GMM ESTIMATION OF ASSET PRICING MODELS

Patrick GAGLIARDINI
University of Lugano and Swiss Finance Institute

Outline

1. Rational expectations and no-arbitrage pricing models
2. Empirical analysis with GMM
3. Weak identification
4. Inference robust to weak identification
5. GMM with optimal instruments
6. Information-theoretic GMM
7. Lack of identification in asset pricing models
8. XMM and efficient derivative pricing

Rational expectations (Lucas (1978))

An agent maximizes expected life-time utility from consumption:

$$\max_{\{C_t\},\{\alpha_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t U(C_t;\gamma) \right] \qquad \text{s.t. } C_t + p_t\alpha_t = p_t\alpha_{t-1}$$

$U(C_t;\gamma) = [C_t^{1-\gamma} - 1]/(1-\gamma)$ (say) is utility from consumption C_t
α_t is the number of shares in $[t, t+1)$ of an asset with price p_t

First-order condition for maximization (Euler equation):

$$\begin{aligned}
&-p_t U'(C_t;\gamma) + \beta E_t\left[U'(C_{t+1};\gamma)p_{t+1}\right] = 0 \\
&\Leftrightarrow p_t = E_t\left[\beta \frac{U'(C_{t+1};\gamma)}{U'(C_t;\gamma)} p_{t+1}\right] \\
&\Leftrightarrow E_t\left[\beta \frac{U'(C_{t+1};\gamma)}{U'(C_t;\gamma)} R_{t+1} - 1\right] = 0, \qquad R_{t+1} = p_{t+1}/p_t
\end{aligned}$$

$E_t[.]$ is conditional expectation given the information at time t

No-arbitrage pricing

Stochastic discount factor (sdf): the no-arbitrage condition implies the existence of a random variable $M_{t,t+1} > 0$ such that

$$p_t = E_t\left[M_{t,t+1}p_{t+1}\right]$$

[Harrison-Kreps (1979), Harrison-Pliska (1981), Hansen-Richard (1987)]

Econometric sdf specification:

$$M_{t,t+1} = m(Y_{t+1}; \theta_0)$$

where Y_{t+1} are relevant state variables and θ_0 is a vector of unknown risk premia parameters yielding

$$E_t\left[m(Y_{t+1}; \theta_0)R_{t+1} - 1\right] = 0$$

Examples of sdf specifications

Time-separable preferences, CRRA utility $U(C_t;\gamma) = \dfrac{C_t^{1-\gamma} - 1}{1-\gamma}$

$$M_{t,t+1}(\theta) = \beta \left(C_{t+1}/C_t \right)^{-\gamma}, \qquad \theta = (\beta, \gamma)$$

Time-nonseparable Epstein-Zin (1989, 1991) preferences

$$M_{t,t+1}(\theta) = \beta^{\lambda} \left(C_{t+1}/C_t \right)^{-\gamma\lambda} R_{0,t+1}^{\lambda-1}, \qquad \theta = (\beta, \gamma, \lambda)$$

where $R_{0,t+1}$ is the gross return of the optimal portfolio

Parameterization such that the risk-aversion is $1 - \lambda(1-\gamma)$ and the elasticity of intertemporal substitution (EIS) is $\psi = 1/\gamma$

Reduced-form sdf for derivative pricing:

$$M_{t,t+1}(\theta) = e^{-r_{f,t+1}} \exp\left(-\theta_1 - \theta_2 \sigma_{t+1}^2 - \theta_3 \sigma_t^2 - \theta_4 r_{t+1} \right)$$

where $r_{f,t}$ is the risk-free rate and σ_t^2 is the stochastic volatility of the underlying asset with logarithmic return r_t

Conditional moment restrictions

Let the information at date t be contained in the stochastic vector W_t following a Markov process, e.g. $W_t = (Y_t, Y_{t-1}, \cdots, Y_{t-q})$

Euler conditions/no-arbitrage conditions yield restrictions on data and parameters in the form of a ...

