

本人学号\_\_\_\_\_

批改人学号\_\_\_\_\_

成绩\_\_\_\_\_

**特注:**

- (1) 你可以与同学共同讨论, 但最后必须亲自完成作业。如发现抄袭, 将以零分计。
- (2) 如果作业中写了姓名, 则该作业无效。
- (3) 只接受纸质版(回归结果须打印), 不接受电子版。手写时请注意书写端正。
- (4) 为了公平起见, 如果逾期未交作业, 将不再接受补交作业。
- (5) **请装订作业, 或在每页都写上学号, 以避免散落遗失。**

**Problem Set 3**

1. 使用数据集 `loanapp.dta` 考察美国的按揭贷款(mortgage loan)是否存在种族歧视。如果申请人的按揭贷款获批, 则被解释变量 `approve` 取值为 1; 反之, `approve` 取值为 0。主要解释变量为 `white`(是否白人)。数据集中的其他种族为 `black`(是否黑人)与 `hispan`(是否拉丁裔)。本题统一使用稳健标准误。

- (1) 把 `approve` 对 `white` 进行 OLS 回归。变量 `white` 的系数是否显著? 此效应有多大?
- (2) 把 `approve` 对 `white` 进行 Probit 回归。此回归结果与线性概率模型有何不同?
- (3) 根据(1)与(2)的回归结果, 是否可认为美国的按揭贷款市场对非白人(nonwhite)存在歧视? 为什么?
- (4) 加入控制变量 `hrat`(房贷占总收入比例), `obrat`(其他债务支出占总收入比例), `loanprc`(贷款额占房价比例), `unem`(所在行业的失业率), `male`(是否男性), `married`(是否已婚), `dep`(家属人数), `sch`(是否受过 12 年及以上教育), `cosign`(是否有担保人), `chist`(1=怠账未及 60 天, 0=怠账 60 天及以上), `pubrec`(是否曾申请破产), `mortlat1`(有 1-2 次逾期付款), `mortlat2`(有 2 次以上逾期付款), 以及 `vr`(所在小区的空置率是否高于平均值), 再次进行 Probit 回归。是否存在歧视非白人的统计证据?
- (5) 使用 Logit, 重复(4)的回归。Logit 模型的系数显著性是否与 Probit 模型相同?
- (6) 使用 Logit, 重复(4)的回归, 但汇报几率比。在给定其他控制变量的情况下, 比较白人与非白人成功申请贷款的几率比。

2. 数据集 `munnell.dta` 包含美国 48 个州, 1970-1986 年的年度数据。为了估计公共资本对经济增长的贡献, 使用此数据集进行以下回归:

$$\ln y_{it} = \beta_0 + \beta_1 \ln k_{1,it} + \beta_2 \ln k_{2,it} + \beta_3 \ln labor_{it} + \beta_4 unemp_{it} + u_i + \varepsilon_{it}$$

其中,  $y$  为州产值(gross state product),  $k_1$  为公共资本(包括高速公路、街道、供水、下水道及其他公共建筑),  $k_2$  为私人资本存量(private capital stock),  $labor$  为非农劳动力,  $unemp$  为州失业率(反映影响产出的经济周期因素)。面板变量为 `state` (州), 而时间变量为 `year` (年份)。

- (1) 进行混合回归, 评论 `lnk1` 的系数符号、显著性与经济意义。
- (2) 对随机效应模型进行 FGLS 估计。`lnk1` 的系数符号与显著性是否有变化? 检验是否存在个体随机效应。
- (3) 对随机效应模型进行 MLE 估计。
- (4) 对固定效应模型进行组内估计。`lnk1` 的系数符号与显著性是否有变化?
- (5) 对固定效应模型进行 LSDV 估计。检验是否存在个体固定效应。
- (6) 进行传统的豪斯曼检验。

- (7) 进行稳健的豪斯曼检验。
- (8) 在组内估计中，加入时间趋势项。时间趋势项是否显著？
- (9) 在组内估计中，加入时间虚拟变量，估计双向固定效应模型。时间效应是否显著？
- (10) 计算组间估计量。此估计量是否可信？

3. 数据集 `phillips.dta` 包含美国 1948-2003 年的失业率(`unem`)、通胀率(`inf`)以及通胀率差分(`dinf`)的年度数据。本题统一使用异方差稳健的标准误。

- (1) 为了预测通胀率的变化，估计有关 `dinf` 的  $AR(p)$  模型。使用序贯  $t$  规则确定滞后阶数  $p$ 。
- (2) 根据菲利普斯曲线，通胀率的变化与失业率有关。为此，将失业率(`unem`)的滞后项加入自回归模型，估计  $ADL(p, q)$  模型。使用序贯  $t$  规则确定滞后阶数( $p, q$ )。
- (3) 失业率的滞后是否有助于预测通胀率的变化？滞后失业率的系数的经济含义是什么？
- (4) 计算失业率每增加一个百分点，对于通胀率变化的短期效应与长期效应。

4. Bai and Kung (2011)使用公元前 219 年至 1839 年的时间序列数据，以每十年为观测单位，考察游牧民族进攻中原王朝的次数(`y1`)与中原王朝进攻游牧民族的次数(`y2`)之间的相互影响。使用数据集 `bai_kung.dta`，估计关于(`y1, y2`)的 VAR 模型。

- (1) 使用信息准则，确定此 VAR 模型的滞后阶数。
- (2) 估计关于(`y1, y2`)的 VAR 模型。
- (3) 游牧民族对中原王朝的进攻可能由气候冲击所驱动。为此，在此 VAR 模型引入以下外生变量：中国北方每十年有旱灾年份的比例(`x1`)，每十年有黄河决堤的比例(`x2`)，中国北方每十年有雪灾年份的比例(`w1`)，中国北方每十年有霜冻灾害年份的比例(`w2`)，平均气温(`w3`)，中原是否由游牧民族所统治(`w4, w5, w6`)<sup>1</sup>，以及时间趋势项 `w7`<sup>2</sup>。提示：使用选择项 `exo(varlist)`。
- (4) `x1` 与 `x2` 对 `y1` 的作用是否显著？评论其经济含义。
- (5) 检验 VAR 模型的各阶系数是否显著。
- (6) 检验 VAR 模型的残差是否有自相关。
- (7) 检验 VAR 模型是否平稳。
- (8) 考察 `y1` 与 `y2` 之间的格兰杰因果关系。
- (9) 画 VAR 模型的脉冲响应图。
- (10) 根据格兰杰因果关系确定变量排序，画 VAR 模型的正交化脉冲响应图。

5. 太阳黑子是否包含单位根？使用数据集 `sunspot_year.dta` 回答此问题。该数据集包含 1700-1988 年间每年太阳黑子的数目(`sunspot`)。

<sup>1</sup> `w4, w5` 与 `w6` 分别对应于 317-589 年，1126-1368 年以及 1644-1839 年三个时期。

<sup>2</sup> 时间趋势项 `w7` 取值范围为-22 至 183，对应于公元前 219 年至 1839 年的每个十年。