Generalized Autoregressive Method of Moments

Marcin Zamojski[1] with: Drew Creal[2], Siem Jan Koopman[1], and André Lucas[1]

[2]University of Chicago, Booth School of Business

[1]VU University Amsterdam and Tinbergen Institute

What is the Generalized Autoregressive Method of Moments (GaMM)?

- GaMM extends Generalized Method of Moments (GMM) to a setting where a **subset of the parameters are expected to vary over time** with unknown dynamics.
- You only need to specify a set of conditional moment conditions and a set of parameters which are assumed to vary over time.
- **We approximate the dynamics by an autoregressive process driven by the score** of the local GMM criterion function to filter out the dynamic path of the time-varying parameter.
- Our approach is **completely observation driven**, such that estimation and inference are entirely straightforward.
- GaMM is flexible, easily implementable, and allows for any parameter dynamics, e.g., structural breaks, sinusoidal waves, AR(p), etc.
- GaMM generalizes GAS (Creal et al., 2013) to settings where densities are unknown or hard to derive.

Motivating example

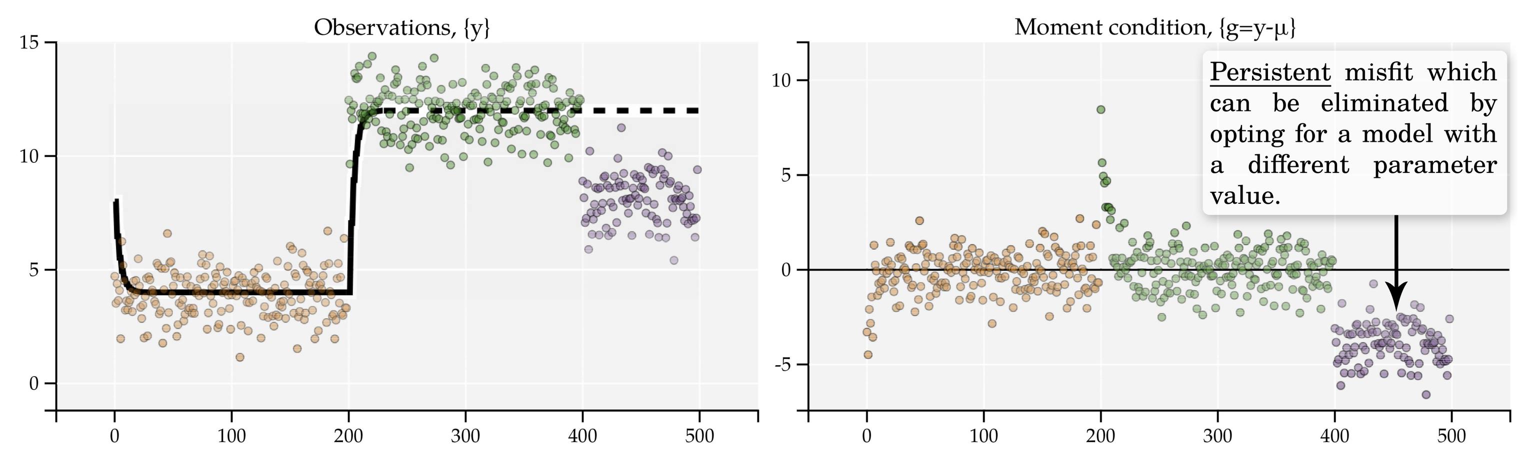
- Consider the problem of estimating the mean $\boldsymbol{\mu}$ of a random variable y_t.
- Given moment condition $\mathrm{E}[y_t - \boldsymbol{\mu}] = 0$, the standard GMM objective function would be $(\sum_{t=1}^{T}(y_t - \boldsymbol{\mu}))^2$.
- What if the true mean of y_t changes at time T_1+1, such that $\mathrm{E}[y_t] = \boldsymbol{\mu}_0$ for $t = 1, \ldots, T_1$, and $\mathrm{E}[y_t] = \boldsymbol{\mu}_1$ for $t = T_1+1, \ldots, T$, where $\boldsymbol{\mu}_0 < \boldsymbol{\mu}_1$?
- The moment condition evaluated at time t provides a signal about the direction in which to adjust $\hat{\boldsymbol{\mu}}$ to locally obtain a better fit.
- We introduce GaMM dynamics for a time-varying parameter $\boldsymbol{\mu}_t$ by considering a GMM criterion function for the observation at time t only, i.e., $\mathrm{E}_{t-1}[y_t - \boldsymbol{\mu}_t]^2$.
- Taking the derivative of this objective function with respect to $\boldsymbol{\mu}_t$ and evaluating it at the tth observation rather than taking the expectation, we obtain the step

$$s_t = -2(y_t - \boldsymbol{\mu}_t). \tag{1}$$

- We use s_t to set-up autoregressive dynamics for the time-varying parameter $\boldsymbol{\mu}_t$:

$$\boldsymbol{\mu}_{t+1} = \omega \cdot (1 - B) + B\boldsymbol{\mu}_t + As_t, \tag{2}$$

where ω, A, and B are static parameters that need to be estimated.