Conditional Moment Restriction (CMR):

$$E\left[h(Y_{t+1};\theta_0)|W_t\right] = 0$$

where

$$h(Y_{t+1};\theta) = m(Y_{t+1};\theta)R_{t+1} - 1$$

From conditional to unconditional MR

Let $Z_t = \varphi(W_t)$ be an instrument

By the iterated expectation theorem:

$$\begin{aligned} E\left[Z_t h(Y_{t+1};\theta_0)\right] &= E\left[E\left[Z_t h(Y_{t+1};\theta_0)|W_t\right]\right] \\ &= E\left[Z_t E\left[h(Y_{t+1};\theta_0)|W_t\right]\right] \\ &= 0 \end{aligned}$$

which yields an ...

Unconditional Moment Restriction (UMR):

$$E\left[g(X_t;\theta_0)\right] = 0$$

where $g(X_t;\theta_0) = Z_t h(Y_{t+1}\theta_0)$

GMM inference in asset pricing models (Hansen and Singleton (1982))

The two-step GMM estimator:

$$\hat{\theta}_T = \arg\min_{\theta \in \Theta} \hat{g}_T(\theta)' \hat{V}_T^{-1} \hat{g}_T(\theta)$$

where

$$\hat{g}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(X_t;\theta), \quad g(X_t;\theta) = Z_t \otimes (M_{t,t+1}(\theta)R_{t+1} - \iota)$$

θ	is $p \times 1$ vector of unknown parameters
R_{t+1}	is $G \times 1$ vector of asset gross returns
Z_t	is $K \times 1$ vector of instruments
ι	is $G \times 1$ vector of ones

and $\hat{V}_T$ is consistent estimator of $V_0 = \lim_{T\to\infty} V\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(X_t;\theta_0)\right]$

GMM inference in asset pricing models (Hansen and Singleton (1982))

Since $E\left[M_{t,t+1}(\theta_0)R_{t+1} - \iota | W_t\right] = 0$, the autocovariances vanish

$$\Gamma(j) = E\left[g(X_t, \theta_0)g(X_{t+j}, \theta_0)\right] = 0, \quad j \geq 1$$

and

$$V_0 = E\left[g(X_t, \theta_0)g(X_t, \theta_0)'\right], \qquad \hat{V}_T = \frac{1}{T}\sum_{t=1}^{T} g(X_t, \tilde{\theta}_T)g(X_t, \tilde{\theta}_T)'$$

The Hansen statistic

$$\xi_T^H = T\hat{g}_T(\hat{\theta}_T)'\hat{V}_T^{-1}\hat{g}_T(\hat{\theta}_T) \sim \chi^2_{GK-p}$$

can be used to test the asset pricing model

Empirical results: Hansen and Singleton (1982)

Monthly data from 1959:2 to 1977:12

Returns of equally-weighted and value-weighted portfolios of NYSE stocks

Instruments include current and lagged values of asset return and consumption growth

Time-separable preferences, CRRA utility $U(C;\gamma) = \dfrac{C^{1-\gamma}-1}{1-\gamma}$

Estimates of risk-aversion γ range between 0.5 and 1 with standard errors of about 0.20

Hansen overidentification test rejects the model

Empirical results: Stock and Wright (2000)

Monthly data on stock and bond portfolios from 1959:1 to 1990:12

Instruments include additionally current and lagged values of bond term spread and dividend yield

CRRA utility: estimated risk-aversion between 0 and 1

Epstein-Zin preferences: estimated risk-aversion is often negative!

Empirical results: Yogo (2004)

Yogo (2004) estimates the EIS ψ by GMM from the regression

$$\Delta c_{t+1} = \alpha + \psi r_{f,t+1} + \varepsilon_{t+1}$$

where $\Delta c_{t+1} = \log(C_{t+1}/C_t)$ is log consumption growth

For time-separable preferences and CRRA utility $\psi = 1/\gamma$

The CMR $E\left[\Delta c_{t+1} - \alpha - \psi r_{f,t+1} | W_t\right] = 0$ corresponds to the linearized Euler condition with CRRA utility written for the riskfree asset and divided by γ

GMM estimates of EIS ψ are in general small (and sometimes negative!), in accordance with Hall (1988)

Results suggest that risk-aversion $\gamma = \dfrac{1}{\psi}$ is (much) larger than 1

Empirical results: Hall (2005)

Monthly stock returns data as in Hansen and Singleton (1982) but on the period 1959:1-1997:12

Instruments include current and past values of asset return and consumption growth only

