Econometrics 710
Midterm Exam
March 12, 2013
Sample Answers

This exam concerns the model

$$\begin{aligned} y_i &= m(x_i) + e_i & (1)\\ m(x) &= \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_p x^p & (2)\\ E(z_i e_i) &= 0 & (3)\\ z_i &= (1, x_i, ..., x_i^p)' & (4)\\ g(x) &= \frac{d}{dx} m(x) & (5) \end{aligned}$$

with iid observations (y_i, x_i), $i = 1, ..., n$. The order of the polynomial p is known.

1. How should we interpret the function $m(x)$ given the projection assumption (3)? How should we interpret $g(x)$? (Briefly)

   The model does not specify that $m(x)$ is the conditional mean. Rather, equation (3) specifies that it is a projection model. Thus $m(x)$ is the best linear predictor of y_i given linear functions of z_i. Equivalently, it is the best predictor in the class of p^{th} order polynomials in x_i. It is also the best mean-square approximation to the conditional mean, in the class of p^{th} order polynomials in x_i. The function $g(x)$ is the derivative of the best linear predictor, and equals

   $$\begin{aligned} g(x) &= \beta_1 + 2\beta_2 x + 3\beta_3 x^2 + \cdots + p\beta_p x^{p-1} \\ &= h(x)'\beta \end{aligned}$$

   where $\beta = (\beta_0, ..., \beta_p)'$ and $h(x) = (0, 1, 2x, 3x^2, ..., px^{p-1})'$.

2. Describe an estimator $\hat{g}(x)$ of $g(x)$.

   Since $g(x) = h(x)'\beta$ is linear in β, the plug-in approach suggests replacing β with the efficient estimator for β. Under the projection assumption (3) OLS is the asymptotically efficient estimator. It equals $\hat{\beta} = (Z'Z)^{-1}(Z'Y)$ where Y and Z are the stacked observations on y_i and z_i. Then the estimator for $g(x)$ is

   $$\begin{aligned} \hat{g}(x) &= h(x)'\hat{\beta} \\ &= \hat{\beta}_1 + 2\hat{\beta}_2 x + 3\hat{\beta}_3 x^2 + \cdots + p\hat{\beta}_p x^{p-1} \end{aligned}$$

3. Find the asymptotic distribution of $\sqrt{n}(\hat{g}(x) - g(x))$ as $n \to \infty$.

   Under the projection assumption (3) plus regularity conditions, we know that as $n \to \infty$, $\sqrt{n}\left(\hat{\beta} - \beta\right) \to_d N(0, V_\beta)$ where $V_b = Q^{-1}\Omega Q^{-1}$ with $Q = E(z_i z_i')$ and $\Omega = E\left(z_i z_i' e_i^2\right)$. Then as $n \to \infty$

   $$\begin{aligned} \sqrt{n}(\hat{g}(x) - g(x)) &= \sqrt{n}\left(h(x)'\hat{\beta} - h(x)'\beta\right) \\ &= h(x)'\sqrt{n}\left(\hat{\beta} - \beta\right) \end{aligned}$$

   $$\to_d h(x)'N(0, V_\beta) = N(0, h(x)'V_\beta h(x)).$$

4. Show how to construct an asymptotic 95% confidence interval for $g(x)$.

We estimate V_β with

$$\begin{aligned} \hat{V}_\beta &= \hat{Q}^{-1}\hat{\Omega}\hat{Q}^{-1} \\ \hat{Q} &= \frac{1}{n}\sum_{i=1}^n z_i z_i' \\ \hat{\Omega} &= \frac{1}{n}\sum_{i=1}^n z_i z_i' \hat{e}_i^2/(1-h_{ii})^2 \\ h_{ii} &= z_i'(Z'Z)^{-1}z_i \end{aligned}$$

[Other estimates for $\hat{\Omega}$ can be used.] An asymptotic standard error for $\hat{g}(x)$ is $\hat{s}(x) = n^{-1/2}\left(h(x)'\hat{V}_\beta h(x)\right)$.

An asymptotic 95% confidence interval for $g(x)$ is $\hat{g}(x) \pm 2n^{-1/2}\left(h(x)'\hat{V}_\beta h(x)\right)$.

The justification is that $(\hat{g}(x) - g(x))/\hat{s}(x) \to_d N(0,1)$ as $n \to \infty$, so $\Pr\left(|\hat{g}(x) - g(x)|/\hat{s}(x) \le 1.96\right) \to 0.95$

5. Assume $p = 2$. Describe how to estimate $g(x)$ imposing the constraint that $m(x)$ is concave.

When $p = 2$ we have $m(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ and $g(x) = \beta_1 + 2\beta_2 x$. The function $m(x)$ is (weakly) concave iff $\beta_2 \le 0$. The constrained least-squares estimator of β is

$$\tilde{\beta} = \operatorname*{argmin}_{\beta:\beta_2 \le 0} SSE(\beta)$$

The solution is

$$\tilde{\beta} = \begin{cases} \hat{\beta} & \text{if} \quad \hat{\beta}_2 \le 0 \\ \bar{\beta} & \text{if} \quad \hat{\beta}_2 > 0 \end{cases}$$

where $\bar{\beta} = (\bar{\beta}_0, \bar{\beta}_1, 0)'$ with $\bar{\beta}_0, \bar{\beta}_1$ obtained by OLS of y_i on $(1, x_i)$. The constrained estimator of $g(x)$ is

$$\begin{aligned} \tilde{g}(x) &= h(x)'\tilde{\beta} \\ &= \begin{cases} \hat{\beta}_1 + 2\hat{\beta}_2 x & \text{if} \quad \hat{\beta}_2 \le 0 \\ \bar{\beta}_1 & \text{if} \quad \hat{\beta}_2 > 0 \end{cases} \end{aligned}$$

This is a common and important problem in applications. If a theory implies that a function is concave, it may be desirable to impose that condition on estimates.

6. Assume $p = 2$. Describe how to estimate $g(x)$ imposing the constraint that $m(u)$ is increasing on the region $u \in [x_L, x_U]$.

There was a typo. I had meant the question listed above, to impose that $m(u)$ is increasing (monotonic), not $g(u) \ge 0$ is increasing. The latter is more complicated to impose, and made it a tricky question. In constrast, my intention was the simpler problem that $m(u)$ is increasing. Now since $m(x)$ is a quadratic, it cannot be globally increasing unless $\beta_2 = 0$, so it does not make practical sense to impose global monotonicity. Instead we might want to impose monotonicity over a range of interest $[x_L, x_U]$, perhaps the support of x. The function $m(u)$ is (weakly) increasing iff $g(u) \ge 0$ [This is the source of the typo.] Since $g(u) = \beta_1 + 2\beta_2 u$ is linear, it is positive on $[x_L, x_U]$ iff it is positive at the endpoints, that is $\beta_1 + 2\beta_2 x_L \ge 0$ and $\beta_1 + 2\beta_2 x_U \ge 0$. The constrained estimator solves

$$\tilde{\beta} = \operatorname*{argmin}_{\beta:\beta_1 + 2\beta_2 x_L \ge 0, \beta_1 + 2\beta_2 x_U \ge 0} SSE(\beta)$$

If the two constraints are not binding, then $\tilde{\beta} = \hat{\beta}$ and $\tilde{g}(x) = \hat{\beta}_1 + 2\hat{\beta}_2 x$. Otherwise, the minimum lies on the boundary of the set described by the inequalities $\beta_1 + 2\beta_2 x_L \ge 0, \beta_1 + 2\beta_2 x_U \ge 0$. Since there are two constraints there is no simple solution, but the answer can be found by quadratic programming.