GaMM(1,1)

$$\mathrm{E}_{t-1}\left[g_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \boldsymbol{\theta}\right)\right] = 0 \tag{3}$$

$$\{\boldsymbol{w}_t\} = \{y_t, \boldsymbol{x}_t, \boldsymbol{z}_t\} \tag{4}$$

GMM objective function at time t:

$$\mathrm{E}_{t-1}\left[g_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \boldsymbol{\theta}\right)\right]^{\intercal} \boldsymbol{\Omega}_t \, \mathrm{E}_{t-1}\left[g_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \boldsymbol{\theta}\right)\right] \tag{5}$$

We use the notion of Fréchet derivative to derive first-order conditions:

$$\boldsymbol{G}_t^{\intercal}\boldsymbol{\Omega}_t \, \mathrm{E}_{t-1}\left[g_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \boldsymbol{\theta}\right)\right] = 0, \tag{6}$$

$$\boldsymbol{G}_t = \mathrm{E}_{t-1}\left[\frac{\partial \boldsymbol{g}_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \boldsymbol{\theta}\right)}{\partial \boldsymbol{f}_t^{\intercal}}\right] \tag{7}$$

Where: $\boldsymbol{\Omega}_t = \left\{\mathrm{E}_{t-1}\left[\boldsymbol{g}_t\boldsymbol{g}_t^{\intercal}\right]\right\}^{\star}$ and $H^{\star}$ is the Moore-Penrose pseudo inverse. The nature of the problem determines whether and how $\boldsymbol{G}_t$ needs to be estimated.

Transition dynamics for $\boldsymbol{f}_t \to \boldsymbol{f}_{t+1}$

General updating step:

$$\boldsymbol{s}_t = \left(\boldsymbol{G}_t^{\intercal}\boldsymbol{\Omega}_t\boldsymbol{G}\right)^{\star}\boldsymbol{G}_t^{\intercal}\boldsymbol{\Omega}_t\boldsymbol{g}_t\left(\boldsymbol{w}_t; \boldsymbol{f}_t, \theta\right) \tag{8}$$

Time-varying parameter recursion:

$$\boldsymbol{f}_{t+1} = \boldsymbol{B}\left(\boldsymbol{f}_t - \boldsymbol{\omega}\right) + \boldsymbol{\omega} + \boldsymbol{A}\boldsymbol{s}_t \tag{9}$$

GaMM instruments:

$$\bar{\boldsymbol{g}}_t = \boldsymbol{g}_t\left(\boldsymbol{w}_t; [\boldsymbol{\theta}, \boldsymbol{f}_t]\right) \otimes \left[1 \;\; \boldsymbol{s}_{t-1} \;\; \boldsymbol{f}_{t-1}\right]^{\intercal} \tag{10}$$

$$\bar{\boldsymbol{g}} = \frac{1}{T}\sum_{t=1}^{T}\bar{\boldsymbol{g}}_t \approx \mathrm{E}\left[\bar{g}_t(\cdot)\right] = 0 \tag{11}$$

Criterion function, where $\boldsymbol{\theta}$ corresponds to all static parameters, including $\boldsymbol{\omega}$, $\boldsymbol{A}$, $\boldsymbol{B}$:

$$\min_{\boldsymbol{\theta}\in\Theta} \bar{\boldsymbol{g}}^{\intercal}\bar{\Omega}\bar{\boldsymbol{g}} \tag{12}$$

Asymptotic theory

GaMM falls entirely within the standard framework of GMM estimation. Therefore, consistency and asymptotic normality results follow from Hansen (1982), including an expression for the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$ under standard high-level regularity conditions.

$$n^{1/2}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} \mathrm{N}\left(0, \boldsymbol{H}^{-1}\boldsymbol{D}\boldsymbol{H}^{-1}\right), \qquad \bar{\boldsymbol{D}} = \bar{\boldsymbol{G}}^{\intercal}\bar{\Omega}\bar{\boldsymbol{V}}\bar{\Omega}\bar{\boldsymbol{G}}. \tag{13}$$

The efficient weighting matrix is $\bar{\Omega} = \bar{\boldsymbol{V}}^{-1}$, in which case the asymptotic covariance matrix collapses to $(\bar{\boldsymbol{G}}^{\intercal}\bar{\Omega}\bar{\boldsymbol{G}})^{-1}$.