Time-separable preferences with CRRA utility

Estimated risk-aversion ranges between 0.7 and 1.3

Confidence intervals for γ are large, e.g. $(-4, 4)$

Summary on empirical analysis

Advantage of GMM:

allows to estimate general nonlinear rational expectations and no-arbitrage asset pricing models “when only a subset of the economic environment is explicitly specified a-priori”

Drawbacks of GMM:

Instability of estimates when changing basic asset returns and/or instruments and/or parameterization

Large confidence intervals for some preference parameters such as risk-aversion coefficient

Confidence intervals and hypothesis tests based on standard asymptotic approximations are often unreliable in finite sample [see e.g. Hansen, Heaton, Yaron (1996)]

Weak identification

Drawbacks of GMM are likely related to weak identification

A parameter θ is weakly identified when the UMR is not very informative to estimate the true value θ_0

Figures 3.1 and 3.2 in Hall (2005), pp. 62 and 64, show that the GMM criterion is very flat over a wide range of values of the risk-aversion parameter!

Locally, weak identification means that the Jacobian matrix

$$J_0 = E_0 \left[\frac{\partial g(X_t; \theta_0)}{\partial \theta'} \right] = E_0 \left[Z_t \frac{\partial M_{t,t+1}(\theta_0)}{\partial \theta'} \right]$$

is near reduced-rank, i.e., the instrument Z_t is only weakly correlated with (some function of) the future state variables

Intuition with CRRA utility

For CRRA utility the sdf is

$$M_{t,t+1}(\theta) = \beta\,(C_{t+1}/C_t)^{-\gamma} = \beta\exp(-\gamma\Delta c_{t+1}), \qquad \theta = (\beta,\gamma)'$$

where $\Delta c_{t+1} = \log(C_{t+1}/C_t)$ is log consumption growth

By linearization around $\Delta c_{t+1} \simeq 0$ we have:

$$\frac{\partial M_{t,t+1}(\theta)}{\partial\theta'} = \exp(-\gamma\Delta c_{t+1})\,(1, -\beta\Delta c_{t+1}) \simeq (1, -\beta\Delta c_{t+1})$$

Consumption growth Δc_{t+1} is difficult to forecast with variables at time t (Hall (1988))

$\Rightarrow$ the risk-aversion parameter γ is likely weakly identified while the time-discount parameter β is strongly identified!

Weak instruments in linear IV model

Consider a linear IV regression model

$$
\begin{aligned}
y &= X\beta + u \\
X &= Z\Pi + v
\end{aligned}
$$

where y and X are $T \times 1$ vectors of endogenous variables and Z is a $T \times K$ matrix of nonstochastic instruments

Errors $(u_t, v_t)'$, $t = 1, \cdots, T$ are $IIN(0, \Sigma)$ with $\Sigma = \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix}$ and $\sigma_{uv} \neq 0$

Matrix Π measures the strength of the instruments Z

Weak instruments in linear IV model

Theoretical and Monte-Carlo insights show that the finite-sample distribution of the 2SLS estimator depends on sample size T, number of instruments K and instrument strength Π through

$$\mu^2/K$$

where μ^2 is the concentration parameter defined by

$$\mu^2 = \Pi' Z Z' \Pi / \sigma_v^2$$

Figure 1 in Stock, Wright, Yogo (2002) shows that for small values of μ^2/K (≤ 10, say), the distributions of the 2SLS estimator and t-statistics are highly nonnormal!

(see also Nelson and Startz (1990))

Weak instruments asymptotics

Usual (fixed-model) asymptotic normal approximations rely on

$$K, \Pi \text{ fixed} \quad \text{and} \quad T \to \infty, \text{ i.e. } \mu^2/K \to \infty$$

and cannot provide a good description for a setting with low μ^2/K !

Weak instruments asymptotics: a sequence of drifting models

$$K \text{ fixed}, \Pi \to 0 \text{ and } T \to \infty \text{ such that } \mu^2/K \to \text{ constant (small)}$$

to provide an approximation for a setting with low μ^2/K

An example of weak IV asymptotics

Linear IV regression:

$$\begin{aligned} y_t &= \beta + \alpha x_t + u_t, \qquad t = 1, \cdots, T \\ w_t &= \pi x_t + v_t \end{aligned}$$

where $(x_t, u_t, v_t)' \sim IIN(0, I_3)$

Orthogonality condition for parameter of interest $\theta = (\beta, \alpha)'$

$$E[g(X_t; \theta_0)] = E[z_t(y_t - \beta_0 - \alpha_0 x_t)] = 0$$

where $z_t = (1, w_t)'$ is the instrument

If $\pi = \pi_0/\sqrt{T}$ the instrument w_t is weakly correlated with regressor x_t and parameter α is weakly identified!

