952	1.227	0.922
99% quantile	1.014	1.006	1.195	1.053	1.448	1.104