# A Local Generalized Method of Moments Estimator

Arthur Lewbel*
Boston College

June 2006

**Abstract**

A local Generalized Method of Moments Estimator is proposed for nonparametrically estimating unknown functions that are defined by conditional moment restrictions.

JEL Codes: C14, C13. Keywords: Generalized Method of Moments, Nonparametric Estimation, Semiparametric Estimation, Conditional Moments.

Assume we observe a sample of observations of a random vector $W$. Let $Z$ and $X$ be random subvectors of $W$, so each contains some or all of the elements of $W$. Hansen's (1982) Generalized Method of Moments (GMM) estimates a vector of parameters $\theta$ from unconditional moments $E\left[g(\theta_0, W)\right] = 0$ given a known vector valued function $g$. Asymptotically efficient estimates of $\theta$ given conditional moments $E[g(\theta, W) \mid Z = z)] = 0$ can be obtained using variants of GMM or empirical likelihood such as Newey (1993), Donald, Imbens, and Newey (2003), Otsu (2003), or Dominguez and Lobato (2004).

Suppose that, in place of a parameter vector $\theta$, the goal is estimation of a vector of unknown functions $q(X)$ where $E[g(q_0(X), W) \mid Z = z)] = 0$. Resulting nonparametric conditional moment estimators include Carrasco and Florens (2000), Newey and Powell (2003), and Ai and Chen (2003). These estimators require verification of relatively elaborate sets of regularity conditions.

*Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467 USA. Tel: (617)-552-3678 fax: (617)-552-2308 email: lewbel@bc.edu url: http://www2.bc.edu/~lewbel

This paper focuses on the special case of the general nonparametric conditional moment problem where $X$ and $Z$ are the same, so we wish to estimate unknown functions $q(Z)$ where

$$E[g(q_0(Z), W) \mid Z = z)] = 0. \tag{1}$$

While estimators such as Ai and Chen (2003) and other can be applied in this case, the local GMM estimator proposed here exploits the assumption that the unknown functions $q$ depend on the same variables as the conditional moments, resulting in a short, simple set of regularity conditions. This local GMM estimator is straighforward to implement and is a natural extension of ordinary GMM estimation of the same model when $Z$ is discrete.

An example application is the nonparametric probit model $W = I[q_0(Z)+e \geq 0]$, where $I()$ denotes the indicator function that equals one if its argument is true and zero otherwise, $q_0(z)$ is an unknown function to be estimated, and $e$ is a standard normal independent of $Z$, or has some other known distribution. Then estimation could be based on equation (1) where the function $g$ is $g(q, w) = w - F_e(q)$ where $F_e$ is the known cumulative distribution function of $-e$. Nonparametric censored or truncated regression would have a similar form. The local gmm estimator could also be used for estimation of Euler equations, which are mean zero conditional on information in a given time period, and may have parameters $q_0(z)$, such as preference parameters, that are unknown functions of observables. Another application is Lewbel (2006), who provides examples of a treatment model that generates conditional moments in the form of equation (1).

The local GMM estimator provided here is very similar to the local nonlinear least squares estimator of Gozalo and Linton (2000), and in fact an early draft of their paper included a GMM variant. In their applications a local version of a parametric model is estimated, which converges at rate root $n$ if the model is correct, but remains consistent as a nonparametric regression if the parametric model is misspecified (analogous to local polynomial regression, which is the special case of their estimator in which the parametric model is a polynomial). In the present paper no parametric model is proposed. Instead, the intended estimand from the outset is the unknown vector of functions $q_0(z)$.

To motivate the estimator, consider first the case where $Z$ is discretely distributed, or more specifically, that $Z$ has one or more mass points and we only wish to estimate $q_0(z)$ at those points. Let $\theta_{z0} = q_0(z)$. If the distribution of $Z$ has a mass point with positive probability at $z$, then

$$E[g(\theta_z, W) \mid Z = z] = \frac{E[g(\theta_z, W)I(Z = z)]}{E[I(Z = z)]}$$

so equation (1) holds if and only if $E[g(\theta_{z0}, W)I(Z = z)] = 0$. It therefore follows that if $\theta_{z0}$ is identified from the moment conditions (1), then we may under standard regularity conditions estimate $\theta_{z0} = q_0(z)$ by the ordinary GMM estimator

$$\widehat{\theta}_z = \arg\min_{\theta_z} \sum_{i=1}^{n} g(\theta_z, W_i)' I(Z_i = z)\boldsymbol{\Omega}_n \sum_{i=1}^{n} g(\theta_z, W_i)' I(Z_i = z) \tag{2}$$

for some sequence of positive definite $\boldsymbol{\Omega}_n$. If $\boldsymbol{\Omega}_n$ is a consistent estimator of $\boldsymbol{\Omega}_{z0} = E[g(\theta_{z0}, W)g(\theta_{z0}, W)' I(Z = z)]^{-1}$, then standard efficient GMM gives

$$\sqrt{n}(\widehat{\theta}_z - \theta_{z0}) \to^d N\left(0, \left[E\left(\frac{\partial g(\theta_{z0}, W)I(Z = z)}{\partial \theta_z{}'}\right)\boldsymbol{\Omega}_{z0} E\left(\frac{\partial g(\theta_{z0}, W)I(Z = z)}{\partial \theta_z{}'}\right)'\right]^{-1}\right)$$

Continue to assume that equation (1) holds and identifies $q_0(z)$ where $g$ is known and $q_0$ is unknown, but now assume that $Z$ is continuously distributed. The proposed local GMM estimator consists of applying equation (2) to this case of continuous $Z$ by replacing averaging over just observations $Z_i = z$ with local averaging over observations $Z_i$ in the neighborhood of $z$.