An example of weak IV asymptotics

Write the GMM=2SLS estimator as:

$$\hat{\alpha}_T = \left(\sum_{t=1}^{T} (w_t - \bar{w})(x_t - \bar{x}) \right)^{-1} \sum_{t=1}^{T} (w_t - \bar{w}) y_t, \qquad \hat{\beta}_T = \bar{y} - \bar{x}\hat{\alpha}_T$$

Then:

$$\hat{\alpha}_T - \alpha_0 = \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (w_t - \bar{w})(x_t - \bar{x}) \right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (w_t - \bar{w}) u_t \xrightarrow{d} \frac{Z_1}{\pi_0 + Z_2}$$

$$\sqrt{T}\left(\hat{\beta}_T - \beta_0\right) = \sqrt{T}\bar{u} - \sqrt{T}\bar{x}\left(\hat{\alpha}_T - \alpha_0\right) \xrightarrow{d} Z_3 - \frac{Z_1 Z_4}{\pi_0 + Z_2}$$

where $(Z_1, Z_2, Z_3, Z_4)' \sim N(0, I_4)$

The estimator of the weakly identified parameter α is inconsistent while the estimator of the strongly identified parameter β is root-T consistent but asymptotically nonnormal!

Weak IV asymptotics: Stock and Wright (2000)

Consider a general GMM setting with UMR $E[g(X_t;\theta_0)] = 0$

Partition $\theta = (\alpha', \beta')'$ and assume a drifting DGP such that

$$E\left[g(X_t;\theta)\right] = m_1(\alpha,\beta)/\sqrt{T} + m_2(\beta) \tag{1}$$

where $m_1(\alpha_0,\beta_0) = 0$ and $m_2(\beta) = 0 \Leftrightarrow \beta = \beta_0$

Parameter α is weakly identified since the UMR is almost uninformative for α when T is large, parameter β is strongly identified

Proposition (Stock-Wright (2000), Thm 1): *Under the weak identification assumption (1) the GMM estimator $\hat{\alpha}$ is inconsistent and the GMM estimator $\hat{\beta}$ is root-T consistent, with non-Gaussian asymptotic distribution*

Weak identification robust inference

The Continuously Updated Estimator (CUE):

$$\hat{\theta}_T^{\text{CUE}} = \arg\min_{\theta\in\Theta} Q_T^{\text{CUE}}(\theta), \qquad Q_T^{\text{CUE}}(\theta) = \hat{g}_T(\theta)'\hat{V}_T(\theta)^{-1}\hat{g}_T(\theta)$$

where $\hat{V}_T(\theta)$ is consistent estimator of $V_0(\theta) := V[g(X_t, \theta)]$

Proposition (Stock-Wright (2000), Thm 2): *Under the null hypothesis $\theta = \theta_0$ we have $T \cdot Q_T^{\text{CUE}}(\theta_0) \xrightarrow{d} \chi^2(m)$, $m = dim(g)$, whether identification is weak or strong*

By inverting the statistic $T \cdot Q_T^{\text{CUE}}$ we can construct a so-called S-confidence set for θ with asymptotic level ε

$$S_T := \left\{\theta : T \cdot Q_T^{\text{CUE}}(\theta) \leq \chi^2_{1-\varepsilon}(m)\right\}$$

which is robust to weak identification!

Application of S-sets for asset pricing

Stock and Wright (2000) apply S-sets to inference in asset pricing models with US data

In their empirical analysis conventional confidence ellipses and S-sets strongly differ for time-separable CRRA utilities, habit formation and Epstein-Zin preferences!

