

Shanghai University of Finance and Economics
School of Statistics and Management
Financial Econometrics, 05/2020, Mr. Wang Xunxiao

Computer-based Exam

Notes:

- All tests are based on the 5% significance level
- **Do not directly hand in computer output. Summarize the output. Enhance statistical and economic interpretations of your results.**

1. 数据集hprice2a.dta包含了美国波士顿506个社区的房屋中位数价格的横截面数据。考虑如下的“特征价格回归”（Hedonic Price Regression, 即认为房价由房屋性能所决定）:

$$lprice_i = \beta_1 + \beta_2 lnox_i + \beta_3 ldist_i + \beta_4 rooms_i + \beta_5 stratio_i + \varepsilon_i$$

其中, $lprice$ 为房价的对数, $lnox$ 为空气污染程度的对数, $ldist$ 为社区到就业中心的距离, $rooms$ 为房屋的平均房间数, $stratio$ 为社区学校的学生——教师比例, 下标 i 表示“第 i 个社区”。

- (1) 使用稳健标准误进行回归, 并分别检验“ $H_0: \beta_2 = \beta_5$ ”与“ $H_0: \beta_4 = 0.33$ ”。
- (2) 以5%的置信度, 使用还特检验, 检验是否存在异方差。
- (3) 以5%的置信度, 使用BP检验, 检验是否存在异方差(假设扰动项为独立同分布, 分别以拟合值 \hat{y} 以及所有解释变量进行检验)。
- (4) 假设条件异方差仅依赖于拟合值 \hat{y} (不含常数项), 进行WLS估计。

2. 使用数据集hprice2a.dta, 对下面的回归方程进行如下操作。

$$lprice_i = \beta_1 + \beta_2 lnox_i + \beta_3 ldist_i + \beta_4 rooms_i + \beta_5 stratio_i + \varepsilon_i$$

- (1) 进行RESET检验(分别使用拟合值 \hat{y} 与rhs解释变量)
- (2) 计算信息准则AIC和BIC
- (3) 计算所有解释变量的方差膨胀因子(VIF)。存在严重多重共线性吗?
- (4) 计算观测数据的lev的最大值与平均值。最大值是平均值的几倍? 列出影响力最大的前3个数据。

3. 使用数据集gasoline.dta估计美国1953-2004年的汽油需求函数。考虑如下回归:

$$lgasq_t = \beta_1 + \beta_2 lincome_t + \beta_3 lgasp_t + \beta_4 lpnc_t + \beta_5 lpuc_t + \varepsilon_t$$

其中, 被解释变量 $lgasq$ 为人均汽油消费量的对数, 解释变量 $lincome$ 为人均收入的对数, $lgasp$ 为汽油价格指数的对数, $lpnc$ 为新车价格指数的对数, $lpuc$ 为二手车价格指数的对数。

- (1) 使用OLS估计模型。评论各变量系数的符号、显著性与经济意义。

(2) 计算残差，并记为 e_1 。将残差与其一阶滞后的散点图与线性拟合图画在一起。根据此图，是否可能存在自相关？

(3) 画残差的自相关图。

(4) 计算DW统计量，判断扰动项是否存在自相关性

(5) 用广义差分法估计模型，并比较(1)与(5)的结果。

4. 使用数据集consumption_china.dta估计一个有关 (c, y) 的VAR模型，画脉冲响应图，并进行格兰杰因果检验（究竟是消费拉动增长，还是钱多了才消费）。

5. 考虑数据集pe.dta，其中变量logpe为1871-2002年美国标准普尔股指（S&P）的市盈率，用ARIMA模型对其建模。

6. 面板数据集mus08psidextract.dta包含595名美国工人1976-1982年的一下变量：lwage（工资对数），ed（教育年限），exp（工龄），exp2（exp的平方），wks（工作周数，weeks worked）。考虑以下模型：

$$lwage_{it} = \alpha + \beta_1 exp_{it} + \beta_2 exp2_{it} + \beta_3 wks_{it} + \beta_4 ed_{it} + u_i + \varepsilon_{it}$$

(1) 进行混合回归，并对各解释变量的作用进行分析。

(2) 考虑固定效应模型并进行估计。

(3) 考虑随机效应模型并进行估计。

(4) 综合比较上述三类模型。