Assumption A1. Let $Z_i, W_i, i = 1, ..., n$, be an independently, identically distributed random sample of observations of the random vectors $Z, W$. The $d$ vector $Z$ is continuously distributed with density function $f(Z)$. For given point $z$ in the interior of $supp(Z)$ having $f(z) > 0$ and a given vector valued function $g(q, w)$ where $g(q(z), w)$ is twice differentiable in the vector $q(z)$ for all $q(z)$ in some compact set $\Theta(z)$, there exists a unique $q_0(z) \in \Theta(z)$ such that $E[g(q_0(z), W) \mid Z = z] = 0$. Let $\boldsymbol{\Omega}_n$ be a finite positive definite matrix for all $n$, as is $\boldsymbol{\Omega} = plim_{n\to\infty}\boldsymbol{\Omega}_n$.

Assumption A1 provides the required moment condition structure and identification for the model, and Assumption A2 below provides conditions for local averaging. Define $\varepsilon[q(z), W]$, $V(z)$, and $R(z)$ by

$$\begin{aligned} \varepsilon[q(z), W] &= g(q(z), W)f(z) - E[g(q(z), W)f(Z) \mid Z = z] \\ V(z) &= E\left[\varepsilon(q_0(z), W)\varepsilon(q_0(z), W)^T \mid Z = z\right] \\ R(z) &= E\left(\frac{\partial g[q_0(z), W]}{\partial q_0(z)^T} f(Z) \mid Z = z\right) \end{aligned}$$

Assumption A2. Let $\eta$ be some constant greater than 2. Let $K$ be a nonnegative symmetric kernel function satisfying $\int K(u)du = 1$ and $\int ||K(u)||^{\eta}du$ is finite. For all $q(z) \in \Theta(z)$, $E[||g(q(z), W)f(Z)||^{\eta} \mid$

$Z = z]$, $V(z)$, $R(z)$, and $Var[[\partial g(q(z), W)/\partial q(z)]f(Z) \mid Z = z]$ are finite and continuous at $z$ and $E[g(q(z), W)f(Z) \mid Z = z]$ is finite and twice continuously differentiable at $z$.

Define

$$S_n(q(z)) = \frac{1}{nb^d}\sum_{i=1}^{n} g[q(z), W_i]K\left(\frac{z - Z_i}{b}\right)$$

where $b = b(n)$ is a bandwidth parameter. The proposed local GMM estimator is

$$\widehat{q}(z) = \arg\inf_{q(z)\in\Theta(z)} S_n(q(z))^T \Omega_n S_n(q(z)) \quad (3)$$

THEOREM 1: Given Assumptions A1 and A2, if the bandwidth $b$ satisfies $nb^{d+4} \to 0$ and $nb^d \to \infty$, then $\widehat{q}(z)$ is a consistent estimator of $q_0(z)$ with limiting distribution

$$(nb)^{1/2}[\widehat{q}(z) - q_0(z)] \to^d N\left[0, (R(z)^T\Omega R(z))^{-1}R(z)^T\Omega V(z)\Omega R(z)(R(z)^T\Omega R(z))^{-1}\int K(u)^2du]\right]$$

Theorem 1 assumes a bandwidth rate that makes bias shrink faster than variance, and so is not mean square optimal. One could instead choose the mean square optimal rate where $nb^{d+4}$ goes to a constant, but the resulting bias term would then have a complicated form that depends on the kernel regression biases in both $S_n(q_0(z))$ and its derivative with respect to $q_0(z)$, among other terms.

Applying the standard two step GMM procedure, we may first estimate $\widetilde{q}(z) = \arg\inf_{q(z)\in\Theta(z)} S_n(q(z))^T S_n(q(z))$, then let $\Omega_n$ be the inverse of the sample variance of $S_n(\widetilde{q}(z))$ to get $\Omega = V(z)^{-1}$, making

$$(nb)^{1/2}[\widehat{q}(z) - q_0(z)] \to^d N\left[0, (R(z)^T\Omega R(z))^{-1}\int K(u)^2du]\right]$$

where $R(z)$ can be estimated using

$$R_n(z) = \frac{1}{nb^d}\sum_{i=1}^{n} \frac{\partial g[\widehat{q}(z), W_i]}{\partial\widehat{q}(z)^T}K\left(\frac{z - Z_i}{b}\right)$$

At the expense of some additional notation, the two estimators (2) and (3) can be combined to handle $Z$ containing both discrete and continuous elements, by replacing the kernel function in $S_n$ with the product of a kernel over the continuous elements and an indicator function for the discrete elements.