S-sets are typically much larger and imply a much greater degree of risk-aversion $\gamma > 20$ (see e.g. Figures 3, 4 and 6 in Stock-Wright (2000))

In a few cases, the S-sets are null i.e. there are no parameter values consistent with the moment restrictions

Yogo (2004) applies S-sets for inference on EIS and concludes that ψ is small, e.g. 95% CI for Switzerland in period 1976:3-1998:4 is $[-1.42, 0.5]$ (see Table 6)

Summary on weak identification

Empirical conclusions based on weak identification robust methodologies substantially differ from those obtained with conventional approaches

Preference-based sdf specifications are less often rejected but confidence sets are much larger and consistent with very high levels of risk-aversion

Informal checks for symptoms of weak identification:

- Criterion functions admitting large plateaus
- Substantial difference between S-sets and conventional GMM confidence sets
- Substantially different estimates obtained from estimators that are asymptotically equivalent under conventional theory
- Monte-Carlo on a calibrated model showing large biases in point estimates and size distortions in test statistics

Optimal instruments

Consider the CMR

$$E\left[h(X_t,\theta)|W_t\right] = 0,\ P\text{-a.s.}$$

to estimate parameter θ

Assume that the true value θ_0 is identified by the CMR, i.e.

Global Identification (GI) assumption:

$$E\left[h(X_t,\theta)|W_t\right] = 0 \ P\text{-a.s.},\ \theta \in \Theta \Leftrightarrow \theta = \theta_0$$

Local Identification (LI) assumption:

$$E\left[\frac{\partial h(X_t,\theta_0)}{\partial \theta'}|W_t\right] \text{ is full rank } P\text{-a.s.}$$

Optimal instruments

Let $Z_t = \varphi(W_t)$ be an admissible instrument and let $\Sigma(Z)$ denote the asymptotic variance-covariance matrix of the best GMM estimator with instrument Z_t

Proposition (Chamberlain (1987)): *There exists an optimal instrument that minimizes* $\Sigma(Z)$. *It is given by*

$$Z_t^* = E\left[\frac{\partial h(X_t,\theta_0)'}{\partial \theta}|W_t\right] V\left[h(X_t,\theta_0)|W_t\right]^{-1}$$

The associated GMM estimator achieves asymptotically the semi-parametric efficiency bound:

$$\Sigma(Z^*) = E\left[E\left[\frac{\partial h(X_t,\theta_0)'}{\partial \theta}|W_t\right] V\left[h(X_t,\theta_0)|W_t\right]^{-1} E\left[\frac{\partial h(X_t,\theta_0)}{\partial \theta'}|W_t\right]\right]^{-1}$$

How to implement GMM with optimal instrument?

Information-theoretic GMM

The optimal instrument Z_t^* involves the true conditional density $f_0(x|w)$ of X_t given W_t by means of the conditional expectation and variance of the moment function

The kernel estimator $\hat{f}(x|w)$ does not take into account the information in the CMR

Basic idea of information-theoretic GMM: estimate jointly θ_0 and $f_0(x|w)$ by looking for the pdf $f(x|w)$ of X_t given W_t which is

the closest to the kernel conditional density estimator $\hat{f}(x|w)$

subject to

the conditional moment restrictions

Information-theoretic GMM

Let $Dist(f|g)$ denote a distance between densities $f > 0$ and $g > 0$ over the support of X, e.g. chi-square distance

$$Dist(f|g) = \int \frac{[f(x) - g(x)]^2}{g(x)} dx$$

or Kullback-Leibler (KL) distance

$$Dist(f|g) = \int \log\left[f(x)/g(x)\right] f(x) dx$$

Definition: *The estimators* $\hat{\theta}_T$ *and* $\hat{f}_t, t = 1, \cdots, T$ *minimize*

$$Q_T(f_1, \cdots, f_T) = \frac{1}{T} \sum_{t=1}^{T} Dist\left(f_t, \hat{f}(\cdot|w_t)\right)$$

subject to

$$\int f_t(x)dx = 1, \quad \int h(x;\theta) f_t(x) dx = 0, \quad t = 1, \cdots, T$$

Euclidean likelihood: Chi-square distance

The Lagrangian is

$$\mathcal{L} = \frac{1}{T}\int \frac{[f_t(x) - \hat{f}(x|w_t)]^2}{\hat{f}(x|w_t)}dx - \sum_{t=1}^{T}\mu_t \int f_t(x)dx - \sum_{t=1}^{T}\lambda_t' \int h(x;\theta)f_t(x)dx$$

