As we know, There are known knowns. There are things we know we know. We also know There are known unknowns. That is to say, We know there are some things We do not know. But there are also unknown unknowns, The ones we don't know We don't know. Donald Rumsfeld, U.S. Secretary of Defence

Suppose that

$$y_{it} = m_i(t/T) + u_{it},$$

where $i = 1, \ldots, n$ and $t = 1, \ldots, T$; both n and T are large. We suppose that there exists J classes of functions

$$\mathcal{G}_j = \{f : f(u) = cg_j((u-a)/b),\ a, b, c \in \mathbb{R}_+,\ g_j \text{ a density}\}.$$

We suppose that for any i, $m_i(\cdot) \in \mathcal{G}_j$ for some j, that is, for some j there exists a_i, b_i, c_i with

$$m_i(u) = c_i g_j((u - a_i)/b_i)/b_i$$

for all $u \in [0, 1]$.

Estimation. Suppose that there is only one group with unknown $g(.)$. Given unrestricted estimates $\widehat{m}_i(\cdot)$, we can estimate c_i by $\int_0^1 \widehat{m}_i(u)du$ and work with the ratio $\widehat{m}_i^*(u) = \widehat{m}_i(u)/\int_0^1 \widehat{m}_i(u)du$. We may make different assumptions here about a, b. For example, a is the mean and b is the standard deviation of the density g. In that case we can estimate a_i by $\int_0^1 u\widehat{m}_i^*(u)du$ and b_i by the square root of $\int_0^1 u^2\widehat{m}_i^*(u)du - (\int_0^1 u\widehat{m}_i^*(u)du)^2$. Alternatively, median and interquartile range work. Then one can estimate $g(u)$ by

$$\frac{1}{n}\sum_{i=1}^{n} \widehat{b}_i \widehat{m}_i^* \left(u\widehat{b}_i + \frac{\widehat{a}_i}{\widehat{b}_i} \right).$$

Now suppose that there are multiple $g's$. The recovery of the constants a, b, c only uses the individual regression function. But now the uncertainty is around which i to average over. This can be addressed by the clustering algorithms.

The parameter estimates $\widehat{a}_i, \widehat{b}_i, \widehat{c}_i$ are $\sqrt{T}$ consistent, whereas the estimates of g_j will be $\sqrt{nTh}$ consistent.