Application: time-varying risk aversion in CCAPM

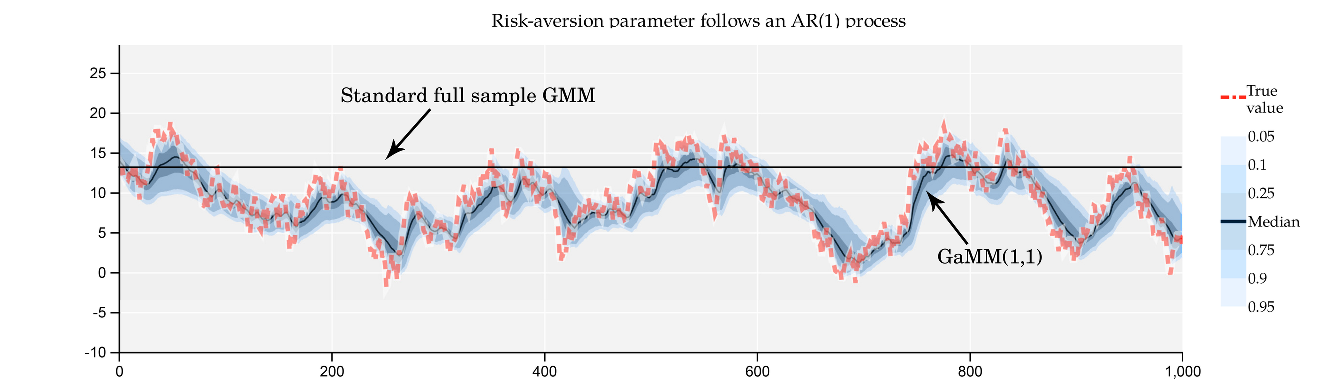
Agents have standard CRRA utility function. We assume that shocks to risk-aversion are exogenous and that agents are myopic in the sense that they consider γ_t to remain fixed forever when making their decision at time t. This results in the Euler equation:

$$\mathrm{E}_t\left[\beta\left(C_{t+1}/C_t\right)^{-\gamma_t} R_{t+1}^{x}\right] = 1. \tag{14}$$

We can directly employ the GaMM framework by taking $\boldsymbol{f}_t = \gamma_t$.

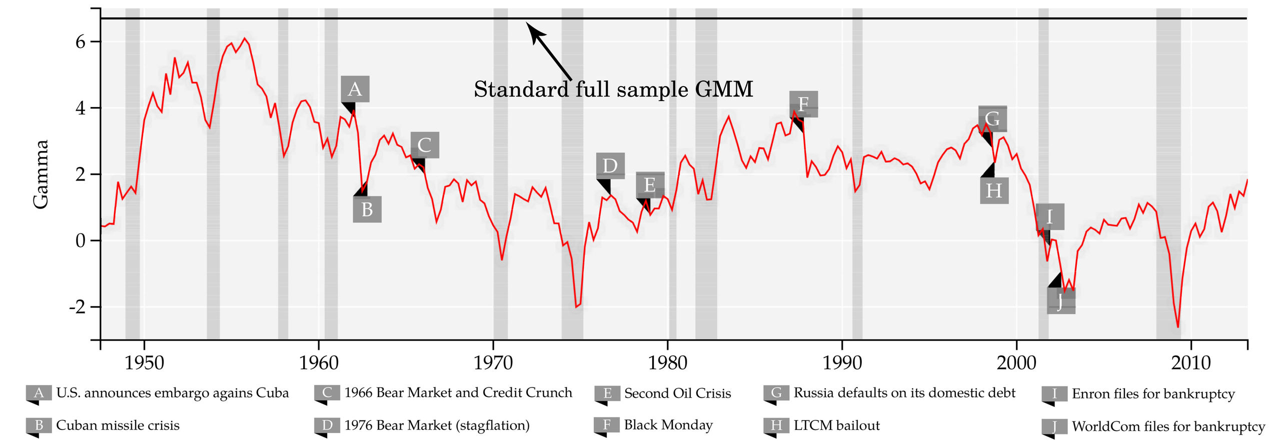
Simulation results

If there is time-variation in γ_t, the full sample GMM estimates are severely biased towards the high end realizations of the parameter. In contrast, the filtered path $\hat{\gamma}_t$ obtained with GaMM tends to follow the true path closely.



Empirical results for U.S. quaterly data (1947–2013)

The are two components in the time-variation of γ_t. A short-term cycle follows the business cycle, agents adjust their consumption slowly and with a delay compared to the reaction of financial markets. The long-term pattern shows a continuous decrease in risk aversion since the 1950s. High risk-aversion estimated with GMM could be due to early observations having high leverage.



Also in the paper...

PDF @ zmks.co/gamm

- Optimal instruments.
- Two additional applications:
  1. Time-varying scale in stable distributions.
     GaMM estimates the time-varying scale and remaining static parameters without bias.
  2. Endogenous regressor problem.
     We benchmark GaMM against Kalman Filter, the Kalman Filter results are biased but with low sampling variability, whereas paths obtained with GaMM are unbiased on average, but at the cost of a higher sampling variance.
- Penalized criterion function which further reduces RMSE.