The optimization w.r.t. functions f_t, $t = 1, \cdots, T$, for given θ can be performed analytically:

$$\hat{f}_t(x;\theta) = \hat{f}(x|w_t)\left\{1 - \widehat{E}[h(\theta)|w_t]'\widehat{V}[h(\theta)|w_t]^{-1}\left(h(x_t,\theta) - \widehat{E}[h(\theta)|w_t]\right)\right\}$$

where $\widehat{E}[h(\theta)|w_t]$ and $\widehat{V}[h(\theta)|w_t]$ denote conditional expectation and variance of $h(x_t,\theta)$ w.r.t. the kernel density $\hat{f}(.|w_t)$

Euclidean likelihood: Chi-square distance

The estimator of θ is computed by minimizing the concentrated criterion

$$\hat{\theta}_T^{\text{CUE}} = \arg\min_{\theta} \frac{1}{T} \sum_{t=1}^{T} \widehat{E}[h(\theta)|w_t]' \widehat{V}[h(\theta)|w_t]^{-1} \widehat{E}[h(\theta)|w_t]$$

and corresponds to a Continuosly Updated estimator

Proposition (Antoine, Bonnal, Renault (2007)): *The estimator $\hat{\theta}_T^{CUE}$ is consistent, asymptotically normal and reaches the semi-parametric efficiency bound* $\Sigma(Z^*)$

The information-theoretic approach to GMM automatically selects the optimal instruments and weighting matrix!

Exponential tilting: KL distance

The estimates of f_t, $t = 1, \cdots, T$, for given θ are

$$\hat{f}_t(x;\theta) = \hat{f}(x|w_t) \exp\left(\lambda_t(\theta)' h(x_t,\theta)\right) / \hat{E}\left[\exp\left(\lambda_t(\theta)' h(x_t,\theta)\right) | w_t\right]$$

where the Lagrange multiplier vector $\lambda_t(\theta)$ is such that the tilted density $\hat{f}_t(x;\theta)$ satisfies $\int h(x,\theta)\hat{f}_t(x,\theta)dx = 0$

The exponential tilting (ET) estimator is

$$\hat{\theta}_T^{\text{ET}} = \arg\min_{\theta} -\frac{1}{T}\sum_{t=1}^{T} \log \hat{E}\left[\exp\left(\lambda_t(\theta)' h(x_t,\theta)\right) | w_t\right]$$

Proposition (Kitamura, Tripathi, Ahn (2004)): *The ET estimator is asymptotically equivalent to CUE, in particular semi-parametrically efficient*

Computation of ET estimator is more cumbersome than CUE but ensures positive estimated densities!

Lack of identification in asset pricing models

No-arbitrage restrictions from a set of fundamental assets can be insufficient to identify the sdf parameter θ, i.e. GI and LI fail

In such a case any (information-theoretic) GMM estimator of θ is inconsistent!

Gagliardini, Gouriéroux and Renault (2005) provide an example in a derivative pricing framework

Under the DGP P_0 the underlying asset return is such that

$$r_t = r_{f,t} + \gamma_0 \sigma_t^2 + \sigma_t \varepsilon_t$$

where $(\varepsilon_t) \sim IIN(0,1)$, the risk-free rate $r_{f,t}$ is deterministic and the stochastic volatility (σ_t^2) follows a discrete-time Heston (1993) model

Lack of identification

The true sdf is exponential affine:

$$M_{t,t+1}(\theta_0) = e^{-r_{f,t+1}} \exp\left(-\theta_1^0 - \theta_2^0 \sigma_{t+1}^2 - \theta_3^0 \sigma_t^2 - \theta_4^0 r_{t+1}\right)$$

where $\theta_0 = (\theta_1^0, \theta_2^0, \theta_3^0, \theta_4^0)'$

The no-arbitrage restrictions for the risk-free asset and the underlying asset are

$$E_0\left[M_{t,t+1}(\theta_0) e^{r_{f,t+1}} | X_t\right] = 1,\ E_0\left[M_{t,t+1}(\theta_0) e^{r_{t+1}} | X_t\right] = 1,\ X_t = (r_t, \sigma_t^2)$$

Proposition (Gagliardini, Gouriéroux, Renault (2005): *There exists a set of parameter vectors $\theta = \theta(\theta_2)$ indexed by $\theta_2 \in \mathbb{R}$ s.t.*

$$E_0\left[M_{t,t+1}(\theta) e^{r_{f,t+1}} | X_t\right] = 1,\ E_0\left[M_{t,t+1}(\theta) e^{r_{t+1}} | X_t\right] = 1,\ P\text{-a.s.}$$

for any $\theta_2 \in \mathbb{R}$

Identification of θ_0 requires the use of derivative prices!

