Problem Set # 3

Econ 722

Spring, 2018

Instructions: Answer each of the following. All solutions must be submitted electronically on Canvas by 11:59pm on Thursday, April 5th. Late problem sets will not be accepted: it is much better to turn in partial solutions rather than nothing at all. You may discuss these problems with your classmates, but if you work together please list the names of the students with whom you have collaborated at the top of your solutions.

1. This question concerns the so-called "GMM-AIC" moment selection criterion from Andrews (1999) which takes the form

$$\text{GMM-AIC}(c) = J_T(c) - 2(|c| - p)$$

   where $J_T(c)$ denotes the J-test statistic for the estimator based on the collection of moment restrictions indexed by c, $|c|$ denotes the number of moment conditions in specification c and p denotes the number of parameters. Characterize the asymptotic behavior of this criterion under the assumptions of the consistency theorem we proved in class. Hint: there are two cases, which parallel our proof from class.

2. The FMSC of DiTraglia (2016) is a moment selection criterion constructed by deriving the asymptotic MSE of an estimator of some "target parameter" μ, under local misspecification. A very similar idea can also be used for *model selection* in maximum likelihood models: this is the so-called "Focused Information Criterion" of Claeskens & Hjort (2003). In this question you will derive the simplest possible example of the FIC. This will require you to "get your hands dirty" with local asymptotics, so you may want to read the beginning of Chapter 4 from the lecture notes before attempting this problem. Consider a linear regression model with two regressors x and z

$$y_t = \theta x_t + \gamma z_t + \epsilon_t$$

where $\{(x_t, z_t, \epsilon_t)\}_{t=1}^{T} \sim$ iid with means $(0,0,0)$ and variances $(\sigma_x^2, \sigma_z^2, \sigma_\epsilon^2)$. For simplicity, assume the errors are homoskedastic. Our goal is to estimate θ with minimum MSE, and the model selection decision is whether or not to include z in the regression. Consider two estimators of θ: the "long" regression estimator $\widehat{\theta}$ calculated from $(\widehat{\theta}, \widehat{\gamma})' = \left\{[\mathbf{x}, \mathbf{z}]' [\mathbf{x}, \mathbf{z}]\right\}^{-1} [\mathbf{x}, \mathbf{z}]' \mathbf{y}$ and the "short" regression estimator $\widetilde{\theta} = (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\mathbf{y}$. Since all random variables are mean zero, you do not have to include a constant.

(a) Suppose that γ is local to zero, in other words $\gamma = \delta/\sqrt{T}$. Under this assumption, derive the asymptotic distributions of $\sqrt{T}(\widetilde{\theta} - \theta)$ and

$$\sqrt{T} \begin{bmatrix} \widehat{\theta} - \theta \\ \widehat{\gamma} - 0 \end{bmatrix}.$$

Note that the limit distribution of $\widehat{\gamma}$ is centered around zero since $\delta/\sqrt{T} \to 0$ as $T \to \infty$. You should find that $\widetilde{\theta}$ has an asymptotic bias that depends on δ.

(b) Under what conditions does $\widetilde{\theta}$ have a lower AMSE than $\widehat{\theta}$? Note that your answer should depend on δ. Explain the intuition for your result.

(c) Propose an asymptotically unbiased estimator of δ constructed from $\sqrt{T}\widehat{\gamma}$.

(d) Combine steps (a) and (c) to propose asymptotically unbiased estimators of the AMSE of $\widehat{\theta}$ and $\widetilde{\theta}$.

(e) The FIC chooses the estimator with the lower estimated AMSE from step (d). How does this rule compare to AIC, BIC, Mallow's C_p, and a t-test of the null hypothesis $H: \gamma = 0$ at the $\alpha \times 100\%$ level? Comment briefly on any relationships you uncover.