PROOF OF THEOREM 1: Define

$$\begin{aligned} S_n'(q(z)) &= \frac{\partial S_n(q(z))}{\partial q(z)^T} = \frac{1}{nb^d}\sum_{i=1}^{n}\frac{\partial g[q(z),W_i]}{\partial q(z)^T}K\left(\frac{z-Z_i}{b}\right) \\ Q_n(q(z)) &= S_n(q(z))^T\Omega_n S_n(q(z)) \end{aligned}$$

Let $S_0(q(z)) = plim_{n\to\infty}S_n(q(z))$ and similarly for $S_n'$ and $Q_n$. Assumptions A1 and A2 give sufficient conditions for consistency of these kernel estimators, so these probability limits exist and

$$\begin{aligned} S_0(q(z)) &= E[g(q(z),W)f(z) \mid Z=z] \\ Q_0(q(z)) &= S_0(q(z))^T\Omega S_0(q(z)). \end{aligned}$$

Now consider consistency of $\widehat{q}(z)$. We have pointwise convergence of $S_n(q(z))$ to $S_0(q(z))$and compactness of $\Theta(z)$. It is also the case that $|S_n'(q(z))| = O_p(1)$, since $|S_n'(q(z))|$ is a kernel estimator, and standard conditions have been provided for its consistency, that is, $plim|S_n'(q(z))| = E[|\partial g(q(z),W)/\partial q(z)|f(Z) \mid Z=z]$. This suffices for stochastic equicontinuity, and therefore we have the uniform convergence

$$plim \sup_{q(z)\in\Theta(z)} |S_n(q(z)) - S_0(q(z))| = 0.$$

It follows that $Q_n(q(z))$ also converges uniformly to $Q_0(q(z))$. The assumptions provide compactness of $\Theta(z)$ and imply continuity of $Q_0(q)$. The quadratic form of $Q_0$ is uniquely maximized at $S_0(q_0(z)) = 0$ and hence at $q(z) = q_0(z)$, so the standard conditions for consistency $plim\widehat{q}(z) = q_0(z)$ are satisfied.

For the limiting distribution, Taylor expanding the first order conditions as in standard GMM gives

$$S_n'(\widehat{q}(z))^T\Omega_n\left[S_n(q_0(z)) + S_n'(\widetilde{q}(z))(\widehat{q}(z) - q_0(z))\right] = 0$$

where $\widetilde{q}(z)$ lies between $\widehat{q}(z)$ and $q_0(z)$. By consistency of $\widehat{q}$, the uniform convergence of $S_n$, and using $R(z) = S_0'(q_0(z))$, this simplifies to

$$R(z)^T\Omega\left[S_n(q_0(z)) + R(z)(\widehat{q}(z) - q_0(z))\right] = o_p(1)$$

Solving for $\widehat{q}(z) - q_0(z)$ and multiplying by $(nb)^{1/2}$ gives

$$(nb)^{1/2}(\widehat{q}(z) - q_0(z)) = (R(z)^T\Omega R(z))^{-1}R(z)^T\Omega(nb)^{1/2}S_n(q_0(z)) + o_p((nb)^{1/2}).$$

Now $S_0(q_0(z)) = 0$ and standard kernel regression limiting distribution theory gives

$$(nb)^{1/2}S_n(q_0(z)) \to^d N[0, V(z)\int K(u)^2 du]$$

and the theorem follows.

**References**

Ai, C. and X. Chen (2003), "Efficient Estimation of Models With Conditional Moment Restrictions Containing Unknown Functions," Econometrica, 71, 1795-1844.

Carrasco, M. and J. P. Florens (2000), "Generalization of GMM to a Continuum of Moment Conditions," Econometric Theory, 16, 797-834.

Dominguez, M. and I. Lobato, (2004), "Consistent Estimation of Models Defined by Conditional Moment Restrictions," Econometrica, 72, 1601-1615.

Donald, S. G., G. W. Imbens, and W. K. Newey (2003), "Empirical Likelihood Estimation and Consistent Tests With Conditional Moment Restrictions," Journal of Econometrics, 117, 55-93.

Florens, J.-P. and L. Malavolti (2003), "Instrumental Regression with Discrete Variables," unpublished manuscript..

Hansen, L., (1982), "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50, 1029-1054.

Lewbel, A., (2006), "Estimation of Average Treatment Effects With Misclassification," Econometrica, forthcoming.

Newey, W. K. (1993), "Efficient Estimation of Models With Conditional Moment Restrictions," in Handbook of Statistics, vol. 11, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod, Amsterdam: North Holland, chapter 16.

Newey, W. K. and J. L. Powell, (2003), "Instrumental Variable Estimation of Nonparametric Models," Econometrica, 71 1565-1578.

Otsu, T., (2003), "Penalized Empirical Likelihood Estimation of Conditional Moment Restriction Models with Unknown Functions," unpublished manuscript.