Extended Method of Moments (XMM): Gagliardini, Gouriéroux, Renault (2005)

XMM is an extension of GMM to accommodate a more general set of conditional moment restrictions:

- uniform CMR, i.e. valid for any value of the conditioning variable (the usual CMR!)
- local CMR, i.e. valid for a given value of the conditioning variable only

Local CMR correspond to no-arbitrage restrictions for cross-sectionally observed prices of actively traded derivatives

Characteristics of actively traded derivatives change from one trading day to the other

Local restrictions from derivative assets provide identification and efficiency in estimation of risk premia and option prices

References

Antoine, B., Bonnal, H. and E. Renault (2007): “On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood”, Journal of Econometrics, 138, 461-487

Chamberlain, G. (1987): ”Asymptotic Efficiency in Estimation with Conditional Moment Restrictions”', Journal of Econometrics, 34, 305-334

Epstein, L. and S. Zin (1989): “Substitution, Risk-Aversion, and the Temporal Behaviour of Consumption and Asset Returns: A Theoretical Framework”, Econometrica, 57, 937-969

Epstein, L. and S. Zin (1991): “Substitution, Risk-Aversion, and the Temporal Behaviour of Consumption and Asset Returns: An Empirical Analysis”, Journal of Political Economy, 99, 263-286

References

Gagliardini, P., Gouriéroux, C. and E. Renault (2005): "Efficient Derivative Pricing by the Extended Method of Moments", Working Paper.

Hall, A. (2005): *Generalized Method of Moments*, Oxford University Press

Hall, R. (1988): "Intertemporal Substitution in Consumption", Journal of Political Economy, 96, 339-357

Hansen, L., Heaton, J. and A. Yaron (1996): "Finite Sample Properties of Some Alternative GMM Estimators", Journal of Business and Economic Statistics, 14, 262-280

Hansen, L. and K. Singleton (1982): "Generalized Instrumental Variable Estimation of Nonlinear Rational Expectations Models", Econometrica, 50, 1269-1286

References

Hansen, L. and S. Richard (1987): ”The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models”, Econometrica, 55, 587-613

Harrison, J. and D. Kreps (1979): ”Martingales and Arbitrage in Multiperiod Securities Markets”, Journal of Economic Theory, 20, 381-408

Harrison, M. and S. Pliska (1981): ”Martingales and Stochastic Integrals in the Theory of Continuous Trading”', Stochastic Processes and Their Applications, 11, 215-260

Heston, S. (1993): ”A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”, Review of Financial Studies, 6, 327-343

References

Kitamura, Y. (2007): ”Empirical Likelihood Methods in Econometrics: Theory and Practice” in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. 3, eds. Blundell, R., Newey, W., and T., Persson, Cambridge University Press, 174-237

Kitamura, Y. and M. Stutzer (1997): “An Information Theoretic Alternative to Generalized Method of Moments Estimation”, Econometrica, 65, 861-874

Kitamura, Y., Tripathi, G. and H. Ahn (2004): “Empirical Likelihood Based Inference in Conditional Moment Restriction Models”, Econometrica, 72, 1667-1714

Lucas, R. (1978): ”Asset Prices in an Exchange Economy”, Econometrica, 46, 1429-1446

References

Nelson, C. and R. Startz (1990): “Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator”, Econometrica, 58, 967-976

Stock, J. and J. Wright (2000): “GMM with Weak Identification”, Econometrica, 68, 1055-1096

Stock, J., Wright, J. and M. Yogo (2002): “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments”, Journal of Business and Economic Statistics, 20, 518-529

Yogo, M. (2004): “Estimating the Elasticity of Intertemporal Substitution When Instruments are Weak”, Review of Economics and Statistics, 86, 797-810